A Semantic Repository Approach to Improve Health Information Management

Jose Gracia-Maldonado[1], **Francisco P. Romero**[2] **and Jose A. Olivas**[2]
[1]Indra-Software Labs, Indra-UCLM Research Center, Ciudad Real, Spain
[2]Dept. Information Technologies and Systems, University of Castilla La Mancha, Ciudad Real, Spain

Abstract— *Nowadays healthcare organization have to daily manage large amounts of information, an optimum use of this is increasingly difficult. This is especially significant when it comes to large and international contexts where the concepts and terminology handled in many cases, are not completely unified. The use of ontologies is widely presented as a solution to this problem because they allow the formal modeling of knowledge within a given domain. However, their implementation in real systems is not yet meaningful because they are not easy to use by non experts and it requires a great experience and training for proper understanding. In this paper we present ASMOR, a system to facilitate the task of organization, retrieval of key information in a healthcare context. The proposed prototype allows the user to monitor in a visually and fully automatic way the evolution of contents stored in a semantic repository.*

Keywords: Ontology, Healthcare Information Management, Semantic Repositories

1. Introduction

Ontologies have proven to be successful in handling a machine processable representation of information and have been used to model domain contexts in several real world applications [1], [2]. They allow capturing domain knowledge such as products, services, markets, etc. in an explicit and formal way such that it can be shared among human and computer systems [3]. An ontology can take the simple form of a taxonomy (i.e., knowledge encoded in a minimal hierarchical structure) or as a vocabulary with standardized machine interpretable terminology supplemented with natural language definitions. Ontology is often specified in a declarative form by using semantic markup languages such as RDF and OWL [4].

The notion of ontology is becoming very useful in various fields such as intelligent information extraction and retrieval, cooperative information systems, electronic commerce, and knowledge management [5]. In nowadays, it is important to deal with the concept of Information Filtering, through which a subset of relevant documents is quickly and effectively selected for further detailed analysis. Information Filtering has emerged with the aim of dynamically adapting the distribution of information where both evolving user's interests and new incoming information are taken into account. Although information retrieval and information filtering share the goal of selecting documents whose content matches a particular information need, they are different in some aspects. Information retrieval systems are designed for supporting users with short term interests (query), and retrieval refers to information selection from relatively static information collections or databases. In contrast, filtering refers to information selection from a dynamic stream of incoming data on the basis of user's long term interests (profile). As a result of their common goal, the techniques applied to information retrieval have been also applied to information filtering task [6].

Although ontologies are used in a wide range of applications and have been instrumental in many interoperability projects, they have so far had only limited success [7]. Early semantic search engines tried to use ontology concepts and structures as controlled search vocabularies, but this was unpractical both functionally and from a usability perspective. Another bottleneck in developing ontology-based systems stems from the fact that the conceptual formalism that supports by common ontologies may not be sufficient to represent uncertainty that is commonly found in many application domains [8]. Moreover, most of ontologies are application-specific and its construction can be a long and costly task. The use of conceptual structures to provide users access useful information requires that there is something bridging the gap between the conceptual and the real world.

In this paper, the tool **ASMOR** (**A**lert **M**anagement System for **O**ntologies **R**epositories) is presented. This tool is a prototype that allows the user to perform a dynamic management of the contents stored in a complex Information System in an easy and intuitive way. The presented work differs from previous approaches that exploit ontologies and SPARQL Queries in a transparent way to the user. The main features of this tool are: (i) Useful and usable in daily use, (ii) Intelligent inbox management, (iii) Visual Interaction (alerts and results graphs) (iv) Flexibility and modularity (it can be easily applied to several domain fields).

The rest of the paper is organized as follows. Section 2 describe the main components of the tool, meanwhile, in section 3, the high-level architectural design of ASMOR and its key functionalities are presented. Section 4 describes some demonstration scenarios, and finally, some conclusions and future works are pointed out in section 5.

2. Components

The use of a publish/subscribe (pub/sub) system can be considered relevant to carry out an active management of semantic repositories. A pub/sub receives information from data sources and notifies users if the messages match the subscriptions criteria. Typically, the publish/subscribe system contains three roles: (i) event providers (ii) event consumers, who subscribe events that they interested, and (iii) the event broker, who is responsible for routing the events from the publishers to subscribers [9]. In our context the event providers are the sources of health information about patients, the event consumers are the professionals, and ASMOR is a tool that allows the user to do an automatic, intuitive and visual event broker with the aim of monitoring of the evolution of the contents stored in the repository.

