Onto4AIR2: a simple ontology to represent theses from open repositories as products of academic collaboration

María Auxilio Medina Nieto [a,*], Jorge de la Calleja Mora [b], Eduardo López Domínguez [c], Ismael Everardo Bárcenas Patiño [d], Delia Arrieta Díaz [e], Araceli Ortiz Carranco [f], and Claudia Zepeda Cortés [g]

[a] *Postgraduate Departament, Polytechnic University of Puebla Juan C. Bonilla, Puebla, México*
E-mail: maria.medina@uppuebla.edu.mx
[b] *Posgraduate Departament, Polytechnic University of Puebla Juan C. Bonilla, Puebla, México*
E-mail: jorge.delacalleja@uppuebla.edu.mx
[c] *Department of Computer Science, Laboratorio Nacional de Informática Avanzada A. C. Xalapa, Veracruz, México*
E-mail: eduardo.lopez@lania.edu.mx
[d] *Computer Engineering Department, National University of México, Ciudad de México, México*
E-mail: ebarcenas@unam.mx
[e] *Faculty of Economics, Accounting and Management, Juarez University of the State of Durango, Durango, México*
E-mail: dad@ujed.mx
[f] *Financial Engineering, Polytechnic University of Puebla, Juan C. Bonilla, Puebla, México*
E-mail: araceli.ortiz@uppuebla.edu.mx
[g] *Faculty of Computer Science, Benemérita Universidad Autónoma de Puebla, Puebla, México*
E-mail: czepeda@gmail.com

Abstract.
This paper describes Onto4AIR2, a simple ontology to represent theses from open repositories as products of academic collaboration. The goal is the construction of machine-readable datasets that are semantically labeled for the further deployment of web services of shared interest to managers, developers, and users within educational organizations. The ontology is populated with sample data of theses from the National Repository of Mexico, an initiative promoted by the National Council of Science, and Technology. The paper suggests practical applications derived from the formalisms of the ontology, and describes an assessment technique where participants were managers, developers, and potential users of the ontology. Developers followed a competency questions-based approach and determined that the ontology represents questions and answers using its terminology; whereas potential users participated in a satisfaction survey; the results showed a positive perception about the ontology. Onto4AIR2 is in English and Spanish languages, this fosters unique and formal definitions of concepts from the Mexican repositories domain.

Keywords: ontologies, repositories, machine readable datasets, semantic web, document management

1. Introduction

For the National Council of Science and Technology (CONACYT), the main Mexican agency that funds research, an open or institutional repository, is a centrali-

*Corresponding author. E-mail: maria.medina@uppuebla.edu.mx.

zed technological framework that complies with international standards to store, manage, preserve, and disseminate scientific, technological, and innovative information derived from educational activities, and research processes [1], [2]. According to [3], repositories collect, index, and share the intellectual capital of faculty staff.

In a developing country like Mexico, endeavors to build repositories have been undergoing for several years like the Mexican Network of Institutional Repositories (REMERI, after the initials in the Spanish statement *Red Mexicana de Repositorios Institucionales*) [4], and the National Repository (NR) [5]. At the time of this writing, the NR interoperates with 106 repositories, integrates 125,824 resources, and has answered 14,832,819 queries; its functionality and technical requirements are specified in [1] and [2]. Technically, the NR implements the 2.0 version of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH protocol) [6], an international interoperability standard that uses unqualified Dublin Core (DC) as the default metadata format [7]. The retrieval mechanisms of the NR match keywords from authors, titles, subjects, and dates.

By exploring theses metadata from repositories, there were differences in the use of DC elements. As a way of illustration of the exposed problematic, the `DC contributor` element sometimes stores the name of a thesis advisor, others the second author of an article, a reviewer of a different educational organization to the first author of the thesis, or even the name of a main-funding organization[1]. Therefore, efforts are required along the lines of unifying uses and meanings in the theses metadata.

