ISSN No:-2456-2165

Ontology-Based Modeling of the Learner in a Web Educational System: Towards Learning Analytics and Adaptive Learning

Ontology-Based Approach for Adaptive Learning

Géraud FOKOU PELAP[1]
Department of Mathematics and Computer Science
University of Dschang
Dschang, Cameroon

Louis FIPPO FITIME[2]
Department of Computer Science
The National Advanced School of Engineering
Yaoundé, Cameroon

Bery Leouro MBAIOSSOUM[3]
Department of Computer Science
University of N'Djamena
N'Djamena, Chad

Abstract:- Web based e-Education systems are an important type of information systems that have benefited from Web standards for implementation, deployment and integration. One of the challenges of these information systems is to personalize and adapt learning process according to the learners. To reach this goal we have to capture and model information about the learner, his pedagogical preferences and his learning activity. In this paper, we propose an ontology-based learner model, based on Semantic Web standards. We add this model to our models of pedagogical resources and, knowledge and skills model and we integrated it in an industrial context. We show how this global model meets industrial requirements in terms of functionalities, opening up new possibilities for learning analytics and adaptive learning.

Keywords:- *e-Education Information System, e-Education Model, Ontology, Semantic Web, Learner Profile Model, Automatic Personalization, Adaptive Learning.*

I. INTRODUCTION

e-Education research field has a wide range of applications including : educational processes management (e.g. Moodle platform), virtual classrooms, pedagogical resources (courses, exercises, etc.), regulations (e.g. official reference standards), users management (learners and teachers) and integration across different systems and actors in particular to ensure compatibility and seamless user experience [29]. e-Education systems leverage state of the art results of Information Sciences and Technologies (IST) as well as the Web architecture and resources to support them. In this context, capabilities to manage and make available through the Web the description of learners as learner profile as well as the description of learners activities are keys issues of e-Education research field. These descriptions meet capabilities of sharing and/or reusing learner profiles and activities between different platforms, between several agents/services of a same platform, and thus achieving interoperability between them.

The key role of learner profile is leading personalization of learning experience and then to fit learning process to learner based on his knowledge and skill. Thus, learning activities are created and/or enriched based on recommendations computed from learner profiles. In addition, learning activities are made from a set of pedagogical resources in order to acquire a given set of knowledge and/or skills called pedagogical objectives of these learning activities. This clearly shows an interdependence between knowledge/skills, pedagogical resources, learner profiles and learning activities. Consequently, description of learner profiles and learning activities will use or reuse description of knowledge/skills and pedagogical resources.

Our proposed learner profile model answers the following questions: (1) Which knowledge and skills have been acquired by a learner? (2) Which pedagogical resources have been used by a learner? (3) Which scores have been achieved by a learner? (4) What is the history of learning activity for the acquisition of a knowledge/skill by a learner? (5) Which learning path fit learner in order to acquire a given knowledge/skill? (6) Which pedagogical resources fit to learner capabilities? We also demonstrate the feasibility of our solution in a real industrial context. We integrate our learner profile model in the Educlever's e-Education system, and we observe that the Semantic Web-based solution meets industrial requirements, in terms of features, and allows Educlever System to address more requirements than its existing system. Moreover, our ontology-based modelling opens up new opportunities for advanced features like adaptive learning or learning trace modelling.

This paper is organized as follows: Section II presents state-of-the-art Educational ontologies and learners profiles modeling. In section III, we recall our earlier proposed model which meets public standards. Section IV presents our Semantic Web based learner profile modeling for e-Education system. Section V shows Semantic Web based integration of our learner profile model in existing e-Educlever solution and describes implementation of Educlever features based on this learner profile model. In Section VI, we propose a solution to implement adaptive and personalization learning on top of our propose model and infrastructure. Section VII summarizes our contributions and provides several perspectives.

II. RELATED WORK

A. *Ontologies in e-Education*

The interest of ontologies in the domain of e-Education has been repeatedly pointed out during the last decade. In [8], authors analyse reasons and ways to use ontologies in e-Education and their goals. One of these goals is implementation in e-Learning platform of features like: management of learning institutions and/or learning platforms and their actors, management of curricula, management of pedagogical resources, management of learning process and assessments [2] .

[9] presented a review and overview of works on ontologies in the domain of e-Education and map existing works to needs that ontologies can address. [9] classify ontologies in e-Learning into four categories: (1) curriculum modelling and management, (2) learning domains, (3) learner data and (4) e-Learning services. Our proposition could be classify in learner data. All these aspects have been addressed by different authors. [10] propose an e-Learning management system based on an ontology and [11] propose ontologies built from French official texts describing curriculum and populate them. Other ontology models like CURONTO [12] are dedicated to curriculum management and to facilitate program review and management.

But, to the best of our knowledge, none of the ontologies reported in the literature has been used in an industrial context, or evaluated on the data of an EdTech company. Moreover, this state of the art works do not integrate public authorities recommendations or standards model, even if there are institutional norms and standards defined by public authorities or standardization committees. We point out that we speak of e-Education when e-Learning is applied in the public institutional context (public school or academic) or when it has to respect the recommendations and standards of the Ministry of Public Education. In French education, as part of the Education Code [4], the Ministry of Education defined and published a common reference base of knowledge and skill[1]. It standardizes the content of courses by specifying knowledge and skill that a student has to acquire at each step of school curriculum. Moreover, the French Ministry of Education specifies a format for digital pedagogical resources description called ScoLOMFR [6], based on the IEEE standard Learning Object Metadata (LOM) [7] and its French version, LOMFR[2]. As a result, any e-Learning environment developed by public institutions or private companies must meet these standards and norms to ensure a wide dissemination in e-Education context. This is precisely one of the purposes of our project: develop solutions based on Semantic Web technologies and compatible with the standards defined by public authorities in the context of e-Education. We proposed a contribution for this goal in [1].

B. *Ontology based User Modelling in e-Learning Systems*

The learner profiles from e-Learning platforms have historically been the descendants of the user profiles of e-Commerce platforms. For a long time, e-Learning platforms represented learner profile as a user preferences model in order to recommend and sell online training courses as it is the case in e-Commerce [28]. These preferences are also used to personalize user requests [13], and/or recommend pertinent answers.