Two main components are the basis of ASMOR tool: the alerts and the ontological model.

2.1 Alerts

An alert is a time saving tool that makes the user aware that a new information of interest has been made available, eliminating the need for constant monitoring of the information repository. In this context, alerts are useful to help the user to keep up-to-date and stay aware of changes in your patients or on new patients in a certain conditions. Automated alerts have recently emerged as a major instrument to influence clinician behavior. In the hospital setting, randomized trials have shown the efficacy of alerts [10].

The alerts systems is feed using the knowledge stored in a semantic repository. An alert means that the underlying data is changing, thus, ASMOR is not only a tool to evaluate the existing data set to find alerting case, it is a tool to monitorize the changes in the contents of the data repository over the time. This process is carried out by defining a set of alerts that will be reported to the user according to their priority.

There are a few different components that make up an alert definition, they include:

- *Priority*: The priority is used for ordering notifications of different types for the same user.
- *Criteria*: The filtering criteria are focused on: entities, entities relationship, and entities attributes.
- *Graphical Representation*: ASMOR offers a user-friendly interface to assist users in the definition of their criteria in a graphical and intuitive way.
- *Query*: ASMOR translates the specified criteria to a SPARQL query. This process is transparent to the user.
- *Recommendations/actions*: It is possible to define actions to carried out when the notification is generated.

Alert descriptions are usually defined by the Semantic Repository Administrator, but the user also have the capacity to create their own alert definitions, that includes building their own criteria, prescribing their own recommendations, and assigning actions.

2.2 Ontological Model

Due to the diversity of such information sources and the increasing amounts of data produced nowadays, it is necessary to build a semantic model that describes the interchanged knowledge in the information system. The model is useful for the system to discover contents and users correlation.

Although there is not a universal consensus on the definition of ontology, it is generally accepted that ontology is an explicit specification of a conceptualization [11]. A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly.

Ontology provides a number of potential benefits in representing and processing knowledge, including the separation of domain knowledge from application knowledge, sharing of common knowledge of subjects among human and computers, and the reuse of domain knowledge for a variety of applications [12]. There are several ontologies found in the literature created for it use in a particular knowledge domain.

In this work, the proof on concept of the proposed tool has been made using the healthcare information expressed on the CCR (Continuity of Health Record) standard [13]. The Continuity of Care Record standard contains information such as patient demographics, insurance and health care provider information, medication lists, allergies and recent medical procedures [14].

The proposed model has been developed in Web Ontology Language (OWL) and is an updated version of the one presented in [15]. Several studies has been demonstrated the utility of this model for semantic interoperability [16]. An synthetic of the proposed ontological model can be seen in Figure 1.

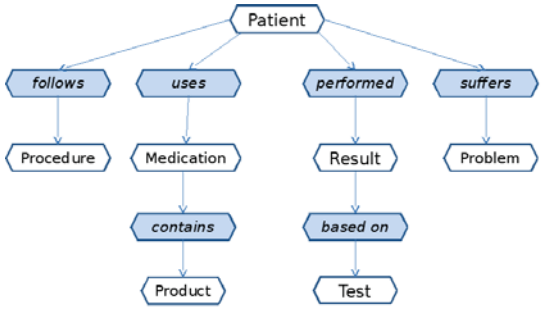

Fig. 1: Synthetic view of the Ontological Model

The modeling process has been carried out using Protege [1] and the repository has been implemented using Sesame OpenRDF [2].

[1] http://protege.stanford.edu/
[2] http://www.openrdf.org/index.jsp

3. Architecture

A global overview of the context and environment of ASMOR tool can be seen in Figure 2. The identified flows of information in the system are the following:

1) Healthcare data is generated at many points in a healthcare organization. Much of the data generated is patient oriented. This data is integrated in a CCR format which is the format in which health information is stored and exchanged across systems.
2) Information from health centers is translated to a RDF format and stored in the semantic repository.
3) The more relevant alerts to the healthcare organization are defined by the administrator.
4) Notifications will be collected automatically in the user's inbox. They were generated by the system when the criteria were verified.

