

Received August 17, 2020, accepted September 7, 2020, date of publication September 18, 2020, date of current version September 30, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024614

# Web Service to Retrieve and Semantically Enrich Datasets for Theses From Open Educational Repositories

MARÍA AUXILIO MEDINA NIETO<sup>1</sup>, PAULO DANIEL VÁZQUEZ MORA<sup>2</sup>,  
JORGE DE LA CALLEJA MORA<sup>1</sup>, MIREYA TOVAR VIDAL<sup>3</sup>,  
EDUARDO LÓPEZ DOMÍNGUEZ<sup>4</sup>, (Member, IEEE),  
DELIA ARRIETA DÍAZ<sup>5</sup>, AND  
ISMAEL EVERARDO BÁRCENAS PATIÑO<sup>6</sup>

<sup>1</sup>Postgraduate Department, Polytechnic University of Puebla, Puebla 72640, Mexico

<sup>2</sup>Information Technologies Department, Technological University of Puebla, Puebla 72300, Mexico

<sup>3</sup>Faculty of Computer Science, Benemerita Universidad Autónoma de Puebla, Puebla 72570, Mexico

<sup>4</sup>Department of Computer Science, Laboratorio Nacional de Informática Avanzada (LANIA), Xalapa 91090, Mexico

<sup>5</sup>Faculty of Economics, Accountability and Management, Juarez University of the State of Durango, Durango 34000, Mexico

<sup>6</sup>Computing Engineering Department, National University of Mexico, Coyoacan 04510, Mexico

Corresponding author: María Auxilio Medina Nieto (maria.medina@up Puebla.edu.mx)

**ABSTRACT** The paper describes the design and implementation of a semantic web service that retrieves theses and extends the keyword based-search of a DSpace repository taking into account the roles of advisors and steering committee members formally represented into a custom-made ontology. The service uses SPARQL queries and the serialization module of RDF DSpace, this links the item submission process and the ontology, thus the more theses are added into a repository, the more instances are inserted into the ontology. The paper provides empirical insights about how to reuse theses metadata and includes the results of an exploratory and self-management survey of usability heuristic evaluation of a web site that enables to access the proposed service. Heuristics were estimated with a purposive sample of students, teachers, and managers, the results indicated a high satisfaction level and showed that the service increased theses accessibility in the web environment. The service also generates semantically enriched datasets that coexist with the repository, they are of utility and value to educational organizations as they give institutional visibility.

**INDEX TERMS** Educational technology, semantic web, web services, information retrieval.

## I. INTRODUCTION

Scientific and academic communities use repositories, they are technological platforms designed to store and preserve digital documents, a dissemination medium (green path) or a medium to publish the produced contents (golden path); [1]. An open educational repository (OER) or institutional repository (IR) is a set of services offered by educational organizations rendered to the community to gather and manage digital documents of any type through the creation of an open, interoperable, and organized collections that use the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH protocol) intending to ensure visibility and

impact of these organizations. Detailed information about this protocol is described in [2].

The impact of OERs is known and reported in the literature and databases such as the Directory of Open Access Repositories (OpenDOAR) [3], where the percentage of technological platforms for OERs is as follows: DSpace (40%) [7], EPrints (11%) [6], WEKO (8%), this platform has been developed by the National Institute of Informatics (NII), Japan (more information about WEKO is available at: <https://weko.wou.edu.my>), Digital Commons (5%), a cloud-hosted institutional repository software (<https://bepress.com/products/digital-commons/>), islandora (3%), a free open-source software framework designed to manage and discover digital assets (<https://islandora.ca>), CONTENTdm (2%), a “software as a service” (SaaS) platform to manage digital

The associate editor coordinating the review of this manuscript and approving it for publication was Farhana Jabeen Jabeen<sup>1</sup>.

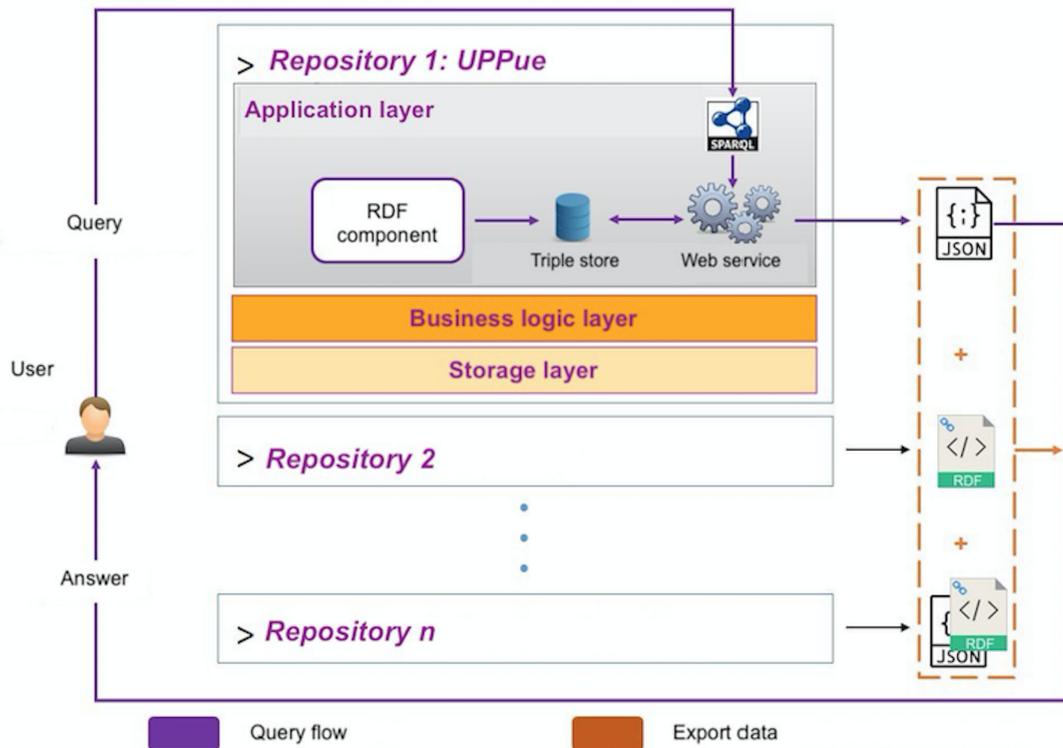


FIGURE 1. Context for SW001 web service.

collections (<https://www.oclc.org/en/content-dm.html>), Open Publication Systems (OPUS) (2%), an open-source software package to create repositories that are compliant with the OAI-PMH protocol, (<https://oapus.nl>), Hyper Articles en Ligne (HAL) (1%), an open archive where authors upload scholarly documents (<http://hal.inria.fr>), dLibra software (1%), a software used to build repositories of digital objects (<https://dingo.psnc.pl/en/dlibra-en/dlibra-funkcje-en/>), while the remaining percentage correspond to tailored software. These platforms have specialized Graphical User Interfaces (GUIs), information retrieval services and the majority use relational models to store metadata [4], [5].

EPrints [6] and DSpace [7] have an open license, so they are widely used in libraries, universities, and cultural heritage organizations, both support the management of digital documents such as master and Ph.D. theses. In Mexico, from the date of writing this paper, the National Repository [8] collect data from 105 open repositories, more than the 90% of them use a version of DSpace.