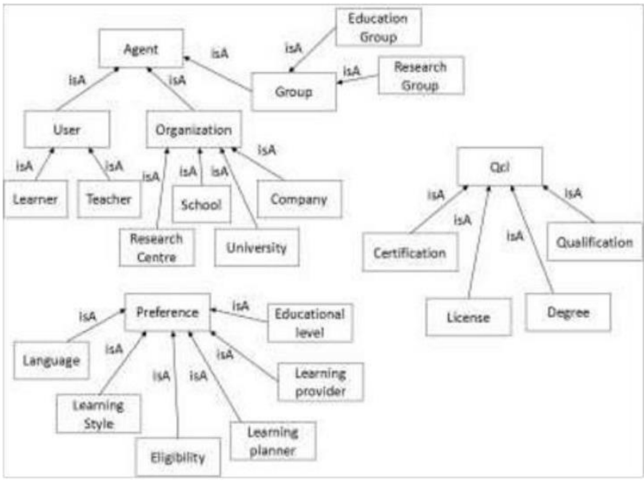

Fig 1 Generic Learner Model.[14]

In e-learning context, [14] proposes an ontology-based model to represent generic user profiles and associated knowledge, depicted in figure 1. The main goals of this model are to share information on different users, to provide a common vocabulary used and also to describe formally user preferences. The learner profile contains contact information, relation information, preference information, goal information, portfolio information. But, there is many features in e-Education systems that a learner model based solely on preferences couldn't help to reach. Among these features we get, learning model, which could be define as the ability to structure the learning path according to the individual skill profile of each learner [15]. We also get the assessing Learner model [15] which defines assessments, with/without misconception, corresponding to a given learner for validating a given skills or knowledge. Learner model should also allow to knows all knowledge and skills validated by a leaner and assist him in the remediation process [16] for those he fails to validate. Another important feature is helping leaner to keep in mind what he has learned,

[1] original name: *Socle commun de connaissance, de compétences et de culture*

[2] http://www.lom-fr.fr

ISSN No:-2456-2165

by implement the forgetting curve [17] and determine when learner need to revisit some knowledge and skills. At last, considering that learners do not have the expertise to define their learning content and that Educlever's target is public educational institutions, the proposed learner model have to allow learning process to meet public standards and recommendations and this could not be possible when using solely preferences.

In the state of the art of learner profile design in e-Learning context, we find works which mainly modelling learner's preferences. [18] proposed a framework to model learner's preferences and recommend learning content based on collaborative and content based approach. In a similar approach, [2] and [19] propose a learning recommendation based on a learning profile represented by a fuzzy tree. In this recommendation system, learner model includes the learner's background, learning goal, required learning categories, and learning activities are used in the recommendation process. They used fuzzy logics to handle similarities between concepts and applied fuzzy tree-structured data model to model the learner profiles. [20] propose an ontology based learner profile modeling which describes the following concepts: knowledge, skills, preferences or habits and interaction. They show the way an educational resource is assigned according to the learner's preference states during learning activities.

Unfortunately, to the best of our knowledge, existing works do not implement important features described above like monitoring learning path, helping learner to achieve institutional pedagogical objectives and keep in mind what he has learned and recommend him/her a fit pedagogical resource. Moreover, we did not find any integration and evaluation of e-Education system in a real industrial context. We propose to overcome these limits with Semantic Web models and technologies to design a learner profile that makes it possible to find the history of learner activities. We integrated this work in the Educlever company industrial context and evaluated it.

III. ONTOLOGY MODELS OF KNOWLEDGE, SKILLS AND PEDAGOGICAL RESOURCES

We had proposed two ontology-based models to describe knowledge and skills referential and pedagogical resources [1]. These models have been integrated in the Educlever software infrastructure. These ontologies reuse and extend ***EduProgression ontology*** [11] which is modelling the official common base of knowledge and skill. Since our learner profile model reuse these two models, we briefly review them.

A. Knowledge and Skills Modelling

We describe knowledge and skills in the ontology called ***Referential*** which contains all the elements of knowledge and skill available through the e-Education solution. The concept *Cocon*, which stands for *compétences et connaissances* in French (knowledge and skills), is the keystone of ***Referential ontology***. a *Cocon* represents an atomic element of knowledge or skills learnt by students on the e-Education solution. In the Educlever system, an example of *Cocon* could be the **multiplication of two integers** identified with URI refeduclever: Multiply Two Integers[3].

Figure 2 shows ***Referential ontology*** model implements in Educlever system. This figure shows that we formalize the concept *Cocon* as an equivalent class to *EKS* from the ontology ***Eduprogression*** [11]. Thanks to formalization, we extend ***Eduprogression*** and then to use properties of concept *EKS* from ***ontology Eduprogression*** (*has Course*, *has Cycle and has Learning Domain*) to describe a *Cocon*.

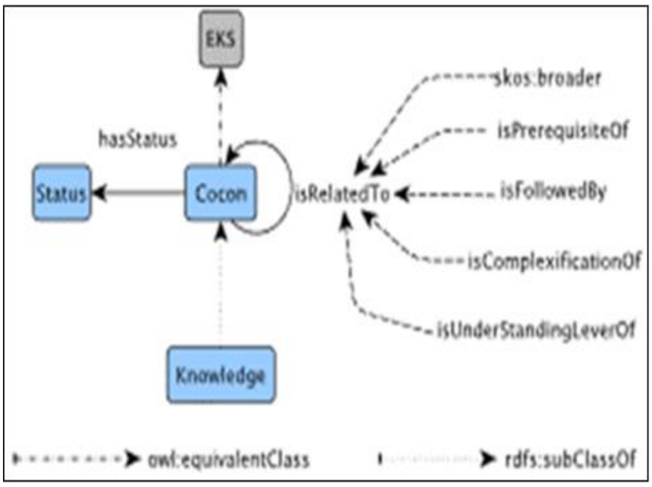

Fig 2 Referential Ontology

For instance, **multiplication of two integers** is a *Cocon* of cycle two (*has Cycle*), its learning domain is the first domain of French education standards, languages for thinking and communicate (*has Learning Domain*) and its course is Mathematics (*has Course*). So, with this formalization, we integrate public standards description, and we improve semantic of *Cocon* by adding new properties, between them, (i) *skos:broader*, for hierarchical relationships between *Cocon*, (ii) *is PrerequisiteOf*, for dependency relationship between *Cocon*, (iii) *is Followed By*, for chronological dependency between *Cocon*, (iv) *is Complexification Of* and (v) *is Under Standing Lever Of* [1] which are specialization of the relation *is Related To*.