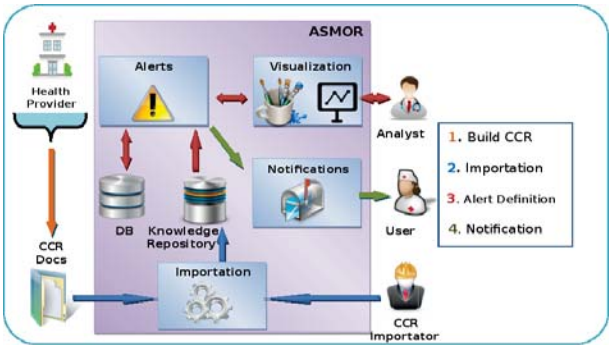

Fig. 2: Global View of ASMOR tool

The monitoring process is carried out by means the user-defined set of alerts. These alerts are reported to the user's inbox according theirs contents and their priority. For each alert, the user can define, in a completely visual way, a set of criteria that, when one is verified, it will generate and automatic end user notification. After this notification the users can interact visually with the content that triggers the alert.

ASMOR has been built according to a modular architecture that comprises two key modules acquisition and exploitation as is detailed in the the following subsections (Figure 3).

3.1 Acquisition

The ontological repository is able to integrate and merge information from different sources in a heterogenous format. Therefore, the content provided by the information sources is transformed to RDF triplets in order to be stored in the semantic repository.

Fig. 3: Asmor Architecture

This procedure is carried out in a three-step process(Figure 4):

1) Knowledge instances are extracted from the documents in a specific format.
2) RDF representation of these knowledge instances.
3) Storing the information in the ontological repository.

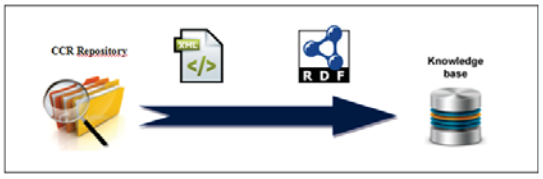

Fig. 4: Acquisition Process

The transformation rules that are used in the CCR to RDF/OWL algorithm are the following:

- Complete elements are transformed into OWL classes.
- Simple elements are transformed into data properties.
- The complex relationships between elements are transformed into class-class relationships, that is, in OWL object properties (see Table 1).

Table 1: OWL object properties.

Property	Source	Target
Bases_on	Result	Test
Contains	Medication	Product
Follows	Patient	Procedure
Performed	Patient	Result
Uses	Patient	Medication
Suffers	Patient	Problem

In compliance with the transformation rules explained above, the second part of the process takes care of transforming CCR to RDF. For this purpose, it is necessary to transform XML data to RDF/OWL by creating instances of the necessary OWL classes, RDFS datatypes, OWL datatype and object properties.

3.2 Exploitation

The knowledge exploitation has made by three modules: alert, evaluation and notification.

3.2.1 Alert Module

In terms of volume, a Healthcare Knowledge Base (KBs) could contains millions of entities and facts about patients, problems and procedures . In this case, such information is stored in RDF format and queried with the SPARQL language. Large KBs are difficult to use as their schema (the ontological model) and its underlying semantics are rarely understood by end users.

ASMOR proposes a graphical interface to construct a SPARQL query used to monitoring the content evolution in a RDF repository. The "Alert Module" allows the user to define alerts in a completely visual way. The visualization component, based on JIT (Javascript InfoVisualization Toolkit [3]) framework, has been used for this purpose. The browser panel displays the ontology as a graph. As an ontology can be large, a search box with auto completion feature is available for finding classes and properties that will be used in the query. When a class or property is selected, the graph is centered around this concept to see all the related concepts. The visual construction of a query is composed of three main steps (similar to the proposed in [17]):

1) Selecting classes and/or properties from the semantic model browser into the query panel.
2) Linking the classes and properties defined in the previous step.
3) Adding query operators by selecting the components of the query on which they must be applied.

An example of alert in a medical ontological repository can be seen in Figure 5.

