Towards a New Scalable Big Data System Semantic Web Applied on Mobile Learning

Mouad Banane (✉), Abdessamad Belangour
Hassan II University, Casablanca, Morocco
`mouad.banane-etu@etu.univh2c.ma`

Abstract—In Web 3.0, semantic data gives machines the ability to understand and process data. Resource Description Framework (RDF) is the liagna franca of Semantic Web. While Big Data handles the problematic of storing and processing massive data, it still does not provide a support for RDF data. In this paper, we present a new Big Data semantic web comprised of a classical Big Data system with a semantic layer. As a proof of concept of our approach, we use Mobile-learning as a case study. The architecture we propose is composed of two main parts: a knowledge server and an adaptation model. The knowledge server allows trainers and business experts to represent their expertise using business rules and ontology to ensure heterogeneous knowledge. Then, in a mobility environment, the knowledge server makes it possible to take into account the constraints of the environment and the user constraints thanks to the RDF exchange format. The adaptation model based on RDF graphs corresponds to combinatorial optimization algorithms, whose objective is to propose to the learner a relevant combination of Learning Object based on its contextual constraints. Our solution guarantees scalability, and high data availability through the use of the principle of replication. The results obtained in the system evaluation experiments, on a large number of servers show the efficiency, scalability, and robustness of our system if the amount of data processed is very large.

Keywords—Semantic Web, MongoDB, Big Data, RDF, Mobile Learning.

1   Introduction

In view of the rapid emergence of new mobile technologies and the growth of the offerings and needs of moving society in training, work is on the increase to identify new relevant learning platforms to improve and facilitate the process of learning. "Distance learning"[1]. The next step in distance learning is, of course, the port of e-Learning to new mobile systems. This is called M-Learning [2,3] (mobile learning). The search for information in the field of m-learning can be defined as an activity whose purpose is to locate and deliver learning contents to a learner according to his need for information and its context. Until now, the learning environment was either defined by a pedagogical framework or imposed by the learning content.

Recent years have been marked by the rise of mobile learning or m-learning, driven by the continued development of new mobile technologies. Learning becomes situated, contextual, and personal. This phenomenon encourages the evolution of learning methods to adapt to this new type of learning. New uses in the field of learning have multiplied in different ways. In the context of learning within companies, we seek to develop an M-Learning system whose main issues are:

- Learning at work whatever the time, place, delivery device, and the technological constraints of the learning process and adapted to the learner's profile
- Learning without breaking.

In this paper, we propose a scalable and powerful Big Data recommendation system based on Semantic Web technologies. This system is composed of two main layers, in the semantic knowledge layer we use an M-Learning domain ontology, and in the storage layer, we use a document-oriented NoSQL database named MongoDB for data management. This system is a case of using our RDFMongo [4] solution, which presents a complete system for managing massive semantic web data.

This paper is organized as follows. Section 1 introduces the notion and emergence of mobile learning technologies and the standard representation of semantic data (RDF). We make a brief state of the art on existing work that deals with the topic of using Semantic Web technologies in the area of M-Learning in Section 3. Section 4 is devoted to the main contribution of our work with the syntax and architecture of our solution. Section 6 focuses on evaluating the performance of the solution implemented on real datasets from databases and standards. Finally, we give the conclusion and perspectives to our work.

2 Related Work

The goal of mobile learning research is to build a learning environment where activities need to adapt to the learner's mobility situation using new digital technologies. The representation of the abstract model in a specific format is called binding. Today there are two bindings of the LOM schema: either of the XML binding or the RDF binding: The XML binding is easy to implement, however, it remains insufficient for the representation of all the elements of LOM since it does not allow to express the semantics of these elements. The RDF binding defines a set of RDF constructs that facilitate the introduction of LOM metadata into the web and is complemented by RDFS for defining classes, properties, and so on. The advantage of this second type of binding is that it adds semantics to the elements of LOM, except that it is not expressive enough to define all the constraints of LOM. Consider the "Title" and "Entry" elements of the "General" category that are mandatory elements in the LOM. Using RDF and RDFS one cannot specify that a property is mandatory or constrain its use at one time for a resource. As a second example, RDF and RDFS do not allow to express the inverse of a relation: thus, to say that a LOx "has a part" a LO y, will not allow inducing that the LO y "is part of "LOx. This lack of expressiveness leads us to think of the use of another more powerful formalism. In order to determine which language

is the most appropriate for solving the expressiveness problem, we have focused on identifying the necessary description logic (LD). The LD is a family of formalisms to represent knowledge in a structured and formal way. A fundamental characteristic of these languages is that they have a formal descriptive semantics. We start from a minimal logic ALC and we add to this logic the constructors necessary to define all the constraints of LOM.