This paper presents Onto4AIR2, a simple ontology to represent theses from open repositories as products of academic collaboration; ontologies are knowledge representation models defined in [8] as formal specifications of shared conceptualizations. The goal is supporting the construction of machine-readable datasets that are semantically labeled for the further deployment of web services of shared interest to managers, developers, and users within educational organizations. Onto4AIR2 is an acronym of *Ontologies for All/Any Institutional Repositories*, this is the newer version of the ontology described in [9], the number *2* denotes a second version and two languages, English and Spanish.

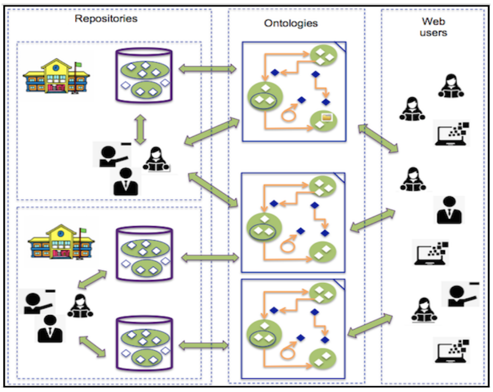

Fig. 1. The context of Onto4AIR2 ontology

Ontologies define relevant concepts for particular applications [10], in this case, Onto4AIR2 models academic collaboration, members of evaluation committees, theses, and their properties, this has been designed to establish a common vocabulary that reduces ambiguity and enables knowledge acquisition. The ontology includes `foaf`[2], `schema.org`[3], and `skos`[4] vocabularies; a set of competency questions expressed as SPARQL[11][5] queries determine its scope.

Figure 1 shows the context of Onto4AIR2 ontology. On one hand, educational organizations have one or more repositories that store theses and other educational resources; on the other hand, these resources are transformed into instances to construct machine-readable and semantically enriched datasets that support information needs of web users. At present, the level of the ontology is proof of concept.

The paper is organized as follows. Section 2 contains related work. Section 3 presents the proposed ontology. Section 4 explains practical applications derived from ontologies. Section 5 describes the results of an assessment technique. Finally, we conclude in Section 6 with a summary of the present work along with further research perspectives.

[1]More information about the usage of DC contributor element is available at: https://www.dublincore.org/specifications/dublin-core/usageguide/appendix_roles/

[2]http://xmlns.com/foaf/spec/

[3]http://schema.org

[4]https://www.w3.org/2004/02/skos/

[5]SPARQL is an acronym for *SPARQL protocol and RDF query language.*

2. Related work

Technological platforms such as Eprints [12], Digital commons [13], existDB [14], Greenstone [15], Fedora Commons [16] or DSpace [17] support a lot of open repositories; databases like OpenDOAR report their impact [18], they use tables to store specific metadata formats [19]. This section presents related works that use semantic technologies for knowledge acquisition, data integration, and metadata management.

SpecINT is a hybrid framework for scenarios where integration is not possible because there are no mapping schemes between bioinformatics repositories, this uses graph eigenvectors, vertices ranking, and existing pattern SPARQL queries to identify data sources [20]. ORDI is an open-source ontology middleware that implements the Resource Description Framework (RDF [21]) data model to integrate structured data sources; this processes and storages metadata and context information regardless of specific ontologies [22]. Authors of [23] use ontologies to integrate data resources and workflows that support online science activities (semantic e-science); ontologies represent theories, models, online tools, and means of innovation.

Semantic repositories are DataBase Management Systems (DBMSs) or other software components that store RDF data, have inference mechanisms and provide users of real-time query-answering mechanisms [24]. OWLIM is a set of semantic repositories used in the life sciences, telecoms, and publishing sectors developed as part of the *Semantic Knowledge Technologies* (SEKT) and *Triple Space Communication* (TRIPCOM) European Research Projects. OWLIM is packaged as a storage and inference layer for the Sesame open RDF framework, this has two variants: 1) *SwiftOWLIM*, the free version that uses an in-memory RDF database, an inference-engine, and a query-answering engine and 2) *BigOWLIM*, the commercial version that handles huge volumes of RDF data.