From the authors' point of view, this paper identifies the following problems related with the management of theses: 1) information retrieval only distinguish between the creator and contributors, that is, this is not possible to establish a role for authors such as the first author, advisor or steering committee members, 2) there is ambiguity in the use of descriptive data (or metadata) during the item submission process, for

example, what kind of information is expected for the contributor element and 3) the interpretation of exported data from repositories is subject to final users and this depends on each educational organization.

Previously, at the Polytechnic University of Puebla (UPPue), an ontology called Onto4AIR had been designed to formally represent domain and operative knowledge for Mexican OERs focusing on documents, users, and their relationships, [9] and a newer version called Onto4UPPue is used for the *semantic web service* described in this paper, from now on, called SW001. This service has been designed to tackle the previous problematic and extends the keyword based-search taking into account the roles of advisors, other steering committee members, and semantic information stored in Onto4UPPue ontology and generates semantically enriched datasets that are downloaded for further analysis. The definition of the ontology concept is adopted from [10]. Figure 1 illustrates the context of SW001.

The feasibility of the service is justified by applying and testing it in the repository of the UPPue university (UPPue-IR) that is supported in the 6.2 version of DSpace, the adaptation to other DSpace repositories will be possible with simple updates. The paper contributions lie in the design of a web site that allows users to access SW001 and a web application to gather data for a heuristic usability testing. We expect that the development methodology for SW001 to be, directly or with slight modifications, reusable by OER managers that seek ways to exploit RDF data enriched with

ontologies; detailed information about RDF is presented in [11]. This is especially true for repositories where the application of metadata schema is kept to a minimum.

The paper is organized as follows. Section II contains related work that offers similar or alternative solutions. Section III describes an in-detail view of SW001. Section IV presents the results of a heuristic usability testing. Finally, we conclude in Section V with a summary of the current work along with further research perspectives.

## II. RELATED WORK

Web data integration and extraction using semantic technologies are extended to many knowledge areas as is illustrated in [12], where ecological data from the Guanabara Bay ecosystem are stored in distinct relational databases, thus there is the heterogeneity, lack of metadata standardization, and reduced interoperability. To tackle these problems, a four-level architecture is proposed to integrate, publish and retrieve ecological data from repositories using linked data; data are published as RDF triples using a relational-Resource Description Format mapping language and an application ontology to integrate a common vocabulary and a global view of the generated datasets. Data views are related to queries over data sources, they associate mappings and query answering to represent workflows for scientists.

In the domain of OERs, the use of metadata and ontologies has been studied since different perspectives, for example, [13] analyzes relationships between collection-level and item-level metadata and proposes a general method for translating them into a set of statements in first-order logic and formal knowledge representation languages using logical inference rules, while [14] proposed an agile method that minimizes the need of expertise when semi-structured data are used, they analyze ontology-based methodologies for integrating and reconciling information due to ontologies deal with syntactic and semantic heterogeneity. Reference [15] proposes a standards-compliant approach than involves a set of mappings between domain vocabularies to transform data in a DSpace repository into linked open datasets.

Reference [16] applied semantic searching techniques on digital repositories supported by the DSpace and their results showed that this type of search is useful to expert and novice users and that it enabled new alternatives for content browsing and retrieval in comparison with the keyword-based search. Reference [4] propose an ontology-facilitated sharing of data as an alternative to integrate different IRs metadata; their method consists of transforming data from relational databases of OERs into ontologies that are queried by users from a unique web page. The specialization of this proposed is reported in [5], where authors describe a system that transforms the DSpace metadata database supported by the DSpace into an intermediate database with a normalized schema that is then transformed into an ontology; the aim is to share information with other systems to discover

common interests. The ontology obtained from the OER and ontologies of other systems is integrated by means of semantic correspondence between entities.

In [17], authors report three methods for exporting data about scientific activity stored in the Current Research Information System (CRIS) at the University of Novi Sad; this system implements the Common European Research Information Format (CERIF). One of these methods involves the OAI-PMH protocol and the OAICat library. The exported data generate datasets that enable the creation of graphs that show links between departments, faculties, and researchers, these links are useful to find research leaders and common research areas. Reference [18] authors propose a user-centered approach and the information of a learning management system to customize ontology web language ontologies (OWL) to attend specific needs of disabled students, ontologies are used to access, index, and retrieve heterogeneous items stored in a DSpace repository. More information about OWL is available in [34].

Reference [19] describes an ontology that models students, teachers, monographs, graduated documents student-advisor, student-topic, and student-assessment committee relationships with a domain, range, and cardinality restrictions. Through interviews with the staff of different universities, the ontology integrates a set of rules that formally models the two stages of the student's graduation process.

To manage a huge amount of data and web services, [35] presents a model that uses multithreading technology with dataset parameters to combine web services with parallel processing of computers that form a wide area network. A framework to automatic semantic web service composition that also uses parallel execution of processes during the preprocessing is also described in [36].

In [20] scholarly resources are harvested from various federated repositories, then a common learning object ontology is used for subject classification to automatically annotated learning objects using keyword expansion, inferences, and standard taxonomic vocabularies expressed in SKOS. Reference [21] presents a literature review about applications of semantic technologies in bibliographic databases such as the definition of semantic models for bibliographic descriptions, approaches to transform existing bibliographic data into machine-readable datasets, data enrichment to facilitate web search and information extraction as well as the construction of knowledge repositories.

Finally, another application that takes into account semantic search is the Arabic semantic search engine based on a domain-specific ontological graph for Colleges of Applied Science, Sultanate of Oman (CASOnto) described in [22], this engine supports the factorial question answering and uses keyword-based search and semantic-based search in Arabic and English languages. A comparative study showed that the second search is better than the first search in both simple and complex queries, the performance and efficiency of this engine are compared with Kngine, Wolfram Alpha, and Google.

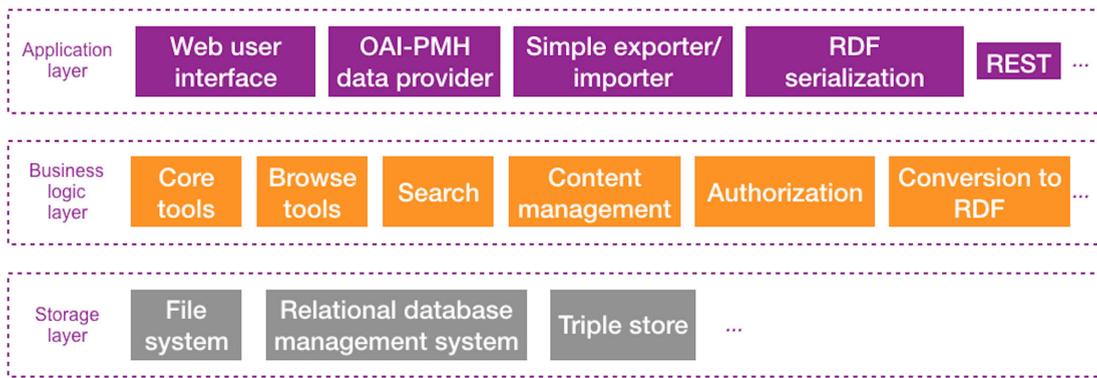


FIGURE 2. Modules of DSpace.