Several research efforts use Big Data technologies for the management of massive RDF data, the [5] work presents a comparative study of RDF data management systems based on NoSQL databases, and the Hadoop HDFS file system (Hadoop Distributed File System)[6]. To manage large volumes of RDF data we proposed RDFMongo [4] an evolutionary system of semantic data management based on MongoDB [7], this system consists of two layers, the first manages the storage of RDF data, these data are stored in MongoDB, using an RDF triple transformation technique to a JSON document. The second layer manages the SPARQL query processing part, each SPARQL [8] request entered by the user is transformed into a MongoDB Query Language program. The SPARQL language is known by the complexity of its queries by that RDF to a graphical structure, for that, we have developed a system [9] for querying complex SPARQL queries using Apache Spark [10], in paper [11] the authors transform SPARQL queries to Hive [12] program.

This work is a case of using a large RDF data management system based on Big Data [13.14] systems like Hadoop and NoSQL databases. In [15] the authors present a case study of the use of the Semantic Web for automatic price management. Vesin et al [16] describe the use of Semantic Web technologies to facilitate the use of mobile devices in e-Learning systems. For information retrieval in Mobile Learning, existing work retrieves a lot of irrelevant information, but with the use of the Semantic Web and precisely the ontologies [17], we try to maximize the relevance of the search results. In [18] the authors aim to reduce the rate of retrieval of irrelevant information through the use of ontology, the execution of a mapping program using this ontology to successfully reduce processing time, as well as then the complexity of the calculations. El-Seoud et al [19] present the impact of the Semantic Web architecture on the development of e-Learning systems, they have derived the role of Semantic Web technologies in the process of development of learning systems through Semantic data processing mechanisms of e-Learning. Bakhouayi et al [20] presents a new Semantic Web-based solution to improve the interoperability of e-Learning systems using the next generation of SCORM specifications [21]. The principle of this system is based on the use of Resource description framework (RDF). more than that the RDF standard is a widely used interchange format for overcoming the interoperability problem by that in a system of mobile-learning several components: programs, software agents, web abuse communicate with each other, the use of RDF at the level of this communication is a means of optimization, and the availability of data used by these components. Rya [22] and RDF-3X [23] are two distributed RDF triplestores using for managing large amounts of RDF data, Rya stores RDF data in the cloud, we use both systems with our approach to measuring the efficient, and the scalability of our system.

3   System Architecture

In this section, we present our approach, the RDF data storage, and processing part is managed by our RDFMongo [4] system. The massive RDF[24] data is stored in the MongoDB database, and for the querying of this data, users use the SPARQL query language [25], it is a semantic Web standard dedicated to RDF data, RDFMongo transforms the SPARQL query of the user in MongoDB Query Language. So it's total management of RDF data and SPARQL queries by the MongoDB system. Figure 1 below shows an overview of RDFMongo, which is the background part of our Mobile-Learning system.
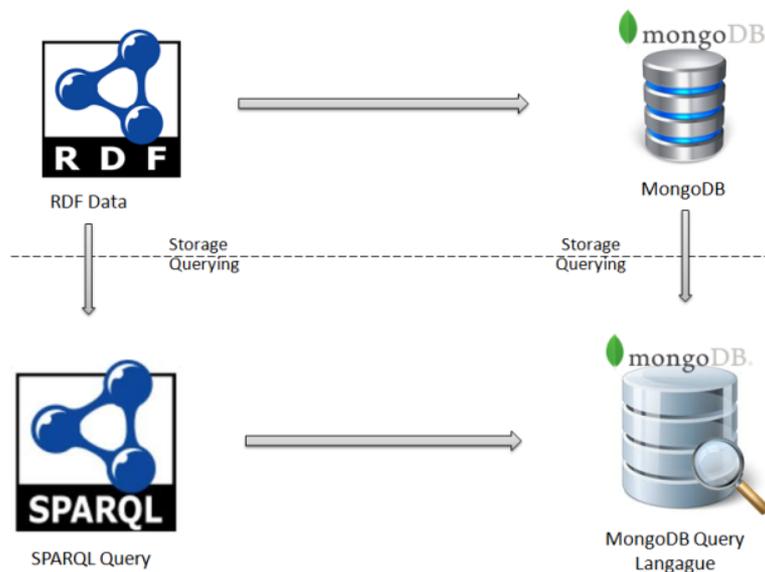

Fig. 1. RDF data management using RDFMongo[4]

As we have already quoted. We worked particularly well with the MongoDB system, which is one of the most used at present [26]. This system relies on the JSON [27] format for data representation MongoDB manages collections of JSON documents. This format allows for great flexibility in structuring data. Collections can store simple documents similar to the n-tuples of a relationship with atomic values, these collections can also include documents with complex structures that recursively nest other documents, or even tables of documents, or reference documents stored in other collections. The system allows to freely combining all these manufacturers and opens the possibility of storing data represented in a very complex way. The MongoDB system offers many features for mass management. It can be deployed on a client-server architecture but also on a cloud. It offers facilities for horizontal scalability and high availability. MongoDB allows partitioning of data, sharding, and implements master-slave replication protocols. MongoDB provides a set of operations for

inserting, deleting, querying and updating documents. Concerning the interrogation, a set of selection, projection, aggregation and classification operators is provided. Logical operators, arithmetic operators, date manipulation and character strings are also available. After seeing the background part, we now describe our two layers that make up our system that aims to handle the M-Learning data semantically.