Ontological design patterns are used during the creation, verification, and validation of the ontologies in industrial scenarios [25], whereas an ontology serves as a reference model for a historical organizational memory (HOM) in the business domain [26]. The main terms represent document types, collections, subjects, the audience, and the dissemination channel of the HOM, they were gathered by interviewing experts. The ontology indexes the resources, and equivalent relationships support multiple names for terms in different languages.

In academic contexts like the Monterrey Technological Institute (ITESM), an information system, a multi-agent system, a knowledge management system, and a knowledge information interpreter that coordinates repositories, domain ontologies, and databases, compose an intelligent platform for knowledge acquisition; the repositories store theses, journal articles, research-based books, patents, technology licensing, trademarks, and documents of technology-based start-up companies [27]. [28] presents an analysis of ontology-based methodologies for integrating and reconciling information, authors proposed an agile method that minimizes the need for ontological expertise when semi-structured data are used; the instances correspond to university ranking data.

Ontologies in [29] represent students, teachers, monographs, theses, and model the graduation process; the main terms come from staff interviews of different universities. The ontology of the Leigh University Benchmark [30] describes departments, individuals, and their relationships, its written in OWL [31]; LUMB(8000) stores data from 8000 universities and 1.1 billion explicit statements, while LUMB(90000) has over 12 billion explicit statements and nearly 21 billion after inference.

The Bowlogna ontology contributes to the gradual adoption of the overall description of the Bologna reform across European universities; this models academic settings and administrative procedures; a graphical user interface (GUI) enables faceted search and browsing for course information [32].

On one hand, for organizations that use the DSpace platform, authors of [33] propose exporting data and semantically enriched by using the OAI-PMH protocol, the OAICat library, and the Current Research Information System (CRIS). The exported data form datasets that link departments, faculties, and researchers; these links are useful to discover research leaders and common research areas. [34] describes a system that converts the DSpace database into an ontology using a normalized schema, the aim is to share information with other systems to discover common interests; the ontology obtained is integrated to other ontologies through semantic correspondence between entities. On the other hand, [19] propose an ontology-facilitated sharing to integrate repositories data, their method consists of transforming data into ontologies accessible from a unique web page. An overview of the ontology engineering evolution that discusses some of the unsolved issues and opportunities for future research is described in [35].

3. Onto4AIR2 ontology

[36] presents an overview of methodologies to construct ontologies, some of them use the IEEE standard 1074-1995. The construction of Onto4AIR2 ontology followed the steps proposed in [37], they are implemented a different order using the Protégé [38] editor as follows:

- *Determine domain and scope*. The ontology is designed to establish a common vocabulary that reduces ambiguity and enables knowledge acquisition, its potential users are document managers, members of evaluation committees of theses, (students, teachers, directors), institutional authorities and web users. Table 1 shows the Competency Questions (CQs) that determine its scope; more information about CQs can be found in [39] and [40].
- *Reuse of existent vocabularies*. Figure 2 shows the vocabularies integrated into the ontology to improve reusability and minimize ambiguity.
- *Enumerate main concepts*. The terms are related to the following main concepts: institutional repository, educational resources, persons, educational organizations, and subjects or knowledge areas.
- *Define classes and construct their hierarchy*. Once the main concepts are enumerated, they are defined as classes, then remaining concepts are obtained by generalization, and specialization. `rdfs:isDefinedby` properties include class definitions as quick guidelines for users, clarity and completeness are supported by CONACYT documents [1] and [2]; `rdfs:comment` and `rdfs:seeAlso` properties also include fruitful comments and notes. Figure 3 shows an excerpt of the class hierarchy; it is worth to mention that classes for educational resources are disjoint.
- *Assign properties for classes*. Classes have properties, for example, see Table 2; according to the NR specifications [2], properties for a repository are mandatory (M) or recommended (R).
- *Define properties between instances*. Table 3 shows the domain and range of properties that link instances; an asterisk means that a property accomplishes a facet. The notation is as follows: functional (F), inverse functional (IF), asymmetric (A), and irreflexive (I). Some of these properties are also used in [41] to describe a collection of posters. Note that: 1) the facets from the `knows` property are also reflexive, symmetric and transitive, and 2) the contribution of authors in a thesis is represented in the sub-properties of `isAuthorOf` called `isFirstAuthorOf` for the student, and `isCoauthorOf` for evaluation committee members.
- *Create instances*. Populating the ontologies with instances allow users to have knowledge bases. Figure 4 shows a thesis modeled as an instance.