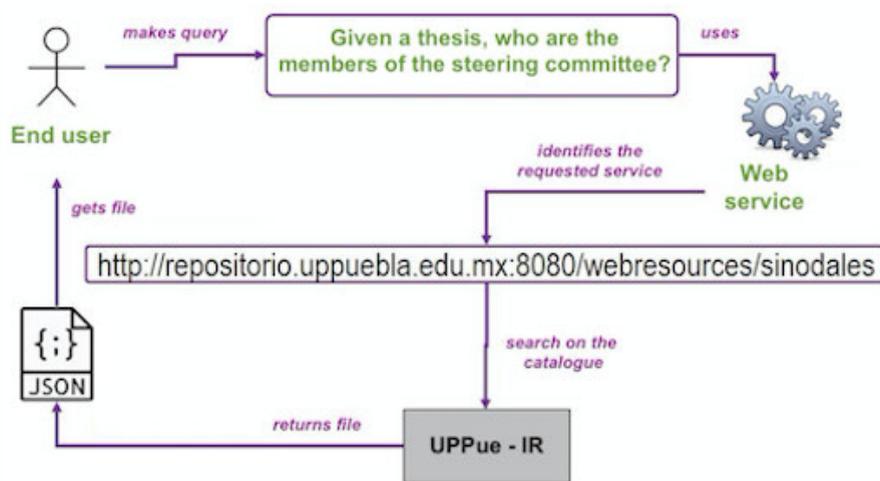


FIGURE 3. High level design for SW001 web service.

### III. DESCRIPTION OF SEMANTIC WEB SERVICE

Some behaviors and Technical Recommendations of the Confederation of Open Access Repositories (COAR) Next Generation Repositories are the following [23]: expose identifiers, declare licenses at the level of resources, discover information through browsing, and expose standardized metrics of use (COAR, 2017). This section describes an in-detail view of the SW001 service that was developed according to the cascade model is in the baseline, then an excerpt of the incremental model referred to small modules, fast tests of functionality, feedback, correction, release, or scaling is used until the implementation stage is finished; the last stage refers to maintenance.

#### A. BASIC REQUIREMENTS

The main requirements for SW001, the proposed semantic web service, are the following: a) store RDF tuples, b) query of semantic information, c) export data in open formats, and d) integrate data from OERs. The installation of a local instance of the 6.2 version of DSpace that emulates the technical requirements of the UPPue repository, the Dspace

modules related to these requirements are showed in Figure 2; it is worth to notice that the RDF module is not enabled in the default DSpace installation.

#### B. DESIGN

Figure 3 shows the high-level design for SW001, its cases of use, and its class diagram are illustrated in Figures 4 and 5, detailed information about these classes and their attributes is described in [24]. By reusing the information of these figures similar services can be constructed.

#### C. IMPLEMENTATION

SW001 is a REST-type service, its implementation, and its feasibility is explained in a scenario that uses a test set of these described using the Dublin Core metadata format, [25], the Java Server Pages (JSP) Interface of DSpace to export the following metadata: id, collection, author, accessioned, available, issued, abstract, provenance, sponsorship, description, citation, URI, iso, publisher, subject, alternative title, title, and type. These metadata are stored in a Comma Separated Values file.

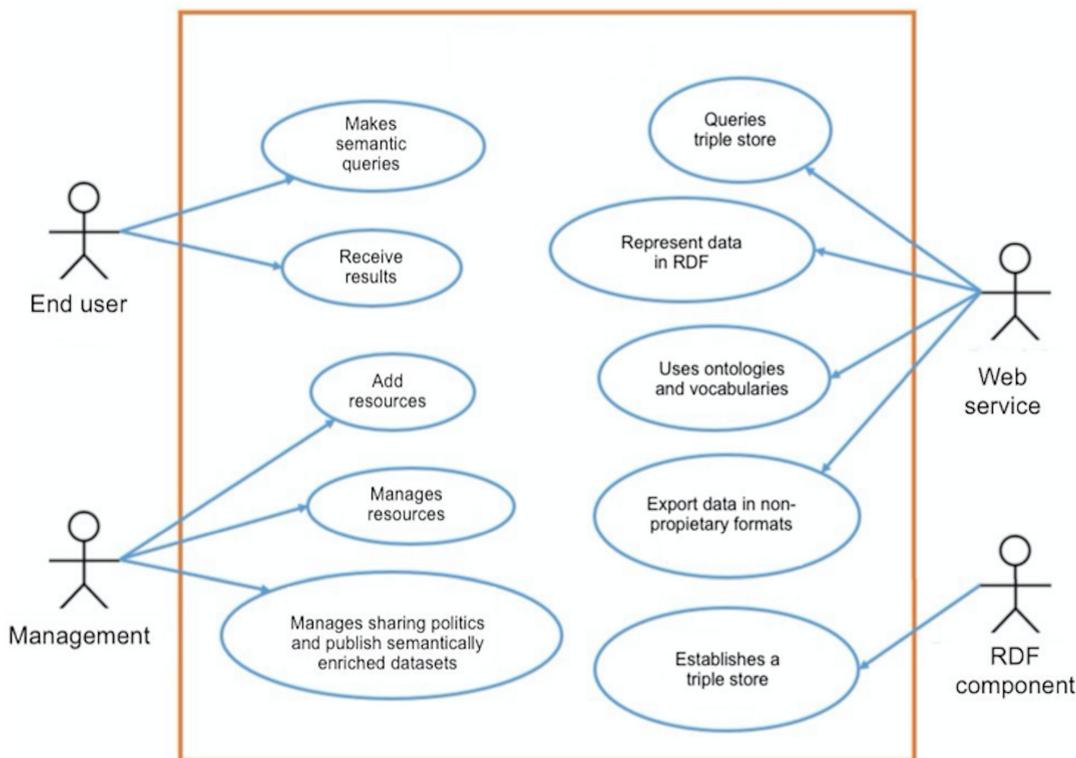


FIGURE 4. Cases of use for SW001 web service.

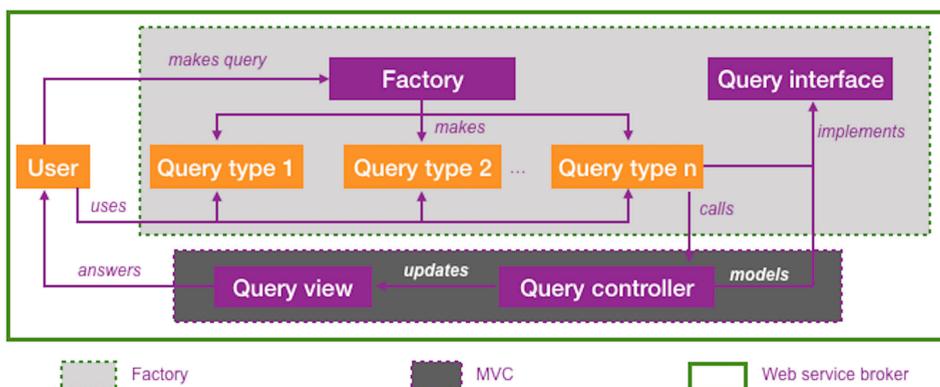


FIGURE 5. Class diagram for SW001 web service.