M-learning is often considered as an extension of e-learning. This extension is not mobile only, but it is also an extension to new forms built into the learning environment that e-Learning does not allow. The learning context is a crucial aspect of mobile learning. It is, therefore, necessary to determine according to the context of what resources to send, in what way, when, on which interface, etc. The whole learning process must adapt to these contextual changes. However, contextualization in learning is not easy to achieve. The diversity of mobile technologies and dynamics in mobile environments complicate the contextualization process.

Context management is an iterative process that uses contextual information at the system level from context detection and acquisition. It involves capturing the context data, storing them and distributing the LOs to the learner according to the contextual information stored. We define the necessary steps in the life cycle of a context-aware system:

- **Context data acquisition:** This is to capture all the contextual information that is available.
- **Storage:** The captured data is stored in a meaningful and understandable way for the intended use.
- **Processing:** In our case, context information processing consists in selecting LOs from a query and applying an optimization method to refine the selected LOs.
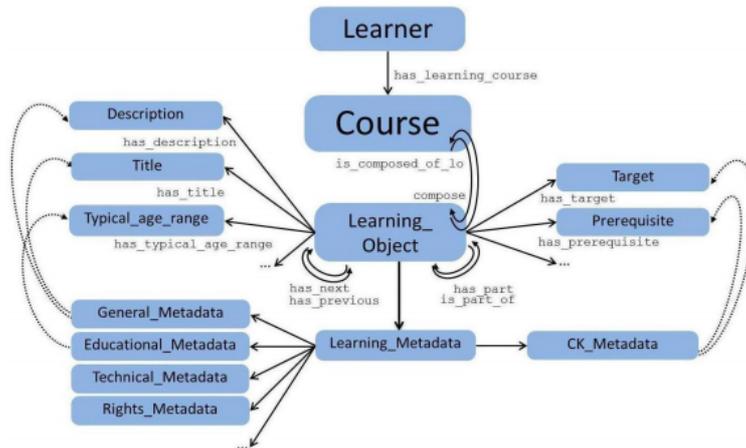

Fig. 2. Simplified preview of the learning object model

The data, as well as its metadata are stored in RDF format for use by our recommendation ontology, storage in RDF format has many advantages as: the semantic processing of the content, a format of exchange between the software agents and the program, and it is considered as the basic standard for ontologies.

For RDF data sharing, we present a new model that aims to preserve consistency through a combination of two techniques: commutativity and dependency relationships. The mechanism of commutativity consists of defining a set of operations that commute to each other on a given data structure. It must ensure consistency regardless of the order of the editing operations upon receipt.

We propose to formulate the contextual information in the form of RDF data, to record it in the MongoDB database and to apply a reasoning based on the inference rules, MongoDB is the NoSQL database that will be used to store the data RDF as a JSON document. To store the RDF triples in MongoDB, we transform the RDF triples into a JSON document using a Linked Data JSON-LD API [28].
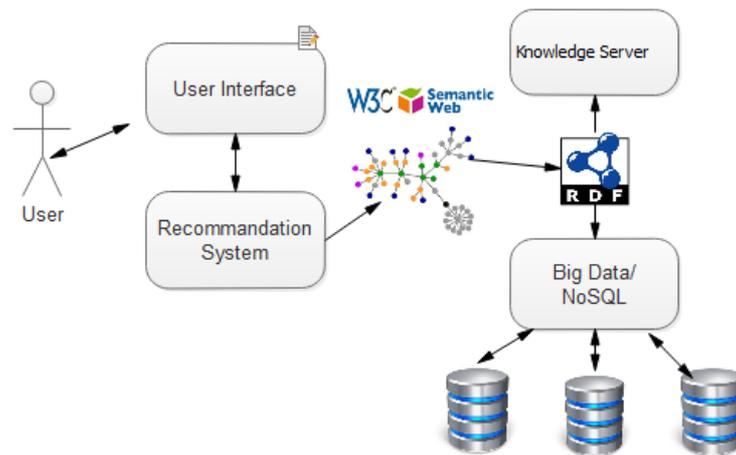

Fig. 3. Overview of System Architecture

3.1 Ontology

The ontologies make it possible to standardize the vocabulary, to standardize the language of exchange between the various actors, to compare the different systems and to structure the knowledge to simplify the analysis and the synthesis of the knowledge of a domain. Building on ontologies, the development of the Semantic Web opens up new possibilities and challenges to the design of a generation of adaptive systems, making it possible to model users' profiles and contexts. As part of the adaptation of learning pathways to the learner, ontology has become an unavoidable solution. It allows the construction of complex knowledge models that can be used to model both the users of the system, their context and the field of application in an intelligible way.