B. Pedagogical Resources Modelling

We describe pedagogical resources available through the e-Education solution in the ***ontology Corpus***, which uses a specific vocabulary. Figure 3 describes ***Corpus ontology*** and we could observe that the *OPD* class is the keystone of the ontology. *OPD* stands for *Objet Pédagogique* in French (*Pedagogical Object*) and it represents pedagogical resource created to learn and acquire knowledge or skills.

[3] refeduclever: http://www.educlever.fr/edumics/refeduclever\#

ISSN No:-2456-2165

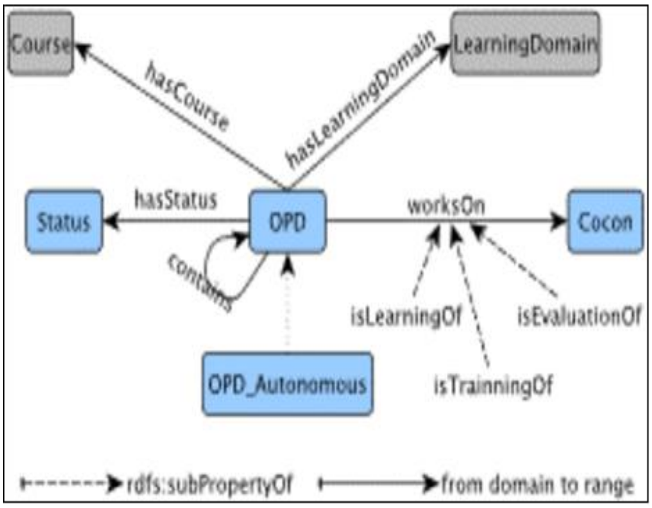

Fig 3 Corpus ontology

There are two key properties, properties *works On* and *has OPD*. Property *works On* relies an instance of *OPD* to an instance of *Cocon*, from the ***Referential ontology***, tackled in the pedagogical resource. *worksOn* is specialized into *is Learning Of, is Trainning Of,* and *is Evaluation Of* properties. These properties describe the role of an *OPD* in the learning process of a *Cocon*. Property *has OPD* links two *OPDs*. It represents partonomies and expressing how some pedagogical resources are built as a combination of other pedagogical resources.

Autonomous OPD is a subclass of *OPD* gathering the resources which do not need any other resources to be used. Thanks to ***Corpus model***, e-Education company could provide pedagogical resources annotated on public standards, class Course and Learning Domain from ***Eduprogression***. Moreover, private companies could share pedagogical resources when theses pedagogical resources allow to learn or evaluate many different skills and knowledge.

IV. ONTOLOGY BASED LEARNER MODEL

This section presents our proposition of learner profile model based on Semantic Web technologies. In addition to the explicit description of a learner, our model allows an integration of public standards and recommendations and effective implementation of Educlever use cases involving learner profile. More precisely, our model describes relationships between learners and their preferences. It also allows the description of the learning progress over time and allows evaluation of this learning process.

A. Learner Profile Modelling

Our proposed learner profile model describes learner concept. To implement this intrinsic description of learners, we reuse the concepts of the ***Friend-Of-A-Friend (FOAF) ontology*** [21]. Indeed, as depicted in figure 5, we observe that the class *Learner* inherits from the class *User* which itself inherits from the class *Person* of the ***FOAF ontology***. We designed the class *User* because *learner* is not the only kind of user of e-Education platform, there are others users like *teachers*. Since *Learner* is a subclass of *User*, *Learner* have property *foaf:topic_interest* and then we are able to manage learner preferences like existing works.

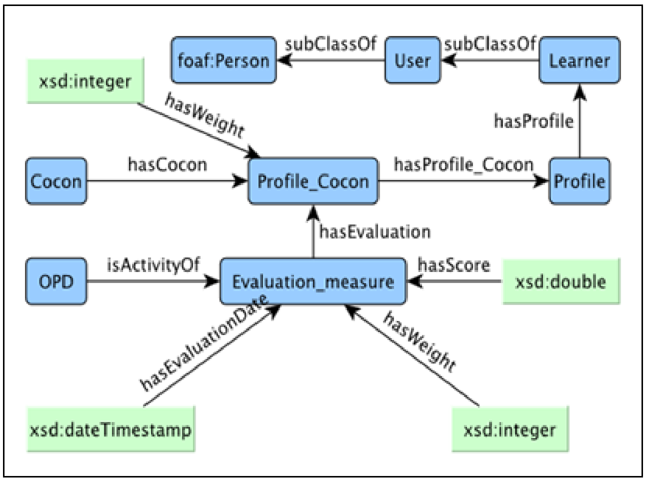

Fig 4 Learner Profile Model

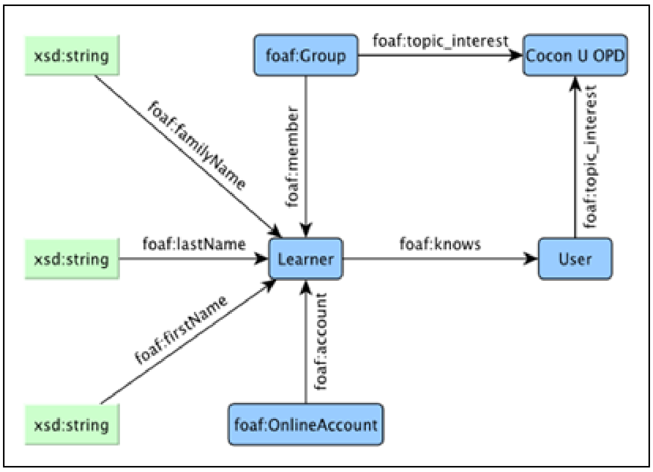

Fig 5 Learner Description with FOAF Ontology

But here, preference will be defined as union of *Cocon* and *OPD*, to show the knowledge and skills appreciated or wanted by the learner as well as the pedagogical resources. Our model describes learners group and allow them to share interested *Cocon* and *OPD* in these groups. A deep analysis of learners activities will also allow us to recommend learner group to learner based on their activities and also their interested topic. We recall that interested topic does not only means a topic like by learner but it also could be a topic that a learner have to acquire or mastered in his curriculum. The ontology depicted in figure 4 allows use to build a learner social network, on top of the ***FOAF ontology***, where learner could share knowledge, pedagogical resource solutions and learning process experience. So, the next step is the description of the learning process.