Table 1
Competency questions for Onto4AIR2 ontology

Identifier	Description
CQ1	What is a repository?
CQ2	What types of resources does a repository store?
CQ3	How to describe a thesis?
CQ4	How are the theses organized?
CA5	What information does a person have in the repositories?
CQ6	Which people have reviewed a thesis?
CQ7	Which is the affiliation of the students and the teachers?
CQ8	Which are the knowledge areas for the theses?

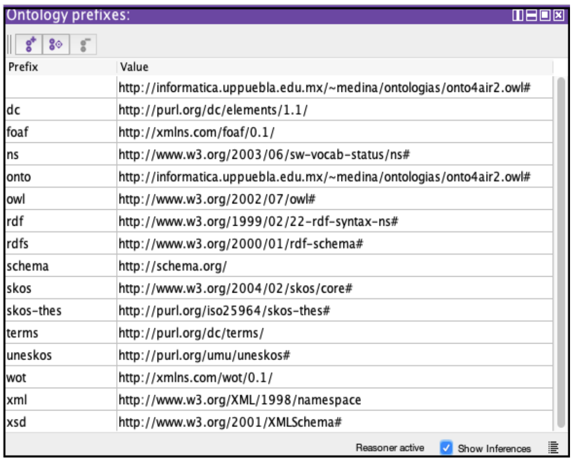

Fig. 2. Vocabularies integrated into Onto4AIR2 ontology

Table 2
Data properties for the `InstitutionalRepository` class

Property	Type	Property	Type
`identifier`	M	`numberOfDocs`	R
`irName`	M	`typeOfDocs`	R
`languageOfDocs`	R	`updateDate`	M
`manager`	M	`website`	M

Table 3
Object properties of Onto4AIR ontology

Object property	Domain	Range	F	IF	A	I
`hasAdvisorOfMasterThesis`	`Student`	`Teacher`	*		*	*
`hasKnowledgeArea`	`CreativeWork, Document`	`ConceptGroup`	*		*	*
`hasKnowledgeField`	`CreativeWork, Document`	`Concept`			*	*
`isAuthorOf`	`Person`	`CreativeWork, Document`			*	*
`isCoauthorOf`	`Person`	`CreativeWork, Document`				
`isFirstAuthorOf`	`Person`	`CreativeWork, Document`	*		*	*
`isManagedBy`	`InstitutionalRepository`	`IR-manager`	*	*	*	*
`isMemberOfEvaluationCommittee`	`Teacher`	`Student`			*	*
`knows`	`Person`	`Person`				
`manages`	`IR-manager`	`InstitutionalRepository`	*	*	*	*
`worksIn`	`Teacher`	`EducationalOrganization`				