By using the Jena-Fuseki version 1.6.0 server and the RDFizer service, SW001 transforms metadata into RDF documents that form a triplestore. On one hand, Jena-Fuseki is integrated with TBD, a component that supports a layer of persistent storage, queries, and transactions in RDF documents (Apache Software Foundations, 2020). On the other hand, RDFizer implements extraction, transformation, and load. The RDF document for each thesis is accessible using a Uniform Resource Locator (URL) such as the following:

```
http://repositorio.uppuebla.edu.mx:8080/rdf/handle/123456789/100/ttl
```

A software module in the Python language uses the RDFlib, library that works with XML and RDF files [26]

to verify that the metadata from DSpace was corrected included in the RDF documents. Data of these documents are semantically enriched and modeled as instances of Onto4UPPue ontology that represents documents, users, and their relationships for OERs, its intended use is the establishment of a common vocabulary that provides the foundations for the deployment of semantic web services. A review of the correct insertion of instances into the ontology required of the version 5.2 of Protégé editor [27], Figure 6 shows the instances of the test set and some data properties for a thesis identified as instance T24, note that instances are associated with DataPropertyAssertion properties and thus they model DSpace metadata. The manager of the OER at the UPPUE validated the insertion

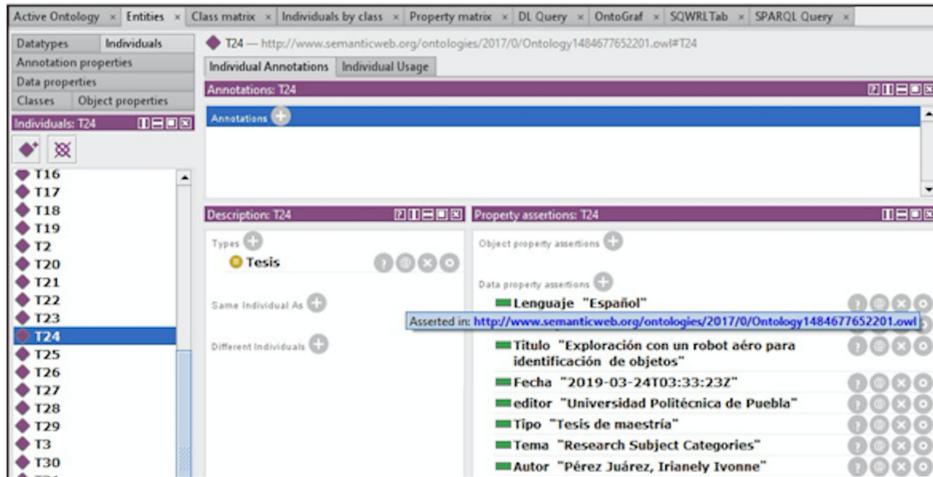


FIGURE 6. Modeling theses as instances of Onto4UPPue ontology.



FIGURE 7. Home page to access SW001 service.

process, the plug-ins for Hermit version 1.3.8 and Pellet version 2.3.0 were used for automatic verification of logical consistency.

The instances for these are linked with other instances that represent students, advisors, and steering committee members by using the object properties of the Onto4UPPUE ontology. As a result, the implementation of the SW001 service offers benefits like the following ones: 1) this new semantically enriched dataset is used in a web site that extends the keyword-based search of DSpace and allows users to retrieve information about the elaboration process of theses, 2) the logical consistency of this dataset has been automatically validated, the use of a common vocabulary reduces the ambiguity of the Spanish language and 4) this dataset is exported into OWL or JSON formats with XML syntax for further analysis. However, there is some drawback such as the

fact that the management of huge datasets will speed up the response times, hence new solutions using parallel processing will be required as well as different alternatives to manage collections.

At present, SW001 service is accessible from a web site composed of six pages in the Spanish language described as follows:

- *Home*. The initial web page that includes a welcoming message and the goal of the service as is illustrated in Figure 7
- *Semantic search*. The initial web page that includes a welcoming message and the goal of the service as is illustrated in Figure 8
- *Test UX*. Displays a test that uses usability heuristics to allow users to evaluate the Semantic search page, see Figure 10

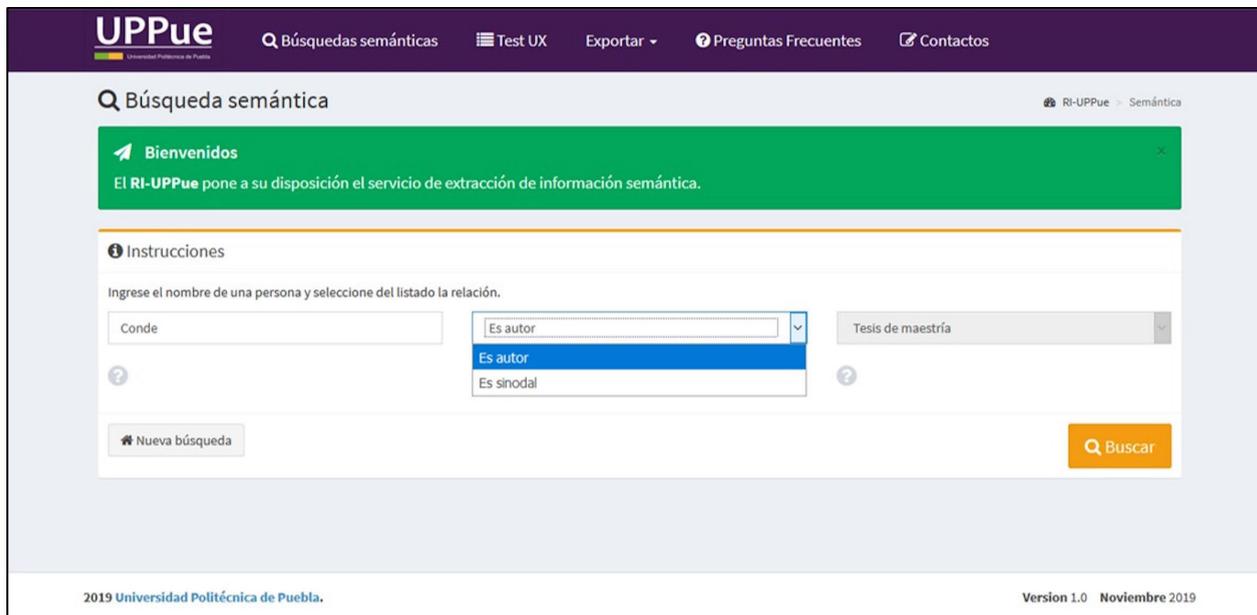


FIGURE 8. The web page for the semantic search of SW001 service.