B. Learner Activities Modelling

Learning process of a learner is defined by the set of activities perform by him. Then, our second contribution is the description of these learner activities in our learners profile model. In order to describe learner activities, we designed the model depicted in figure 5. A key concept of this model is the concept *Profile cocon* which represent acquisition of a *Cocon*, mentioned using property *has Cocon*, through a set of *OPD* evaluation. The concept *Profile* represents the profile of a learner, it is a set of *Profile_cocon* that learner begins acquisition or already acquires. The concept *Profile* contains learning path of a learner for a given *Cocon* and for a given period of time. Indeed, learning path for a given *Cocon* is the set of *Cocon* that learner have to acquire and also activities on these *Cocon*. This set of *Cocon* are *Cocon* related to the given *Cocon* with properties *is Parent Of*, *is Prerequisite* and so on, from ***Referential ontology***. A curriculum of a learner contains many *Cocon* and all of them do not have the same pertinence in the curriculum. The property *has Weight* of *Profile_cocon* represent the weight of the *Cocon* in the curriculum. Similarly, acquisition of a *Profile_cocon* need many evaluation represented by the concept *Evaluation_measure*. Since, all *Evaluation_measure* do not have the same pertinence for the acquisition of the given *Cocon*, each *Evaluation_measure* have a weight denoted by the property *has Weight.*

Now, we consider the example depicted in figure 6. This figure 6 shows a part of learning activities of learner *learner_1* who learns Identify Base Sentence Components, which is a *Cocon.*

In order to acquire this knowledge and skill, a learning path is Identify Base Sentence Components T.id-ent. CI. ident. sujet→ Identify Base Sentence ComponentsT. ident. CI→Identify Base Sentence Components, where they are related with property *is Parent Of*.

Figure 6 also shows that *profile_learner_1* of *learner_1* has two *Profile_cocon (the way the user learns a Cocon)*: *profile_cocon_1*, for *Cocon* Identify Base Sentence Components, and *profile_cocon_2* for *Cocon* Identify Base Sentence ComponentsT. ident. CI. Each of these *Profile_cocon* have two *Evaluation_measure* which represent assessments performed online by *learner_1* on different pedagogical resources in different day with their score. A *Profile_cocon* could have more than two *Evaluation_measure*, or less than two. These scores of an assessment are over 100 marks and allow Educlever system to compute average using weight and then validate or not acquisition of the learning path.

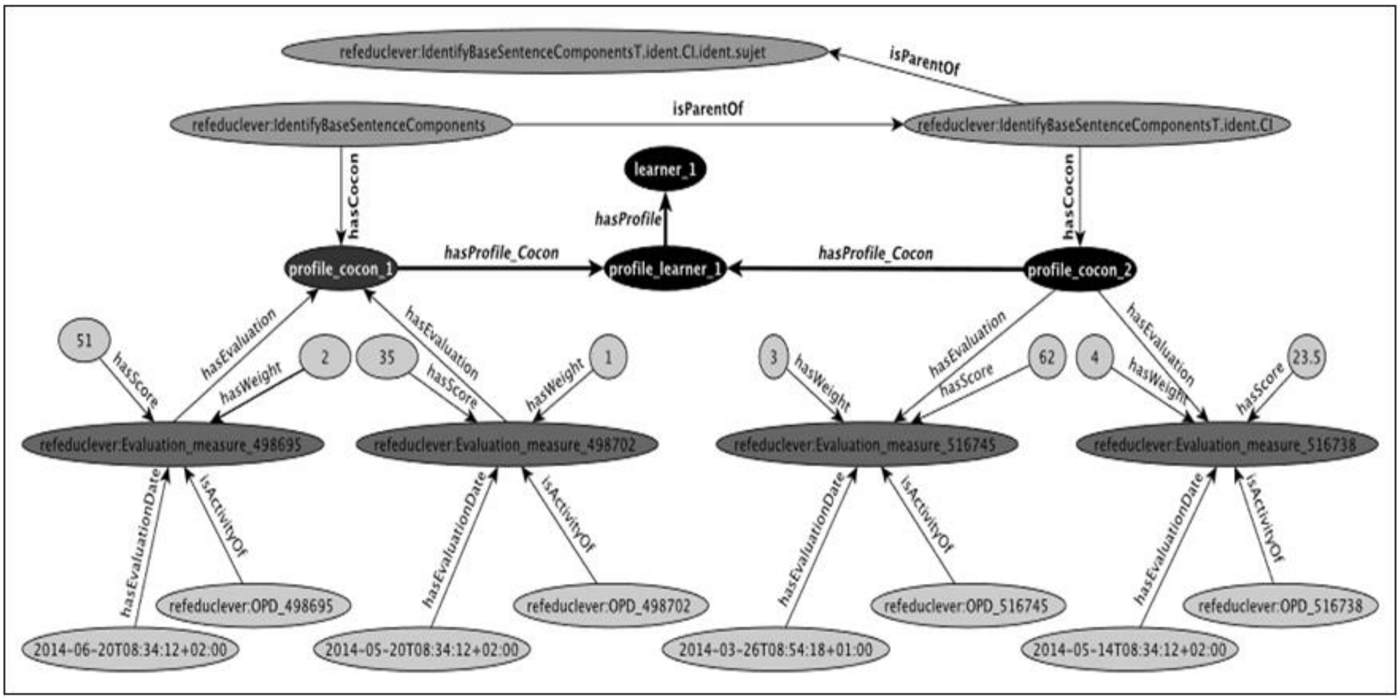

Fig 6 Learner Profile Instances Example

V. INTEGRATION IN SEMANTIC WEB BASED EDUCATIONAL SYSTEM

Integration of our model takes place at two levels: (i) at the level of the architecture of the Educlever platform and (ii) at the functional level.