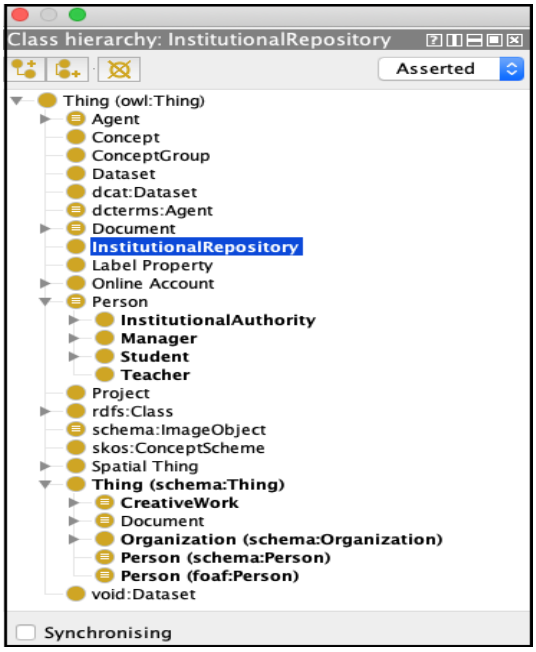

Fig. 3. An excerpt of the class hierarchy of Onto4AIR2 ontology

4. Practical applications derived from the formalisms of the Onto4AIR2 ontology

Some practical applications derived from the formalisms of Onto4AIR2 ontology useful for managers and developers of repositories are the following:

- *Capacity to insert incomplete information.* Examples of use:
  - ∗ Insertion of theses or students without any scholar identifier
  - ∗ Insertion of theses without subject

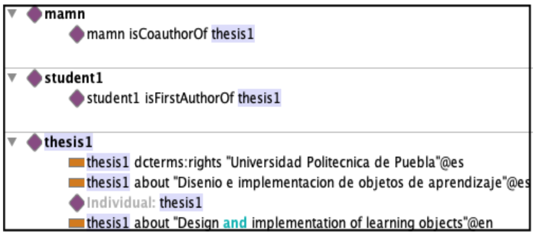

Fig. 4. A thesis modeled as an instance.

- *Modeling persons as instances.* An instance can belong to one or more subclasses of the Person class, some subclasses are disjoint while others have common elements. Examples of use:
  - ∗ Insertion of information about university applicants
  - ∗ An undergraduate teacher can also be a graduate student
  - ∗ A person directs a thesis and reviews another
- *Working with disjoint classes.* The type of educational resource is unique because its subclasses are disjoint. Examples of use:
  - ∗ An educational resource is a thesis or an article but this can not be of both types simultaneously
  - ∗ A thesis associated with two or more types of educational resource will produce an inconsistence, that enables managers to identify automatically incorrect information
- *Management of cardinality restrictions.* Unlike relational database schema, cardinality restrictions can have a minimum, maximum, or exact value. Examples of use:
  - ∗ A thesis has only a first author that belongs to the `Student` class

  - ∗ An evaluation committee has exactly a fixed number of members
  - ∗ An advisor directs a thesis
  - ∗ An advisor collaborates or works in one or more educational organizations
  - ∗ A thesis has exactly one title, although there can be alternative titles

- *Data properties management*. Properties between instances and data types have domain and range restrictions. Examples of use:
  - ∗ The theses have a title and a publication date
  - ∗ The first author of a thesis is an instance of the `Student` class
  - ∗ The members of evaluation committees are instances of the `Teacher` class
  - ∗ A thesis is associated with states such as registered, in process, or finished
- *Inference mechanisms*. The reasoners use universal and existential restrictions to infer new knowledge. Examples of use:
  - ∗ A person is an advisor if there is a thesis that has been directed by him or her
  - ∗ The members of an evaluation committee are related to at least one thesis, this activity is considered academic collaboration
  - ∗ Any thesis has an evaluation committee
- *Functional facets*. Examples of use:
  - ∗ A student belongs to an educational organization
  - ∗ A thesis is associated with a unique academic program

Equivalence relationships between classes support English and Spanish languages, different names for the same instance can also be used.

5. Results and discussion

This section describes an assessment technique to explore the quality and relevance of the Onto4AIR2 ontology. Participants were organized in two groups: 1) managers - developers and 2) potential users.

5.1. Assessment by managers and developers

Three managers and developers experts in semantic web applications and repositories used the CQ-based approach described in [40] to verify the satisfiability of CQs, (see Table 1). Figures 5, 6 and 7 relate to the answers from CQ1 to CQ4; `skos:note` and `rdf:isDefinedBy` properties include class definitions.