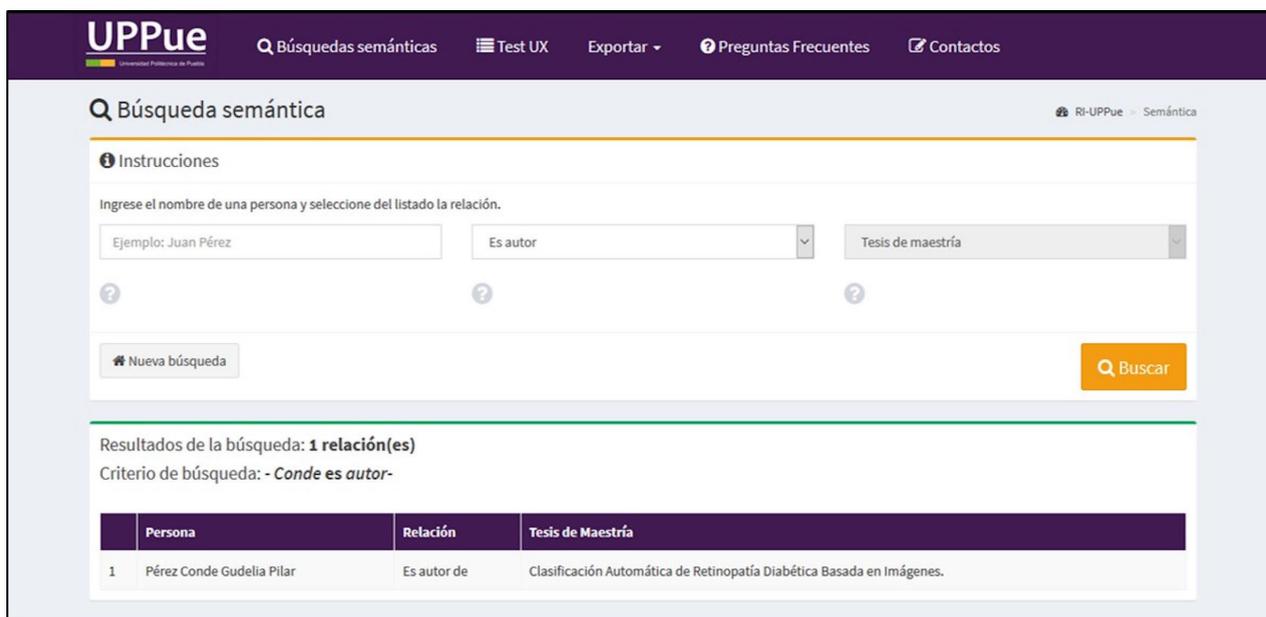


FIGURE 9. Display of a result set for a semantic search in tabular form.

- *Export data.* Enables users to export the ontology with instances in JSON and OWL formats
- *Frequent questions.* Shows frequent questions and answers to support knowledge acquisition of the OERs domain
- *Contact.* Contains contact information of developers

A personal computer and the Google Chrome web browser were used to display Figures 7 to 10. Due to the web site was designed to be responsive, this also can be accessed from a mobile device as is illustrated in Figure 11.

To add other relationships from Onto4UPPue ontology developers need to implement simple modifications to the source code. In summary, the SW001 service exports and enrich bibliographic metadata of the UPPue repository as open linked data.

The technologies used to implement SW001 are illustrated in Figure 12; some versions are 3.6 for Python, 4.2.2 for RDFLib library (RDFlib team, 2013), and 2.4.15 for AdminLTE. The information of Onto4UPPue ontology was extracted with xml.etree.ElementTree (available at

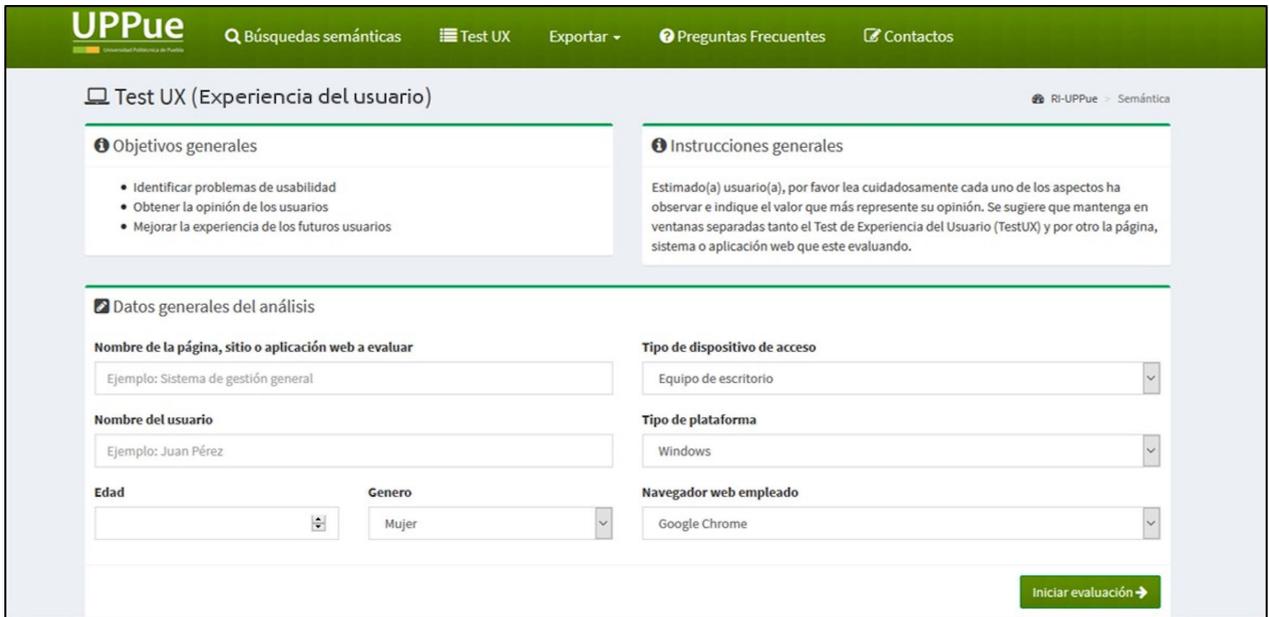


FIGURE 10. Display of a UX test to evaluate the semantic search page.

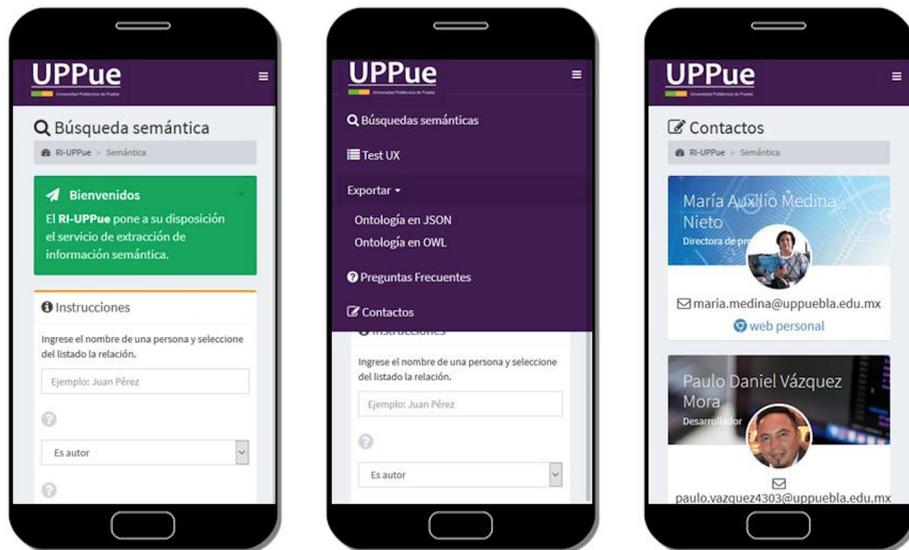


FIGURE 11. Access to SW001 from a mobile device.