In the semantic knowledge layer of our proposal, and for reasoning purposes, only one ontology is considered that we call domain ontology of m-Learning. However, we distinguish two subparts of this ontology: (1) the LO model and (2) the context model. In the next section, we begin by describing the model of LO.

3.2 Replication data

In this section, we present a context-guided replication model for mobile RDF data.

Replication of semantic web data on mobile enables applications and services to operate independently of the quality of network connections. Nevertheless, this replication must take into account the limited resources of mobiles such as storage capacity. For this, a selective replication is desirable in order to make available to users only useful data. The selection process is based on contextual information about the user and their environment. Let's begin by illustrating the context of which we can find several definitions in the literature. In general, context can be defined as any information that can be used to characterize the situation of an entity. Another definition describes the context as anything surrounding a user or device and gives meaning to something. Our replication model helps ensure high data availability. In this model we propose to formulate the contextual information in the form of RDF data, to record them in triplestores and to apply reasoning based on the rules of inference. Costly processes in terms of energy consumption such as the evaluation of contextual information as well as the selection of partial graphs run on the clone (to unload the mobile). The mobile is only able to acquire contextual information and send it to its clone without any additional processing.

4 Evaluation

In this section, we first introduced the configuration, test environment, and datasets used to evaluate our system based on the combination of Semantic Web and Big Data technologies applied to the m-Learning domain. This system has been tested on real databases, and benchmarks of performance comparison according to the function of relevance and the response time have been achieved. In order to show the efficiency of our system compared to existing systems, we compared our system with the two existing solutions: RDF-3X, and Rya. The results illustrated in the figures above show the robustness and efficiency of our system whatever the volume of data.

To evaluate our system we use Open University Courses[29] dataset of 54,584,125 triples, the second dataset is Coursera MOOC with a size of 4,927,697 triples, and LUBM [30] Benchmark and to show the scalability of our system our latest test are to realize on the Latest Wikidata [31] its size is 69GB and 7.2 Billion triples.

4.1 Evaluation using LUBM benchmark, open university, and coursera MOOC

Tables 1, 2, and 3 below show the execution times for LUBM Benchmark queries on the three LUBM Benchmark dataset instances created.

Table 1. Execution time using LUBM1000

LUBM1000	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
RDF-3X	2,5	1,95	1,1	1,87	3,6	3,13	3,51	4,4	6,36
Rya	2,69	1,4	0,98	2	3,4	2,7	4,05	3,2	5,1
Sem-Learning	1,7	2,2	0,73	1,52	2,91	2,4	2,63	2,4	4

Table 2. Execution time using LUBM2000

LUBM2000	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
RDF-3X	11,3	8,45	4,7	14,9	7,11	25,1	10,3	12	16,4
Rya	9,5	12	6,2	11,4	5,8	18,6	9,5	8	13,3
Sem-Learning	10,4	6,1	3,9	11,7	3,2	14	8,7	7,6	10,5

Table 3. Execution time using LUBM5000

LUBM5000	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
RDF-3X	430	2852	560	543	1587	1850	321	506	5276
Rya	308	2707	311	452	1305	996	376	467	3741
Sem-Learning	165	2164	176	265	1387	698	188	345	2870

The following table presents the results obtained from loading time for the five datasets: LUBM1000, LUBM2000, LUBM5000, Open University, and Coursera MOOC.

Table 4. Datasets loading time.

LUBM1000			LUBM2000			LUBM5000			Open University			Coursera MOOC		
RDF-3X	*Rya*	*Sem-Learning*	*RDF-3X*	*Rya*	*Sem-Learning*	*RDF-3X*	*Rya*	*Sem-Learning*	*RDF-3X*	*Rya*	*Sem-Learning*	*RDF-3X*	*Rya*	*Sem-Learning*
43	17	11	48	18	11	62	21	16	51	20	15	46	18	10