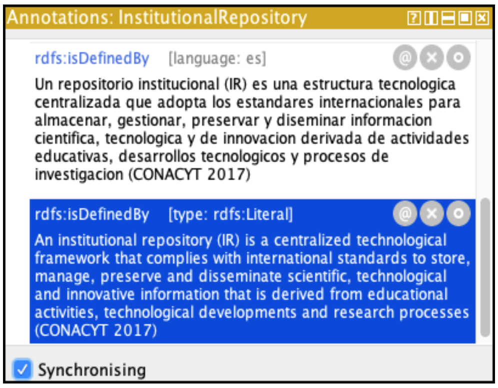

Fig. 5. Definition of a repository

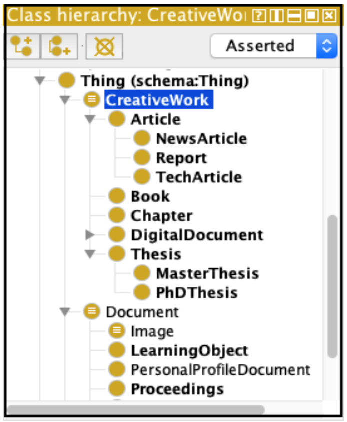

Fig. 6. Types of educational resources

Figure 8 shows information related to CQ5 and CQ6, while the answer to CQ7 is retrieved by the following SPARQL query:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX onto: <http://localhost:8080/mauxOntologies/onto4AIR2.owl#>
SELECT ?subject ?object ?name where { ?subject onto:worksIn ?object .
?subject foaf:firstName ?name . } order by ?subject
```

Figure 9 shows the labels of a SKOS for the knowledge areas in Spanish and English languages of

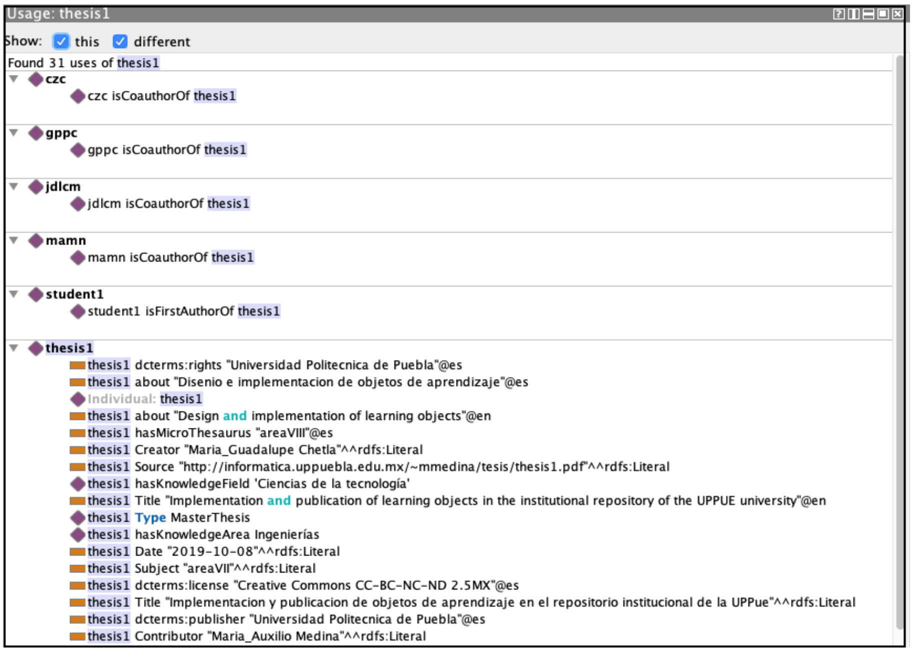

Fig. 7. Description of a thesis using ontology properties

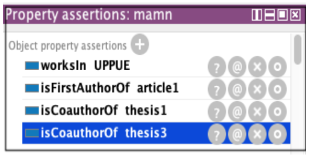

Fig. 8. An excerpt of information related to CQ5 and CQ6

CONACYT, this SKOS and the ontology are available at:

- http://www.mauxmedina.com/re3/re3.html,
- http://www.mauxmedina.com/re4/re4.html

In summary, managers and developers determine that Onto4AIR2 ontology represents CQs and their answers using its terminology, they also correct all the inconsistencies detected by the reasoners Fact++ 1.6.5, Hermit 1.4.3.456, and Pellet.

5.2. Assessment by potential users

A group of students and teachers participated in an exploratory and self-management survey designed to gather their perceptions about the Onto4AIR2 ontology, (see Table 4). The values high, medium, and low represent the experience.

A questionnaire gathered information about correctness and language expressiveness, this had five closed questions, three dichotomy questions, a multiple-

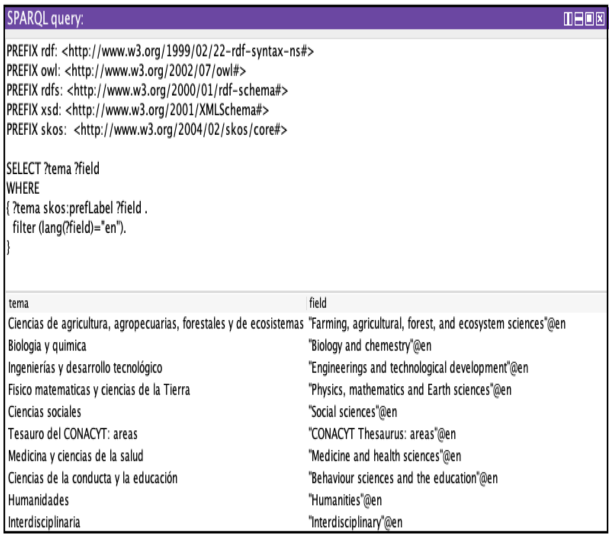

Fig. 9. Knowledge areas proposed by CONACYT

Table 4

Profiles of potential users

Number of teachers:	12 (6 female, 6 male)
Number of students:	18 (10 female, 8 male)
The average age of teachers:	47 years
The average age of students:	25 years
The average daily use of the internet:	8.5 hours
Experience with repositories:	26 high, 4 medium
Experience with ontologies:	28 high, 2 medium

Table 5

Perception of users

Element	Average
Class names:	4.8
Class hierarchy, (see Figure 3) :	3.8
Definitions of classes:	4.3
Description of an IR:	4.1