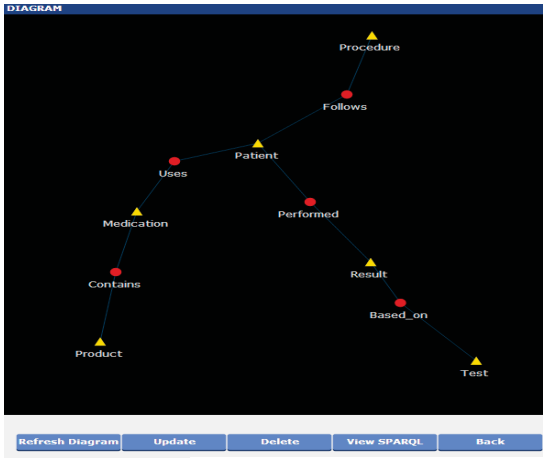

Fig. 5: Example of alert definition

[3]http://philogb.github.io/jit/

Each alert is internally translated to a SPARQL query as Figure 6 shown.

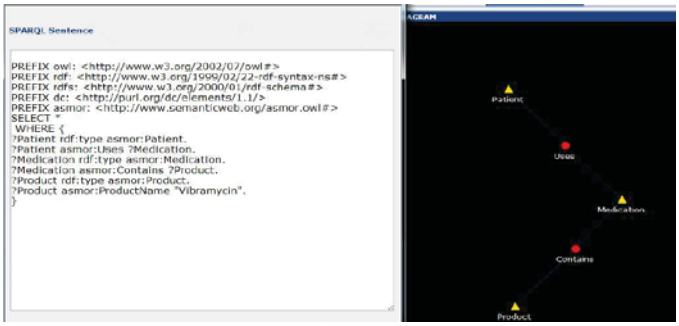

Fig. 6: SPARQL Query

3.2.2 Evaluation Module

The evaluation module consists of those features that allow ASMOR system evaluating the active alert criteria in a given period, in a completely automatic way. The module provides the necessary algorithms for the analysis and evaluation of every alert defined in the system. The tool provides a mechanism to automatically run the analysis algorithms in the background. The management of the results generated by the SPARQL queries is also part of it.

The tool evaluates only active alert criteria that are also associated with active alerts in the system. These alert criteria are executed automatically and evaluated on the semantic repository. If a criterion is satisfied based on the stored information, an action to be performed is determined. The event dispatcher is the component responsible for collecting definitions and forwarding alert to the users. The architecture of the event dispatcher in this tool is centralized.

3.2.3 Notification Module

The funcionality of the notification module consists of managing and visualizing the different "inboxes" which receive alert notifications. An example is shown in Figure 7, the notification "inbox" can be observed with two notifications that are pending to be examined.

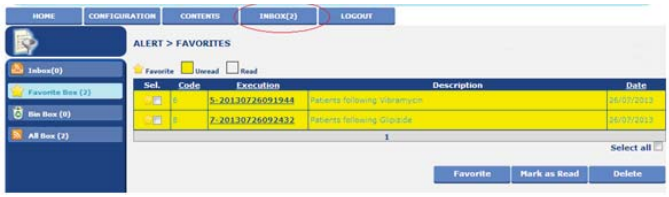

Fig. 7: Example of User Notification

One the notification is read, the tool display a graph with the results that have been activated according to the alert criteria. The user can visualize and interact with these results (Figure 8).

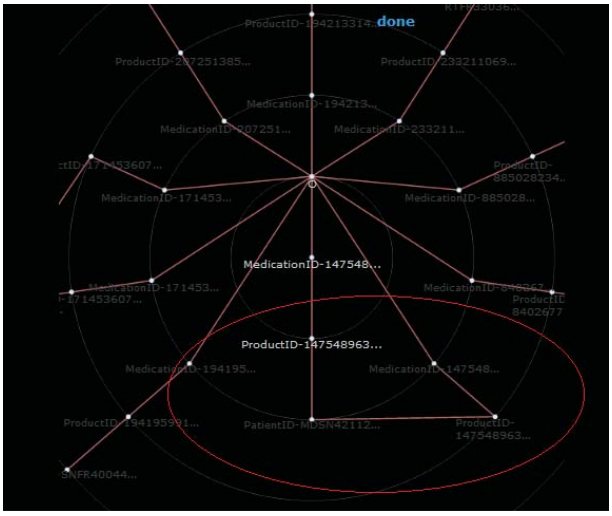

Fig. 8: Example of Result

4. Demonstration Scenarios

Three scenarios has been defined to demonstrate the feasibility of the proposed tool (Figure 9). This demonstration will include groups of cases with the aim of showing all the capabilities of the system in three different contexts:

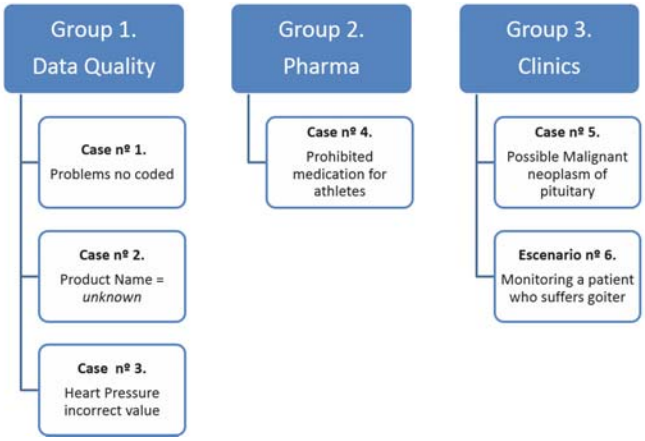

Fig. 9: Considered Cases

- *Data Quality*: This scenario illustrates the use of the tool for detecting erroneous or inconsistent data. The type of user involved in this scenario is the *Health Documentalist*. This user is responsible for ensuring the quality of data managed by the e-Health systems. In this scenario, the user can monitorize the quality of the data that would feed the semantic repository.
- *Pharma*: This scenario illustrates the use of the tool for detecting medications that containing products considered illegal or whose use must be under control. The role involved in this scenario is the *Pharmacist*, i.e., the responsible for ensuring that medication to be supplied to patient is under current legislation.
- *Clinics*: This scenario illustrates the use of the tool to define associated alerts to a number of symptoms that have high probability of being evidence of serious illness. The user involved in this scenario is the *Doctor*. This user is responsible for analyze the possibility of having a serious illness from severe detected symptoms. For example, those patients whose medical history concur problems such *Hypercalcemia* and *Cyst and pseudocyst of pancreas* has a high risk of the Wermer syndrome (also known MEM Type 1). These cases would have to request a Nuclear Magnetic Resonance (NMR) to rule out a pituitary tumor.

Fig. 10: Example of Alert

The goal of the demo is to show how ASMOR can be used to manage healthcare information in a very efficient manner. In absence of real world data for evaluating the tools, real anonimyzed data has been used during the testing process.

During the demonstration, the focus will be on showing the details of the process involved, starting from alert definition to the visualization of the notification results. During alert definition, the graphical user interface will assist users to issue or refine the alert query by adding, editing, or deleting semantic links among entities. Upon retrieval of the notification related data, users will view the graphical representation of the results. Also, they could be able to organized them in a number of different ways.

5. Conclusions and Future Work

A prototype of semantic repository management tool based on a graphical way for query management and results visualization is presented. The proposed tool, ASMOR, is an example of advanced decision support system based on semantic models. The basic idea is to allow the user to define alerts by identifying entities and values which are detected together and to derive a notification if they occur in the repository.

The system core is characterized by the fact that the graphically obtained SPARQL query is submitted to a ontological repository that stores information from heterogeneous data sources.

The prototype has been developed based on open source standards and following a modular design and architecture, therefore, it has given him the flexibility to be perfectly used for any ontology. This tool can be used as an alert

and notification system in other areas of expertise such as e-learning if they have a related ontological model like the proposed in [18].

The prototype has been tested in different scenarios within the field of healthcare in collaboration with medical professionals. After these tests, it can be concluded that all the system features were accepted. Therefore, it has demonstrated the feasibility of building a tool that allows users to monitor visual, intuitive and automatic evolution of those contents represented in an ontological model for proper management of them on a active way.

ASMOR features are being extended to support a machine learning module to automatically define complex intelligent rules for content management.

References

[1] L. Han and G. Chen, "A fuzzy clustering method of construction of ontology-based user profiles," *Advances in Engineering Software*, vol. 40, no. 7, pp. 535–540, 2009.