The following figure 4 presents the results of the execution of the three systems RDF-3X, Rya, and our system, to better evaluate and compare these three systems we used the five datasets.
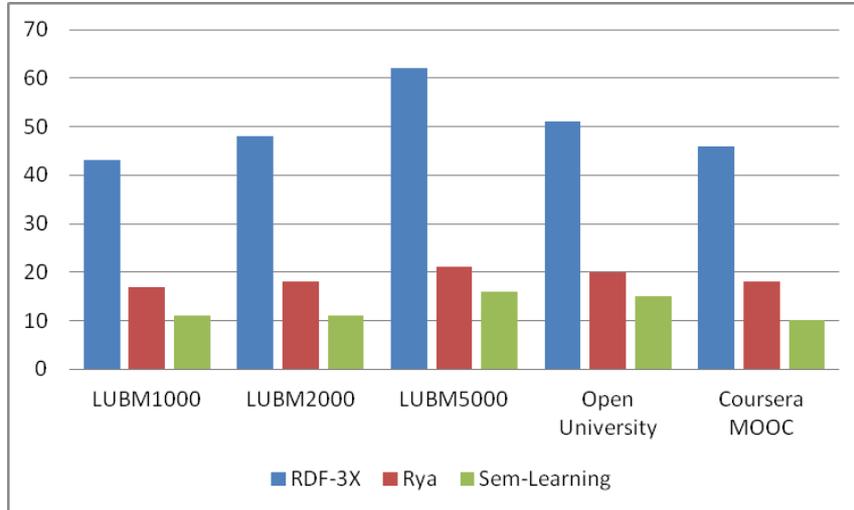
Fig. 4. RDF Data Load Time.

The data loading results show the efficiency of our system compared to the RDF-3X and Rya, for all data-sets tested, thanks to MongoDB which allows loading the data in a fast way compared to others Big Data system[32]. The following figure graphically illustrates the results in the previous table.
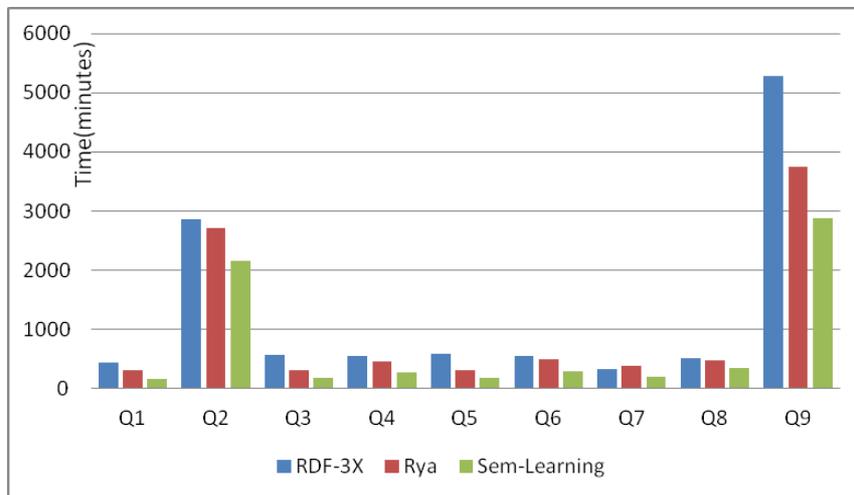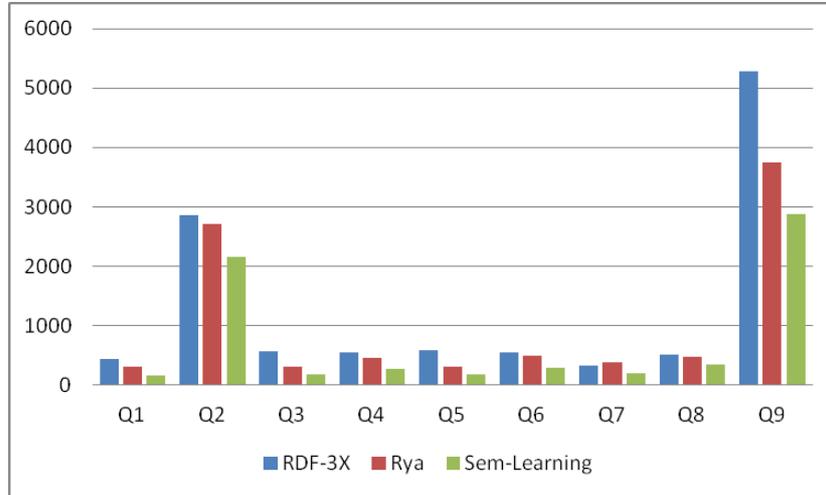