<https://pypi.org/project/elementree/>) while the insertion of new instances was supported by the OWLReady2 version 0.19 [29], OWLReady2 is compatible with RDFLib and can be linked with Hermit reasoner. The Flask framework (available at <https://pypi.org/project/Flask/>) was used to implement the Model-View-Controller design pattern whereas the dashboard for the front end is provided by AdminLTE version 2.4.15.

Section IV presents the preliminary results of a survey to evaluates the usability of the described web site.

#### IV. USABILITY HEURISTIC EVALUATION

An exploratory and self-management survey was conducted to evaluate the usability of the web site that allows users to access the SW001 service, this survey adopted the heuristics Torres-Budiel template [30], the application of these heuristics is reported in [31]–[33], they are the following: generalities, identity, and information, language and readability, labels, the structure and browsing, structure of the web pages, searching, multimedia elements, help, accessibility, control, and feedback.

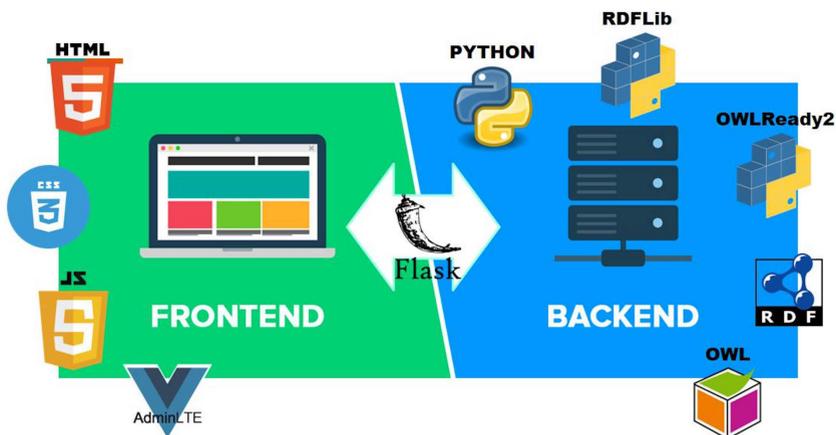


FIGURE 12. Technologies for the front-end and back-end of SW001 service.

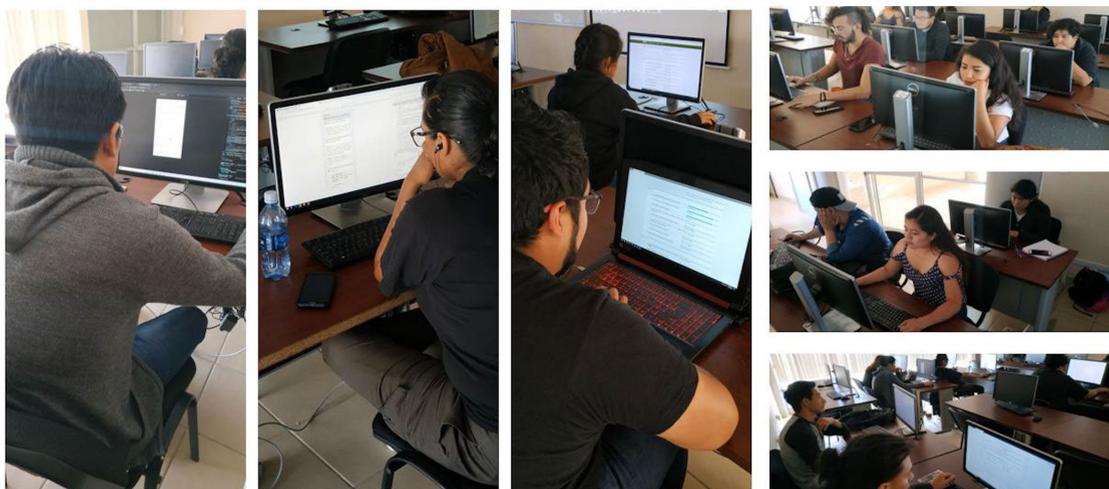


FIGURE 13. Participants of the usability testing.

Each heuristic is associated with a set of questions, for example, the questions for the heuristic of language are the following [30]:

- 1) Does the web site is in the same language than the language of users?
- 2) Are the language and wording concise and clear?
- 3) Are the language and wording friendly, familiar and close?
- 4) Each paragraph has an idea?

Those heuristics were estimated with a purposive sample of 16 undergraduate students from a group of 30 computer science students who participated in the survey, teachers, and managers, these students are frequent users of the UPPue repository, 10 men and 6 women, they are between 21 and 24 years old. Students accessed the web site by using personal computers during a face-to-face session of about 30 minutes. Two teachers and two repository managers participated in the analysis of results. Students expressed a level of satisfaction for each heuristics using the questionnaire

included in the TestUX page according to the following Likert scale:

- It is not applicable (0)
- Strongly disagree (1)
- Disagree (2)
- Neither agree or disagree (3)
- Agree (4)
- Strongly agree (5)

Table 1 shows the average for each heuristic, note that the heuristic with the maximum value is accessibility. The questions related to accessibility are the following [30]:

- 1) Do the images include the ‘alt’ attribute that describes their content?
- 2) Is the web site compatible with different browsers? Is it visible with different screen resolutions?
- 3) Can the users browse for all the web site without downloading any plug-in?
- 4) Does the weight of the web site has been controlled?
- 5) Can the web site be printed without any problems?

TABLE 1. Results for each usability heuristic.

Heuristic of usability	Value
Identity	4.03
Language	4.19
Label	4.08
Structure	4.07
Layout	4.30
Searching	4.08
Multimedia	4.29
Help	2.55
Accessibility	4.35
<b>Average</b>	<b>4.02</b>

The average of the 11 heuristics was 4.02, this value is very close to the agreed value of the Likert scale, as a conclusion, the students reported that they were very satisfied with the usability of the web site. For the heuristic with the lowest value, that is, help, a section for frequently asked questions (FAQs) was added as well as links to the help section were added after the usability test.

Figure 13 displays photographic evidence of students during the usability test.

## V. CONCLUSION

The paper outlines concepts related to OERs and described SW001, a web service that retrieves theses and extends the keyword based-search of the UPPUE repository taking into account the roles of advisors and steering committee members formally represented into the Onto4UPPUE ontology. The proposed REST-type service was designed and implemented taking into account scalability, transparency, and some technical requirements of COAR. By using a triple store, SPARQL queries and technological tools such as the RDF serialization module of DSpace platform, the service links the repository and the Onto4UPPue ontology in the sense that the more thesis and descriptive data are inserted into the repository, the more instances are integrated into the ontology.

On one hand, the use of the ontology reduce ambiguity in the interpretation of descriptive data, its vocabulary is shared between users and computers and represents a tool to verify automatically inconsistencies in the data. On the other hand, the ontology supports the generation of semantically enriched datasets that coexist with the repository.

The functionality of SW001 was verified through the implementation of a software module in charge of checking that all exported metadata from DSpace were integrated into ontology instances, the Protégé editor was used to explore information about theses. The logical consistency of the ontology was checked by using reasoners.