A. Knowledge Extraction and Integration

After modelling the learning profile, we have faced the challenge of deploying this model in Educlever context and migrating existing data, stored in relational databases. We upgrade our proposed architecture [1]: (i) Simple Architecture and (ii) Federated. We remind that these architectures were built on top of triple stores to process RDF data from the ***Referential*** and ***Corpus*** datasets.

In simple architecture, we load ***learner profile ontology*** and instances in a single triplestore as depicted in Figure 7. Thanks to the flexibility of the Semantic Web solution and precisely the use of URIs to identify resources, this update does not affect existing features already implemented. Thus, SPARQL queries implementing these features return exactly the same results. All that remains is to implement the functionalities relating to learner profiles. But with this architecture, in case of failure of the triple store data will no longer be available. In industrial context this is an important risk which we have to prevent.

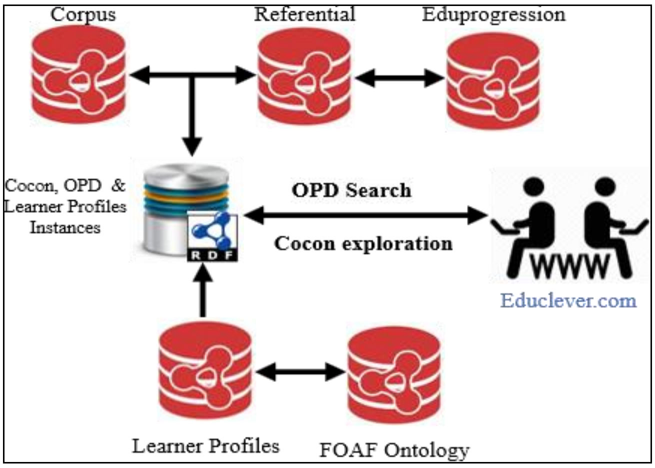

Fig 7 Simple Architecture

In the federated architecture, we add to our federation a third triple store for learner profile ontology and instances as depicted in Figure 8. Thus, the federated endpoint allows to query, in a transparent way, the three datasets. Thanks to Semantic Web technologies, this operation is made with configuration instructions in an industrial production context. Moreover, this architecture prevents failure of one of the triple store, and allows to query others as a single dataset. This context and scenario is typical of the need to take into account legacy software and information system from real industrial contexts as well as the service quality constraints, etc.

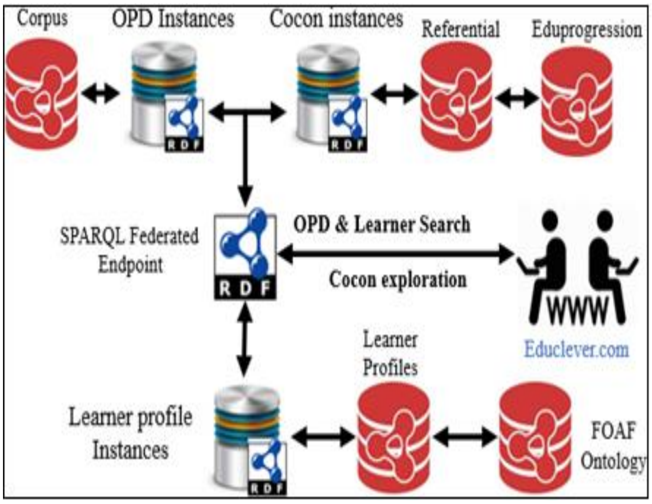

Fig 8 Federated Architecture

B. Knowledge Engineering: Learning Analytics

Now, we are able to address Educlever's existing needs and the features they wish to add to their system. Among Educlever's existing needs dealing with learner profile, we find: R_1 The search for administrative information on a learner, R_2 The number of evaluations carried out by a learner, R_3 . The score of the evaluations carried out by the learner, thus this uses case allows to know if the evaluation was passed successfully or not, R_4 . The most recent evaluations carried out by the learner and used to measure his acquisition of the knowledge and skill assessed (*Cocon).*

In addition to the above listed use cases implemented in the Educlever platform, using Semantic Web technologies allows to implement additional features. Among these new features we have: R_5 find information on learning resources that have been used to assess a learner, this feature avoids redundancies in a given time interval and thus avoids distorting the assessment. R_6 Retrieve all the knowledge and skill (*Cocon)* acquired by the learner. R_7 Determining whether a learner has all the prerequisites to start learning of a given *Cocon*, it is a direct consequence of the previous feature. R_8 Our model also allows us to assist a learner in the remediation process [16] in the event of failure on a given *Cocon* learning. For this, we are looking for *Cocon* already acquired by the learner or for which he has all the prerequisites and which are levers for understanding (*is Under Standing Lever Of*) of the *Cocon* to acquire. R_9 . We have integrated an implementation of the forgetting curve [17] in order to accurately get acquisition level of a *Cocon* over time

Implementation of these features has been done through one SPARQL query or functions executing several SPARQL queries. While existing features have been implemented in Educlever platform with only functions and execute many queries. This shows useful of Semantic Web based platform since it implement more features than the existing Educlever platform. Table 1 compares existing Educlever Platform and Semantic Web platform effectiveness in uses cases implementation. This table shows use case implemented on the existing Educlever platform and our proposed upgrade.

Table 1 Implementation of the use Cases

	Existing Features				**Added Features**				
	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9
Existing Educlever Platform	✓	✓	♦	♦					
Educlever Semantic Web Based Platform	✓	✓	✓	✓	✓	♦	♦	♦	♦

In the previous table 1, case with symbol ✓ means that we implement these uses cases with one query. While symbol ♦ means the implementation use a function which itself executes several queries in order to fill the use case. In table 1, empty cells means that use case is not implemented.

Then, we could observe that using Semantic Web Based platform we meet all the uses cases. The new platform, could also lead the personalization of learning experience [25, 26]. For a best learning experience, predicting learner's success on pedagogical resources is an important step.