[2] M. Golemati, A. Katifori, C. Vassilakis, G. Lepouras, and C. Halatsis, "Creating an ontology for the user profile: Method and applications," in *In Proceedings of the First International Conference on Research Challenges in Information Science (RCIS)*, 2007.

[3] R. Lau, "Fuzzy Domain Ontology Discovery for Business Knowledge Management," *IEEE Intelligent Informatics Bulletin*, vol. 8, no. 1, pp. 29–41, 2007.

[4] D. L. McGuinness and F. van Harmelen, "Owl web ontology language overview," W3C, Tech. Rep. REC-owl-features-20040210, 2004.

[5] A. Sieg, B. Mobasher, and R. Burke, "Web search personalization with ontological user profiles," in *CIKM '07: Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*. New York, NY, USA: ACM, 2007, pp. 525–534.

[6] R. Kassab and J.-C. Lamirel, "Towards a synthetic analysis of user's information need for more effective personalized filtering services," in *SAC '07: Proceedings of the 2007 ACM symposium on Applied computing*. New York, NY, USA: ACM, 2007, pp. 852–859.

[7] G. Solskinnsbakk and J. A. Gulla, "Combining ontological profiles with context in information retrieval," *Data Knowl. Eng.*, vol. 69, no. 3, pp. 251–260, 2010.

[8] Q. T. Tho, S. C. Hui, and T. H. Cao, "FOGA: A Fuzzy Ontology Generation Framework for Scholarly Semantic Web," in *Proceedings of the 2004 Knowledge Discovery and Ontologies Workshop (KDO 2004)*.

[9] F. Tian, B. Reinwald, H. Pirahesh, T. Mayr, and J. Myllymaki, "Implementing a scalable xml publish/subscribe system using relational database systems," in *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '04. New York, NY, USA: ACM, 2004, pp. 479–490.

[10] "The use of patient and physician reminders for preventive services: results from a national study of physician organizations," *Preventive Medicine*, vol. 39, no. 5, pp. 1000 – 1006.

[11] T. R. Gruber, "A translation approach to portable ontology specifications," *Knowl. Acquis.*, vol. 5, no. 2, pp. 199–220, 1993.

[12] K. Dramé, G. Diallo, F. Delva, J. F. Dartigues, E. Mouillet, R. Salamon, and F. Mougin, "Reuse of termino-ontological resources and text corpora for building a multilingual domain ontology: An application to alzheimer's disease," *Journal of Biomedical Informatics*, vol. 48, no. 0, pp. 171 – 182, 2014.

[13] M. Botsivaly, B. Spyropoulos, K. Koutsourakis, and K. Mertika, "Enhancing continuity in care: An implemantation of the ASTM E2369-05 standard specification for continuity of care record in a homecare application," in *AMIA 2006, American Medical Informatics Association Annual Symposium, Washington, DC, USA, November 11-15, 2006*. AMIA, 2006.

[14] P. Plastiras, D. O'Sullivan, and P. Weller, "An ontology-driven information model for interoperability of personal and electronic health records," in *Proceedings of eTELEMED, The Sixth International Conference on eHealth, Telemedicine, and Social Medicine*, March 2014, pp. 130–133.

[15] J. Puustjärvi and L. Puustjärvi, "Personal health ontology: towards the interoperation of e-health tools," *IJEH*, vol. 6, no. 1, pp. 62–75, 2011.

[16] ——, "Using ontology-based registry and SPARQL engine in searching patient's clinical documents," in *HEALTHINF 2014 - Proceedings of the International Conference on Health Informatics, ESEO, Angers, Loire Valley, France, 3-6 March, 2014*, 2014, pp. 151–158.

[17] G. Fokou, S. Jean, A. Hadjali, and M. Baron, "Qars: A user-friendly graphical tool for semantic query design and relaxation," in *Proceedings of the 18th International Conference on Extending Database Technology, EDBT 2015, Brussels, Belgium, March 23-27, 2015.*, G. Alonso, F. Geerts, L. Popa, P. Barceló, J. Teubner, M. Ugarte, J. V. den Bussche, and J. Paredaens, Eds. OpenProceedings.org, 2015, pp. 553–556.

[18] C. Vidal-Castro and M. F. Satler, "Modelo ontológico para la secuenciación de objetos de aprendizaje," *IEEE-RITA*, vol. 5, no. 2, pp. 63–67, 2010.