A web site was designed to access SW001 from web browsers that allow users to retrieve content in the Spanish language by using relationships between steering committee members, however, the development methodology enables developers to implement simple updates to the source code to integrate other ontological relationships and similar services. The TestUX web page was constructed for usability testing, sixteen undergraduate students of computer science and fre-

quent users of the UPPUE repository participated in the testing, the information of 11 usability heuristics was gathered, the results were presented by heuristic and they indicated that students reported a high level of satisfaction. After the testing, some support elements were added into the web site due to the help heuristic had gotten the minimum value.

Since the web site, users can download the ontology in JSON and OWL formats, thus this can be reused by other semantic web applications. Furthermore, reasoning will enable intelligent application development and further exploitation of the generated datasets, for instance, to infer implicit information.

The paper provides the basis for the development of semantic services, this contributes to spread the benefits of open access and the construction of semantically enriched data sets. The current work is focused on the integration of any type of document store in the repository to be considered as an ontology instance. As future work, we plan to design semantic web services addressed to obtain output indicators for authors that support decision-makers in the OER domain.

## REFERENCES

- [1] J. A. Merlo, *Ecosistemas del Acceso Abierto*. Salamanca, Spain: Ediciones Universidad de Salamanca, 2018.
- [2] C. Lagoze, H. Van de Sompel, M. Nelson, and S. S. Warner, "Implementation guidelines for the open archives initiative protocol for metadata harvesting," Open Arch. Initiative, Ithaca, NY, USA, Tech. Rep. 2005/05/03T22:51:00Z, 2005. [Online]. Available: <http://www.openarchives.org/OAI/2.0/guidelines.htm>
- [3] Jisc Open Access Services. (2020). *OpenDOAR: Directory of Open Access Repositories*. [Online]. Available: <http://v2.sherpa.ac.uk/pendoar/>
- [4] H. Fari and S. Khan, *Exchanging Data From Institutional Repositories to the Semantic Web*. Riga, Latvia: Lap Lambert Academic Publishing, 2015.
- [5] H. Farid, S. Khan, and M. Y. Javed, "DSont: DSpace to ontology transformation," *J. Inf. Sci.*, vol. 42, no. 2, pp. 179–199, Apr. 2016, doi: 10.1177/0165551515591406.
- [6] M. R. Beazley, "Eprints institutional repository software: A review," *Partnership, Can. J. Library Inf. Pract. Res.*, vol. 5, no. 2, pp. 1–6, Aug. 2018, doi: 10.21083/partnership.v5i2.1234.
- [7] M. Smith, M. Barton, M. Bass, M. Branschofsky, G. McClellan, D. Stuve, R. Tansley, and J. H. Walker, "DSpace: An open source dynamic digital repository," *D-Lib Mag.*, vol. 9, no. 1, pp. 87–97, 2003.
- [8] Mexican Government. National Council of Science and Technology. (2020). *National Repository*. [Online]. Available: <https://www.repositorionacionalcti.mx>
- [9] M. A. Medina, J. A. Sánchez, O. Cervantes, R. C. Medina, J. De la Calleja, and A. Benitez, "Semantic representation of operative and domain knowledge for institutional repositories," Mexican Copyright Nat. Inst., Mexico city, Mexico, Tech. Rep. 03-2017-042511235500, 2017, pp. 1–9.
- [10] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?" *Int. J. Hum.-Comput. Stud.*, vol. 43, nos. 5–6, pp. 907–928, Nov. 1995, doi: 10.1006/ijhc.1995.1081.
- [11] F. Gandon and G. Schreiber. (Feb. 2014). *RDF 1.1 XML syntax: W3C Recommendation*. [Online]. Available: <http://www.w3.org/TR/rdf-syntax-grammar/>
- [12] A. M. D. C. Moura, F. Porto, V. Vidal, R. P. Magalhaes, M. Maia, M. Poltosi, and D. Palazzi, "A semantic integration approach to publish and retrieve ecological data," *Int. J. Web Inf. Syst.*, vol. 11, no. 1, pp. 87–119, Apr. 2015, doi: 10.1108/IJWIS-08-2014-0028.
- [13] K. Wickett, "A logic-based framework for collection/item metadata relationships," *J. Documentation*, vol. 74, no. 6, pp. 1175–1189, Oct. 2018, doi: 10.1108/JD-01-2018-0017.
- [14] B. A. Casas and G. H. Ceballos, "Integrating semi-structured information using semantic technologies: An evaluation of tools and a case study on university rankings data," in *Proc. 3rd Int. Conf. Data Manage. Technol. Appl. (DATA)*, M. Helfert, A. Holzinger, and C. Francalanci, Eds. Setúbal, Portugal: ScitePress, 2014, pp. 357–364, doi: 10.5220/0005004203570364.