Fig. 5. LUBM1000 Execution Time.
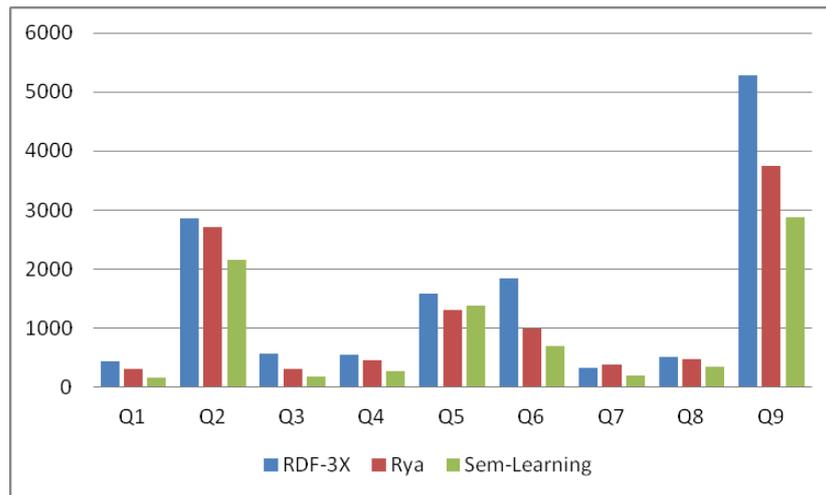
Fig. 6. LUBM2000 Exection Time.

Fig. 7. LUBM5000 Execution Time.

4.2 Evaluation using the latest wiki data

Our latest tests are made using the Latest Wiki Data which contains more than 69GB and 7.200.000.000 triples, we have executed the following 5 queries and the results obtained are illustrated in figure 7.

Q1

```
1 SELECT ?item ?itemLabel
2 WHERE
3 {
4   ?item wdt:P31 wd:Q146.
5   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
6 }
```

Q2

```
1 SELECT ?item ?itemLabel
2 WHERE
3 {
4   ?item wdt:P31 wd:Q2934.
5   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
6 }
```

Q3: This query retrieves Mathematical proofs

```
SELECT ?proof ?proofLabel
WHERE
{
  ?proof wdt:P31 wd:Q11538.
  SERVICE wikibase:label { bd:serviceParam wikibase:language "en". }
}
```

Q4: Query 4 gives a map of hospitals.

```
1 #added 2017-08
2 #defaultView:Map
3 SELECT * WHERE {
4   ?item wdt:P31*/wdt:P279* wd:Q16917;
5         wdt:P625 ?geo .
6 }
```

Q5: Query 5 retrieves the Number of humans in Wikidata Using the count function

```
SELECT (COUNT(?item) AS ?count)
WHERE {
    ?item wdt:P31/wdt:P279* wd:Q5 .
}
```

The following figure illustrates the execution results of these 5 queries in (ms), we notice that the queries 1 and 4 take more time compared to the other requests this time deference is justified by the joining of these two requests which requires more time to join the recovered data. The query 5 of the count is executed quickly since it returns only a number.

Fig. 8. Execution time using the latest Wikidata

5 Discussion

This M-Learning system allows you to recommend courses to learners according to their profiles, thanks to the use of M-Learning domain ontology and a knowledge server, all data is stored in a single format which is the triple RDF. After evaluating our system using datasets that contain large volumes of RDF data, we noticed the efficiency of our approach if the amount of RDF data is very large, as well as an optimal recommendation for learners according to their interests and their relationships through the benefits of RDF graphs and vocabularies like FriendOfFriend. The predicates of RDF triples are properties that express the relationships between subjects and objects of triples, in our system the properties of relations like reflexivity, transitivity and other are very simple to represent and store them using the M-Learning domain ontology because ultimately these data from this ontology are transformed and stored as triple RDFs. The efficiency of our approach will be reached when it is going to be used by a large number of learners, teachers, and others, and also when the database of courses and user information increases through scalable management provided by MongoDB, this document-oriented NoSQL database gives us scalability, partition tolerance, and high data availability for that it is considered the most used NoSQL database management system in the world. Another strong point of our system that is replication, the replication of semantic web data on mobiles allows applications and services to work independently of the quality of the network connections. Nevertheless, this replication must take into account the limited resources of mobiles such as storage capacity. For this, a selective replication is desirable in order to make available to users only useful data. The selection process is based on contextual information about the user and their environment. Finally, in this paper, we presented our approach for a recommendation system applied to the field of mobile-learning combining semantic Web technologies based on MongoDB database management system.

The architecture we propose is composed of two main parts: a knowledge server and an adaptation model. The knowledge server allows trainers and business experts to represent their expertise using business rules and ontology to ensure heterogeneous knowledge. Then, in a mobility environment, the knowledge server makes it possible to take into account the constraints of the environment and the user constraints. The adaptation model corresponds to combinatorial optimization algorithms, whose objective is to propose to the learner a relevant combination of LO according to its contextual constraints.