ISSN No:-2456-2165

VI. ADAPTIVE LEARNING IN SEMANTIC WEB BASED EDUCATIONAL SYSTEM

A. Predicting Learner Success based on the Learning Resources Knowledge Graph

Let $U=\{u_1, u_2,\ldots, u_{|U|}\}$, $I=\{i_1, i_2,\ldots, i_{|I|}\}$ and $S=\{s_1, s_2,\ldots, s_{|K|}\}$ represent the sets of learners , items (*OPD*) and Skills (*Cocon*) respectively. The task of this section focus on predict the outcome of learner towards a specific item, and then recommend learner items that he/she will need to improve his/her master of a certain knowledge. Each learner is associated with a sequence of items from the items set I. Each learner will try a single item for a several times. By sorting interaction records in a chronological order, we can form the interaction sequence $I_u = \{i_1^u, \ldots, i_2^u, i_{|I^u|}^u, \}$ for user u. The index t denotes the relative time index, instead of absolute time index in temporal recommendation. With the above notations, given learner u history sequence I^u, we aim to infer the item(s) that a learner is likely to do which will help him/her quickly learn a certain knowledge.

To perform this task we going to setup a knowledge graph embedding model. The knowledge graph embedding algorithm transforms the knowledge graph into a low-dimensional dense real number vector which is used as input of existing machine learning algorithms: the vector representation of nodes can be sent as a feature to support vector machine and other classifiers for node classification and prediction tasks. We used the knowledge tracing to encode training records. The most popular model is Bayesian knowledge tracing (BKT), which is a hidden Markov model [27]. It is used in many intelligent tutoring systems to model each learner's mastery of the knowledge being tutored and models student knowledge in a Hidden Markov Model as a latent variable, updated by observing the correctness of each student's interaction in which they apply the skill in question. But for this work, we choose to use Knowledge Tracing Machines (KTM).

Knowledge Tracing Machines (KTM) [22,23] is a sequence prediction problem where the goal is to predict the outcomes of students over questions as they are interacting with a learning platform. By tracking the evolution of the knowledge of some student, one can optimize instruction. ***It uses factorization machines (FMs), a model for regression or classification***, encompasses several existing models in the educational literature as special cases to estimate student knowledge accurately and fast even when student data is sparsely observed, and handle side information such as multiple knowledge components and number of attempts at item or skill level. This approach allows to fit student models of higher dimension than existing models, and provides a testbed to try new combinations of features in order to improve existing models. In this work, we use KTM as state of the art algorithm, to predict learner's behavior towards *OPD* learner has never interacted before, and based on the outcome, to recommend learner *OPDs* to improve his/her mastery of a certain *Cocon*.

Factorization machine (FM) is a machine learning algorithm based on matrix decomposition proposed by Steffen render. The form of Factorization machine is like:

$$y = w_0 + \sum_{j=1}^{p} w_j x_j + \sum_{j=1}^{p} \sum_{i=j+1}^{p} x_j x_i \sum_{f=1}^{k} v_{j,f} v_{i,f}$$

Among them $v_{j,f}$ and $v_{i,f}$ are respectively a hidden factor of the corresponding hidden vector of character i and j. Usually, because of the sparse data, we cannot learn w, but we can learn the parameter vector v of i and j features respectively through the data of i features and other features. So we predict the value of w through the product x_j*x_i, which solves the data sparsity problem.

B. Experimentation and Evaluations

In this section, we implement the prediction model and evaluate its capacity to predict success on learner training.

- *Datasets*

We used two dataset in these experiments: (i) **Assistmens09:** The dataset of Assessments [24] described in (Feng, Heffernan and Koedinger 2009) which is one of the dataset used in state of the art work KTM with 4217 students over 26688 questions, 123 knowledge components (*OPD*) and 347k interactions. (ii) **Educlever knowledge graph** 1330 users attempting 334k items, 17127 skills with 85270 times interactions.

- *Data and Encoding of Interaction*

Here, we present how to encode the observed data into sparse vector x. First, we need to choose features which will be used in modeling :

- **Users:** Assume there are n learners, the first n feature will be for n learners, and in KTM, it use one-hot encoding. Let's say if learner i is involved in this interaction, then its x_i value will be 1, the rest for the other learners set to 0.
- **Items:** Assume there are m items (*OPD*). So there will be m features to represent m items. If item j is involved in one interaction, then y_j are setting to 1, the rest items remain 0.
- **Skills:** Let us assume there are s skills (*Cocon*). There will be s features to represent s skills. If one interaction involves several skills, the corresponding index of skills are setting to 1.
- **Wins & Fails** Allocate s features to distinguish if a user learn a skill is success if the attempt was correct, s more features as opportunities to have learned a skill if attempt was incorrect.

Table 2 is an example for encoding of users, items, skills, wins, and fails. Here, we have n=2 learners, m=3 questions, s=3 skills. The first row is an interaction of learner 2 tried question 2, and question 2 involves skill 1 and skill 2. At the beginning, learner 2 has no interaction with any question before, so he/she doesn't have the chance to learn any skill, so the count of wins & fails for any skill are all 0. Then we can see the outcome is 1 for question 2, so in the second row, the wins for skill 1 and skill 2 add 1. Therefore, we encode triplets with N = n + m +3s = 14 features.

Table 2 Example of Encoding for Training in KTM

User		Items			Skills			Wins			Fails			
1	2	Q1	Q2	Q3	KC1	KC2	KC3	KC1	KC2	KC3	KC1	KC2	KC3	out
0	1	0	1	0	1	1	0	0	0	0	0	0	0	1
0	1	0	1	0	1	1	0	1	1	0	0	0	0	0
0	1	0	1	1	0	1	1	0	1	1	0	1	1	1
0	1	0	0	1	0	1	1	0	2	0	0	1	0	0
0	1	0	0	1	0	1	1	0	2	0	0	2	1	1
1	0	0	1	0	1	1	0	0	0	0	0	0	0	1
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0

➢ *Extract Data from Educlever Knowledge Graph.*

We use the SPARQL query to extract user interactions from the Educlever Knowledge Graph, more specific, from profile graph. In order to be able to use Educlever dataset in the state of the art work KTM, we get user, item, skills from the graph, item and skill corresponding to *OPD* and *Cocon* in Educlever Knowledge Graph. And the interactions should be grouped based on user and item in a chronological order. So, we use the following SPARQL query to get the data:

```
select ?user ?opd ?profile_cocon ?cocon ?score
where{
    ?user refeduclever:hasProfile ?profile .
    ?profile refeduclever:hasProfile_Cocon ?profile_cocon.
    ?profile_cocon refeduclever:hasCocon ?cocon .
    ?profile_cocon refeduclever:hasEvaluation ?eval.
    ?eval refeduclever:hasScore ?score1 .
    ?eval refeduclever:hasEvaluationDate ?date .
    ?eval refeduclever:isActivityOf ?opd .
    BIND(?score1/100 AS ?score) .
}
GROUP BY ?user ?opd ?profile_cocon ?cocon ?score
ORDER BY ASC (?user) (?date)
```

Note that in the encoding of KTM, we use, correct = 1 or 0 to denote learner's performance towards item. So, we convert the score from Educlever to 0 or 1 to match the algorithm.