- [15] N. Konstantinou, D.-E. Spanos, N. Houssos, and N. Mitrou, "Exposing scholarly information as linked open data: RDFizing DSpace contents," *Electron. Library*, vol. 32, no. 6, pp. 834–851, Nov. 2014, doi: [10.1108/EL-12-2012-0156](https://doi.org/10.1108/EL-12-2012-0156).
- [16] G. Solomou and D. Koutsomitropoulos, "Towards an evaluation of semantic searching in digital repositories: A DSpace case-study," *Program*, vol. 49, no. 1, pp. 63–90, Feb. 2015, doi: [10.1108/PROG-07-2013-0037](https://doi.org/10.1108/PROG-07-2013-0037).
- [17] L. Ivanovic, D. Ivanovic, and S. B. Dimic, "Improving dissemination of human knowledge by exporting data from research information systems," in *Proc. Manage., Knowl. Learn. Int. Conf.*, Portorož, Slovenia: ToKnowPress, 2015, pp. 711–718. [Online]. Available: <https://ideas.repec.org/h/tkp/mkpl14/711-718.html>
- [18] C. Skourlas, A. Tsolakidis, P. Belsis, D. Vassis, A. Kampouraki, P. Kakoulidis, and G. A. Giannakopoulos, "Integration of institutional repositories and e-learning platforms for supporting disabled students in the higher education context," *Library Rev.*, vol. 65, no. 3, pp. 136–159, 2016, doi: [10.1108/LR-08-2015-0088](https://doi.org/10.1108/LR-08-2015-0088).
- [19] G. A. P. García, "Aplicación de la Lógica de predicados para la Realización de inferencias en documentos de Pasantía descritos a Nivel Ontológico," *Ingenierías USBMed*, vol. 7, no. 1, pp. 11–19, Jun. 2016, doi: [10.21500/20275846.1802](https://doi.org/10.21500/20275846.1802).
- [20] D. A. Koutsomitropoulos, "Semantic annotation and harvesting of federated scholarly data using ontologies," *Digit. Library Perspect.*, vol. 35, nos. 3–4, pp. 157–171, Nov. 2019, doi: [10.1108/DLP-12-2018-0038](https://doi.org/10.1108/DLP-12-2018-0038).
- [21] T. Georgieva-Trifonova, K. Zdravkov, and D. Valcheva, "Application of semantic technologies in bibliographic databases: A literature review and classification," *Electron. Library*, vol. 38, no. 1, pp. 113–137, Dec. 2019, doi: [10.1108/EL-03-2019-0081](https://doi.org/10.1108/EL-03-2019-0081).
- [22] A. Sayed and A. Al Muqrishi, "CASANTO: An efficient and scalable Arabic semantic search engine based on a domain specific ontology and question answering," *Int. J. Web Inf. Syst.*, vol. 12, no. 2, pp. 242–262, Jun. 2016, doi: [10.1108/IJWIS-12-2015-0047](https://doi.org/10.1108/IJWIS-12-2015-0047).
- [23] COAR Building a Global Knowledge Commons. (2017). *Next Generation Repositories: Behaviors and Technical Recommendations of the COAR Next Generation Repositories Working Group*. Accessed: Jul. 20, 2018. [Online]. Available: <https://www.coar-repositories.org/files/NGR-Final-Formatted-Report-cc.pdf>
- [24] D. Vázquez, M. A. Medina, J. De la Calleja, A. Benitez, T. M. Vidal, and D. Alanís, "Diseño de un servicio Web para la recuperación de información semántica del repositorio institucional de la Universidad Politécnica de Puebla," in *Proc. 33rd Congreso de Instrumentación (SOMI)*, 2018, vol. 5, no. 1.
- [25] Dublin Core Metadata Initiative. (2014). *DCMI Metadata Terms*. [Online]. Available: <http://www.dublincore.org/specifications/dublin-core/dcmi-terms/>
- [26] RDFLib Team. (2020). *RDFLib 5.0.0*. [Online]. Available: <https://rdflib.readthedocs.io/en/stable/>
- [27] M. A. Musen, "The protégé project: A look back and a look forward," *AI Matters*, vol. 1, no. 4, pp. 4–12, 2015, doi: [10.1145/2757001.2757003](https://doi.org/10.1145/2757001.2757003).
- [28] E. Prud'Hommeaux and A. Seaborne. (2007). *SPARQL Query Language for RDF: W3C Recommendation 15 January 2008*. [Online]. Available: <https://www.w3.org/TR/rdf-sparql-query/>
- [29] L. Jean-Baptiste. (2019). *Owready2 Version 0.24*. Python Package Index. [Online]. Available: <https://pypi.org/project/Owready2/>
- [30] Torres Burriel Estudio. (2008). *Test Heurístico De Torres Burriel*. [Online]. Available: <http://www.torresburriel.com/weblog/2008/11/28/plantilla-para-hacer-analisis-heuristicos-de-usabilidad/>
- [31] G. E. C. Golondrino, D. P. Oliveros, and M. Y. C. Muñoz, "Automation of usability inspections for websites," in *Human-Computer Interaction (Communications in Computer and Information Science)*, vol. 1114, H. P. Ruiz and V. Agredo-Delgado, Eds. Puebla, Mexico: Springer, 2019.
- [32] O. D. Pérez, G. E. Chanchi, and M. I. Vidal, "Propuesta de un test heurístico de accesibilidad para sitios Web basados en la norma NTC 5854," *Revista Ibérica de Sistemas e Tecnologías de Informação*, vol. 17, no. 1, pp. 170–182, 2019.
- [33] E. Serna, *Desarrollo e innovación en ingeniería*. Instituto Antioqueño de investigación, 2018.
- [34] P. F. Patel-Schneider, P. Hayes, and I. Horrocks. *OWL Web Ontology Language Semantics and Abstract Syntax*. Accessed: Jun. 13, 2019. [Online]. Available: <http://www.w3.org/TR/owl-semantics/>
- [35] A. Kut and D. Birant, "An approach for parallel execution of Web services," in *Proc. IEEE Int. Conf. Web Services*, Jul. 2004, pp. 812–813, doi: [10.1109/ICWS.2004.1314831](https://doi.org/10.1109/ICWS.2004.1314831).
- [36] P. Bartalos and M. Bielikova, "Semantic Web service composition framework based on parallel processing," in *Proc. IEEE Conf. Commerce Enterprise Comput.*, Vienna, Austria, Jul. 2009, pp. 495–498, doi: [10.1109/CEC.2009.27](https://doi.org/10.1109/CEC.2009.27).



**MARÍA AUXILIO MEDINA NIETO** received the M.Sc. and Ph.D. degrees from the Universidad de las Americas Puebla, Mexico, in 2008. She is currently a Researcher with the Post-graduate Department, Polytechnic University of Puebla (UPPuebla), Puebla. She has higher profile recognition (PRODEP). Her research interests include ontologies, semantic web, and open educational repositories. She has been a member with the National System of Researchers (SNI) Level 1 since 2019.



**PAULO DANIEL VÁZQUEZ MORA** received the M.Sc. degree from the Polytechnic University of Puebla (UPPuebla), Puebla, Mexico, in 2019. He has been a Web Developer (Experienced) with public and private sector since 2006. He is currently a part-time Professor with the Information Technologies Department, Technological University of Puebla (UTP), Puebla. His research interests include open access, information exchange, web services, and open educational repositories.

He received several certifications from the National Council of Normalization and the Certification of Working Skills.



**JORGE DE LA CALLEJA MORA** received the M.Sc. and Ph.D. degrees from the National Institute of Astrophysics, Optics and Electronics (INAOE), Mexico. He has been a full-time Professor with the Computer Science Department, Polytechnic University of Puebla (UPPuebla), Mexico, since 2008. His research interests include machine learning, computer vision, and data mining with applications in medicine, education, and astronomy. He is a member of the National System of Researchers (SNI) Level 1.



**MIREYA TOVAR VIDAL** received the Ph.D. degree in computer science from CENIDET, Mexico, in 2015. She is currently a full-time Professor with the Faculty of Computer Science, Benemerita Universidad Autónoma de Puebla (BUAP). Her research interests include ontology, clustering, information retrieval, computational linguistics, logic, and natural language processing. She is a member of the National System of Researchers (SNI) Level 1 in 2016.



**EDUARDO LÓPEZ DOMÍNGUEZ** (Member, IEEE) received the Ph.D. degree from the National Institute of Astrophysics, Optics and Electronics (INAOE), Mexico, in 2010. He is currently a Researcher with the Department of Computer Science, Laboratorio Nacional de Informática Avanzada (LANIA), Mexico. His research interests include mobile distributed systems, partial order algorithms, and multimedia synchronization. He is a member of the National System of Researchers (SNI) Level 1.



**ISMAEL EVERARDO BÁRCENAS PATIÑO** received the Ph.D. degree in computer science from the University of Grenoble. He is currently an Assistant Professor with the Computing Engineering Department, National University of Mexico. His main research interests include the theory of automated reasoning and its application in areas, such as knowledge representation and formal verification. He is a member of the National System of Researchers (SNI) Level 1.

...



**DELIA ARRIETA DÍAZ** received the master's degree in quality of public management, the master's degree in gestalt therapy, and the Ph.D. degree in government and public management. She is currently a full-time Teacher of bachelor's and master's degrees with the Faculty of Economics, Accountability, and Management, Juarez University of the State of Durango (UJED). She has coordinates the academic group called Management and Development of Organizations. She has higher profile recognition (PRODEP). She received the Teacher Certification from the National Association of Faculties and Schools of Accountability and Management (ANFECA).