6 Conclusion

The recommendation in the field of M-Learning is a very important tool for suggesting courses to the learner, in this paper we presented an evolutionary recommendation system based on Semantic Web technologies; RDF, Ontology, and SPARQL, the advantage of this system is robustness, scalability and fault tolerance since it is based at the data management level on a Big Data tool named MongoDB, this NoSQL database will be used to store RDF data in JSON document format.

7 References

[1] Simpson, Ormond. Supporting students in online, open and distance learning. Routledge, 2018.

[2] Briz-Ponce, Laura, Anabela Pereira, Lina Carvalho, Juan Antonio Juanes-Méndez, and Francisco José García-Peñalvo. "Learning with mobile technologies–Students' behavior." Computers in Human Behavior 72 (2017): 612-620. https://doi.org/10.1016/j.chb.2016.05.027

[3] Sulisworo, Dwi, Ishafit Ishafit, and Kartika Firdausy. "The Development of Mobile Learning Application using Jigsaw Technique." iJIM 10, no. 3 (2016): 11-16. https://doi.org/10.3991/ijim.v10i3.5268

[4] Mouad Banane, and Abdessamad Belangour. « RDFMongo: A MongoDB Distributed and Scalable RDF management system based on Meta-model». International Journal of Advanced Trends in Computer Science and Engineering 8, nᵒ 3 (2019): 734 – 741. https://doi.org/10.30534/ijatcse/2019/62832019

[5] Mouad Banane, and Abdessamad Belangour. « An Evaluation and Comparative study of massive RDF Data management approaches based on Big Data Technologies». International Journal of Emerging Trends in Engineering Research. 7, nᵒ 7 (2019): 48 – 53. https://doi.org/10.30534/ijeter/2019/03772019

[6] Ghazi, Mohd Rehan, and Durgaprasad Gangodkar. "Hadoop, MapReduce and HDFS: a developer's perspective." Procedia Computer Science 48 (2015): 45-50. https://doi.org/10.1016/j.procs.2015.04.108

[7] Chodorow, Kristina. MongoDB: the definitive guide: powerful and scalable data storage. "O'Reilly Media, Inc.", 2013.

[8] Wu, Buwen, Yongluan Zhou, Pingpeng Yuan, Ling Liu, and Hai Jin. "Scalable SPARQL querying using path partitioning." In 2015 IEEE 31st International Conference on Data Engineering, pp. 795-806. IEEE, 2015. https://doi.org/10.1109/icde.2015.7113334

[9] Mouad Banane, and Abdessamad Belangour. « Querying massive RDF data using Spark». International Journal of Advanced Trends in Computer Science and Engineering 8, nᵒ 4 (2019): 1481 - 1486. https://doi.org/10.30534/ijatcse/2019/68842019

[10] Zaharia, Matei, Reynold S. Xin, Patrick Wendell, Tathagata Das, Michael Armbrust, Ankur Dave, Xiangrui Meng et al. "Apache spark: a unified engine for big data processing." Communications of the ACM 59, no. 11 (2016): 56-65. https://doi.org/10.1145/2934664

[11] Mouad Banane and Abdessamad Belangour, "New Approach based on Model Driven Engineering for Processing Complex SPARQL Queries on Hive" International Journal of Advanced Computer Science and Applications(IJACSA), 10(4), 2019. https://doi.org/10.14569/ijacsa.2019.0100474

[12] Du, Dayong. Apache Hive Essentials: Essential techniques to help you process, and get unique insights from, big data. Packt Publishing Ltd, 2018.

[13] Banane M., Belangour A., El Houssine L. (2019) Storing RDF Data into Big Data NoSQL Databases. In: Mizera-Pietraszko J., Pichappan P., Mohamed L. (eds) Lecture Notes in Real-Time Intelligent Systems. RTIS 2017. Advances in Intelligent Systems and Computing, vol 756. Springer, Cham. https://doi.org/10.1007/978-3-319-91337-7_7

[14] Banane M., Belangour A. (2019) A Survey on RDF Data Store Based on NoSQL Systems for the Semantic Web Applications. In: Ezziyyani M. (Ed.): Advanced Intelligent Systems for Sustainable Development. AI2SD 2018. Advances in Intelligent Systems and Computing, vol 915. Springer, Cham. https://doi.org/10.1007/978-3-030-11928-7_40

[15] El-Seoud, Samir Abou, Hosam El-Sofany, and Islam Taj-Eddin. "Mobile Applications and Semantic-Web–A case study on Automated Course Management." International Journal of Interactive Mobile Technologies (iJIM) 10.3 (2016): 42-53. https://doi.org/10.3991/ijim.v10i3.5770

[16] Vesin, Boban, Aleksandra Klašnja-Milićević, and Mirjana Ivanović. "Application of Semantic Web Technologies to Facilitate Use of E-Learning System on Mobile Devices." Smart Education and e-Learning 2016. Springer, Cham, 2016. 473-484. https://doi.org/10.1007/978-3-319-39690-3_42

[17] Hitzler, Pascal, Aldo Gangemi, and Krzysztof Janowicz, eds. Ontology Engineering with Ontology Design Patterns: Foundations and Applications. Vol. 25. IOS Press, 2016.