➢ *Data Processing*

Note that data stored in Educlever Knowledge Graph are mostly URIs, but according to KTM, every learner, every item and skill should have a unique ID which could be used to do one-hot encoding. So, after got data, the first step is to process the data to match the algorithm. For learner, the URI is like http://www.educlever.fr/edumics/ refeduclever# USER_10. The last part of URI is number, only extract number will be fine. The same works for items. However, in Educlever Knowledge Graph knowledge graph, there is no id for skills. Therefore, we perform a query to get all skills in the graph, and then randomly mapping allocate them unique id.

```
SELECT ?cocon
WHERE{
    ?cocon rdf:type refeduclever:Cocon
}
```

➢ *Metrics and Evaluation*

To evaluation our model we going to use two standards metrics:

- **ACC:** Stand for **ACC**uracy**.** The accuracy rate represents the proportion of correctly classified samples to the total number of samples. If we predict 10 samples and 8 samples are correct, the accuracy rate is 80%.
- **AUC:** Stands for **A**rea **U**nder the **C**urve. It is defined as the area below the ROC curve: The horizontal axis of ROC curve is true positive rate (TPR), the vertical axis is false positive rate (FPR), also known as true positive rate. AUC can only be used for the evaluation of two classification model.
- **NLL:** Stands for **N**egative **L**og-**L**ikelihood. This metric becomes unhappy at smaller values, where it can reach infinite unhappiness, and becomes happy at larger values. Because it is summing the loss function to all the correct classes, what's actually happening is that whenever the network assigns high confidence at the correct class, the unhappiness is low, but when the network assigns low confidence at the correct class, the unhappiness is high.

For our evaluation, we run experiments for 5 times and average the result. We first run the experiment on dataset Assessments09. When encode interactions only with user and item, got the result:

Table 3 Experiment Results on Assessment09

Model	ACC	AUC	NLL
Users, Items	0,726	**0,730**	0,571
Users, Items, Skills	0,703	0,709	0,580
Users, Items, Skills, Win, Fail	**0,737**	**0,730**	0,545

For the experiment on Educlever Knowledge Graph, the results are in the table below:

Table 4 Experiments Results on Educlever KG

Model	ACC	AUC	NLL
Users, Items	0,833	0,695	0,424
Users, Items, Skills	0,832	0,80	0,383
Users, Items, Skills, Win, Fail	**0,855**	**0,906**	0,280

We observe that using KTM to make prediction on Educlever Knowledge Graph more relevant and get satisfied results with 0.885 accuracy and 0.906 AUC when encoded with wins and fails. And, from the two experimentation, we observe that the last model, where all the feature are encoded, is the more efficient for the prediction.

After running the experiment, we will get a result file which consist of probabilities that a user will do it correctly towards the item. Based on these probabilities, we sort these probabilities, and select the top 10 items that user tend to do badly, and recommend user correspondent skills to practice.

Because user will try the same item for several times, so there has more than one records for a user towards a single item. Thus, during the prediction, we will get different probabilities for the same user-item. Therefore, the first step is to keep only one record with the highest probability for a single user-item. And then sort all the items that the user has interacted with.

VII. CONCLUSION

This work reported a knowledge modelling experience in industrial context to propose an e-Education solution based on Semantic Web models and technologies. We recalled our previous work on modeling knowledge and skill (*Cocons*), as well as pedagogical resources (*OPD*). Then we presented our proposal of an ontology describing a shared conceptualization of a learner, learner activities over time and learner preferences, that can smoothly extend our modelling in order to implement user-oriented use cases.

We also proposed two extensions of Educlever architectures proposed in our previous work, which allows us to integrate and process learner profiles ontology and instances in the Educlever platform. Thanks to Semantic Web technologies, we integrate learner profiles without any change on features already implemented which continue to work properly. This integration allows us to implement existing features on learner profiles, as well as features which could not be implemented on the existing Educlever infrastructure. To implement these features using Semantic Web technologies we use, depending on the complexity of the feature, either a SPARQL query or a function executing several SPARQL query. Based on these features, we are able to provide algorithms to personalize learning path for a given learner. The most immediate continuation of this work is the experimental evaluation of our model in Educlever context.

Next to all the previous contributions, we manage to get real data from Educlever Knowledge Graph and format them in order to perform prediction. In this work, we are able to extract useful information to make predictions and successfully make recommendations based on existing state of the art algorithm.

One of the next challenges is implementation of adaptive and personalized learning based on learner profiles. Indeed, we plan to customize state of the art prediction algorithm and combine it with rule-based reasoning mechanisms in order to propose learning path and recommend relevant pedagogical resources for his/her training. We will also propose measures to evaluate the acquisition of a given knowledge and skills (*Cocons*). Based on this measure, we are going to consider creation of student groups based on pedagogical objectives, such that they have complementary knowledge and skills (*Cocons*) or same level of acquisition for a given set of knowledge and skills.

ACKNOWLEDGMENT

Thanks to Educlever[4] to be part of this project and share data with us. Thanks to WIMMICS/INRIA[5] to hire me for this project. And finally, thanks to the company Modelling Innovative Intelligence Artificial (M2IA) for the technical and financial support.

REFERENCES

[1]. G. Fokou Pelap, C. Faron-Zucker et F. Gandon, «Semantic Models in Web based Educational System Integration,» chez 14th International Conference on Web Information Systems and Technologies (WEBIST), Seville, 2018.