[18] Krishnan, Kalyani, Reshmy Krishnan, and Ayyakannu Muthumari. "A semantic-based ontology mapping–information retrieval for mobile learning resources." International Journal of Computers and Applications 39.3 (2017): 169-178. https://doi.org/10.1080/1206212x.2017.1309223

[19] El-Seoud, S.A., El-Sofany, H. and Karam, O., 2015. Semantic Web Architecture and its Impact on E-learning Systems Development. International Journal of Emerging Technologies in Learning (iJET), 10(5), pp.29-34. https://doi.org/10.3991/ijet.v10i5.4754

[20] Bakhouyi, Abdellah, Rachid Dehbi, Mouad Banane, and Mohamed Talea. "A Semantic Web Solution for Enhancing The Interoperability of E-learning Systems by Using Next Generation of SCORM Specifications." International Journal of Emerging Technologies in Learning 14, no. 11 (2019). https://doi.org/10.3991/ijet.v14i11.10342

[21] Ruano, Ildefonso, Pablo Cano, Javier Gámez, and Juan Gómez. "Advanced LMS integration of SCORM Web laboratories." IEEE Access 4 (2016): 6352-6363. https://doi.org/10.1109/access.2016.2587805

[22] Punnoose, Roshan, Adina Crainiceanu, and David Rapp. Rya: a scalable RDF triple store for the clouds. Proceedings of the 1st International Workshop on Cloud Intelligence. ACM, 2012. https://doi.org/10.1145/2347673.2347677

[23] Neumann, Thomas, and Gerhard Weikum. The RDF-3X engine for scalable management of RDF data.The VLDB Journal—The International Journal on Very Large Data Bases 19.1: 91-113. https://doi.org/10.1007/s00778-009-0165-y

[24] Ma, Zongmin, Miriam AM Capretz, and Li Yan. "Storing massive Resource Description Framework (RDF) data: a survey." The Knowledge Engineering Review 31, no. 4 (2016): 391-413. https://doi.org/10.1017/s0269888916000217

[25] Peng, Peng, Lei Zou, M. Tamer Özsu, Lei Chen, and Dongyan Zhao. "Processing SPARQL queries over distributed RDF graphs." The VLDB Journal—The International Journal on Very Large Data Bases 25, no. 2 (2016): 243-268. https://doi.org/10.1007/s00778-015-0415-0

[26] DB-Engines. https://db-engines.com. Accessed: 2019-02-13.

[27] Afsari, Kereshmeh, Charles M. Eastman, and Daniel Castro-Lacouture. "JavaScript Object Notation (JSON) data serialization for IFC schema in web-based BIM data exchange." Automation in Construction 77 (2017): 24-51. https://doi.org/10.1016/j.autcon.2017.01.011

[28] Kellogg, Gregg, and Pierre-Antoine Champin. "JSON-LD 1.1–A JSON-based Serialization for Linked Data (W3C Working Draft)." (2019).

[29] Jordan, Katy. "Massive open online course completion rates revisited: Assessment, length and attrition." The International Review of Research in Open and Distributed Learning 16, no. 3 (2015). https://doi.org/10.19173/irrodl.v16i3.2112

[30] Guo, Yuanbo, Zhengxiang Pan, and Jeff Heflin. "LUBM: A benchmark for OWL knowledge base systems." Web Semantics: Science, Services and Agents on the World Wide Web 3.2-3: 158-182. https://doi.org/10.1016/j.websem.2005.06.005

[31] Latest Wikidata, available on http://www.rdfhdt.org/datasets/

[32] Gandomi, Amir, and Murtaza Haider. "Beyond the hype: Big data concepts, methods, and analytics." International journal of information management 35, no. 2 (2015): 137-144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

8 Authors

Mouad Banane is a PhD at the Faculty of Sciences Ben M'Sik, Laboratory of Information Technology and Modeling (LTIM), in Hassan II University of Casablanca, Morocco. His research fields: Model Driven Engineering, Semantic Web, E-learning, and Big Data.

Abdessamad Belangour is an associate professor at the Faculty of Science Ben M'Sik, and member of LTIM laboratory (Laboratory of Information Technology and Modeling). He is responsible of research strategies and doctoral education training process at Hassan II University of Casablanca. His work focuses on E-learning, M-learning, Model Driven Engineering, Database Security, Big Data,and Internet of Thing