[2]. K. Balaji et G. Poorni, «An Effective Recommendation System for E-Learning Using Fuzzy Tree,» Semantic Scholar Journal, 2017.

[3]. M. Srivastava, H. Pandey, S. Kumar Shukla et B. Kumar Thakur, «A Literature Review of E-Learning Model Based on Semantic Web Technology,» Semantic Scholar Journal, 2014.

[4]. Nationale Ministère de L'éducation, «Code de l'éducation consolidé au 1er Janvier 2018,» Service Publique, 2018.

[5]. Nationale Ministère de L'éducation, «Le socle commun de connaissances, de compétences et de culture,» Service Publique, 2017.

[6]. Réseau Canopé, «ScoLomFR : Outil de description des ressources numériques,» Service Publique, 2011.

[7]. IEEE-LTSC-LOM Committee, «IEEE Standard for Learning Object Metadata,» IEEE Standards Association, 2002.

[8]. L. Jaffro, «Les objets de l'éducation : quelle ontologie ?,» Revue de métaphysique et de morale, vol. 4, n°56, pp. 429-448, 2007.

[9]. A.-Y. Maha, G. Remya et A.-F. Auhood, «Ontologies in E-Learning: Review of the literature,» International Journal of Software Engineering and its Applications, vol. 9, pp. 67-84, 2015.

[10]. M. Rani, K. V. Srivastava et O. P. Vyas, «An Ontological Learning Management System,» Computer Applications In Engineering Education, vol. 24, n° %15, p. 706–722, September 2016.

[11]. O. R. Rocha, C. Faron-Zucker et G. Fokou Pelap, «A Formalization of the French Elementary School Curricula,» chez Knowledge Engineering and Knowledge Management – (EKM) and Drift-an-LOD, Bologna, Italy, Revised Selected Papers, Bologna, 2016, pp. 82--94.

[12]. A.-Y. Maha, A.-F. Auhood et G. Remya, «CURONTO: An Ontological Model for Curriculum Representation,» chez Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education, ACM, 2013, pp. 358--358.

[4] https://www.educlever.com/

[5] https://www.inria.fr/fr/wimmics

ISSN No:-2456-2165

[13]. G. Koutrika, E. Pitoura et K. Stefanidis, «Preference-Based Query Personalization,» chez Advanced Query Processing, vol. 36, Heidelberg, Springer International Publishing, 2013, pp. 57--81.

[14]. K. Rezgui, H. Mhiri et K. Ghédira, «An Ontology-based Profile for Learner Representation in Learning Networks,» International Journal of Emerging Technologies in Learning (iJET), vol. 9, n°3, pp. 16--25, 2014.

[15]. M. C. Desmarais et R. S. Baker, «A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments,» User Modeling and User-Adapted Interaction, vol. 22, n° 1-2, pp. 9--38, April 2012.

[16]. C. Reverdy, «L’accompagnement à l’école : dispositifs et réussite des élèves,» Dossier de veille de l’Institut Française de l'Education (IFE), vol. 119, pp. 1--31, 2017.

[17]. J. S. Fisher et G. A. Radvansky, «Patterns of forgetting,» Journal of Memory and Language, vol. 102, pp. 130 - 141, 2018.

[18]. K. Abhinav, V. Subramanian, A. Dubey, P. Bhat et A. Divakaruni Venkat, «LeCoRe: A Framework for Modeling Learner's preference,» chez Proceedings of the 11th International Conference on Educational Data Mining, (EDM) 2018, Buffalo, NY, USA, July 15-18, 2018, International Educational Data Mining Society (IEDMS), 2018.

[19]. D. Wu, J. Lu et G. Zhang, «A Fuzzy Tree Matching-Based Personalized E-Learning Recommender System,» IEEE Transactions on Fuzzy Systems, vol. 23, n° 6, pp. 2412-2426, 2015.

[20]. A. Korchi, N. El Amrani El Idrissi et L. Oughdir, «Modeling and Implementing Ontology for Managing Learners’ Profiles,» International Journal of Advanced Computer Science and Applications, vol. 8, n° 8, pp. 144--152, 2017.

[21]. D. Brickley et L. Miller, «FOAF Vocabulary Specification 0.99,» 2014.

[22]. C. Piech, J. Bassen, J. Huang, S. Ganguli, M. Sahami, L. J. Guibas et J. Sohl-Dickstein, «Deep Knowledge Tracing,» chez Advances in Neural Information Processing Systems, vol. 28, Curran Associates, Inc., 2015.

[23]. H. Su, X. Liu, S. Yang et X. Lu, «Deep knowledge tracing with learning curves,» Frontiers Psychology, vol. 14, March 2013.

[24]. E. Lavoué, B. Monterrat, M. Desmarais et S. George, «Adaptive Gamification for Learning Environments,» IEEE Transactions on Learning Technologies, vol. 12, n° 1, pp. 16 - 28, March 2019.

[25]. A. Ettorre, O. R. Rocha, C. Faron, F. Michel et F. Gandon, «A Knowledge Graph Enhanced Learner Model to Predict Outcomes to Questions in the Medical Field,» chez Knowledge Engineering and Knowledge Management: 22nd International Conference, EKAW 2020, Italy, September 16–20, 2020, Proceedings, Bolzano, Springer-Verlag, 2020, p. 237–251.

[26]. A. Ettorre, F. Michel et C. Faron, «Prediction Of Students’ Performance In E-Learning Environments Based On Link Prediction In A Knowledge Graph,» chez Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners’ and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Durham, Springer-Verlag, 2022, p. 432–435.

[27]. A. Bordes, N. Usinier, A. Garcia-Duran, J. Weston et O. Yakhnenko, «Translating Embeddings for Modeling Multi-relational Data,» chez Advances in Neural Information Processing Systems, vol. 26, Curran Associates, Inc, 2013.

[28]. M. Dastani, N. Jacobs, C. M. Jonker et J. Treur, «Modelling user preferences and mediating agents in electronic commerce,» Knowledge-Based Systems, vol. 18, n° 7, pp. 335 - 352, 2005.

[29]. Bery L. Mbaiossoum, Atteib D. Mahamat, L. Dionlar, «Building a Smart education in developing countries: Case of Chad», International Journal of Innovation and Scientific Research, vol. 39, n° 2, pp. 170 - 179, 2018.