Object properties, (see Table 3):	5.0
Overall average:	**4.4**

Table 6

Results for dichotomy questions

Task	Number of answers
Complete the description of a thesis	28 true, 2 false
Find the types of educational resources	30 true, 0 false
Check if a person is a co-author of a specific thesis	26 true, 4 false
Percentage:	**93.33%** true, **6.67%** false

choice question, and a section for suggestions. Unlike the last section, the others were mandatory. There were three days to answer the questionnaires sent by mail. Then, a manager integrated the results of closed and dichotomy questions, while the answers to open questions were analyzed by the managers and developers group during online sessions. The suggestions and comments produced newer versions of the ontology; users received the latest version and the questionnaires for a second time.

Table 5 shows the elements referred to in closed questions and the average results, (a Likert scale of five values varying from 1 = total disagree to 5 = total agree was used for this purpose). The overall average suggests that users had a positive perception.

Table 6 shows the tasks referred to dichotomy questions, the answers were true or false according to task accomplishment. The incorrect answers in row 2 come from two users that only wrote the title of the thesis; in row 4, four users did not distinguish between author and co-author. The overall percentage of correct answers was 93.33%. Table 7 shows the possible choices and the number of users that identified each one, additional uses were democratization and visualization of information. Furthermore, the following suggestions were gathered:

- Include other types of persons such as designers and librarians

Table 7

Use	Number of users
Develop semantic web services	24
Construct machine readable datasets	26
Enrich data from repositories	23
Add new information retrieval mechanisms	23
Knowledge acquisition	28
Reuse knowledge	23
Share knowledge	25
Total of potential users:	**30**

- Add `Play` and `Work` as other types of educational resources
- Change the label of `updateDate` data property to `modificationDate` or `lastUpdate` of the Institutional Repository class
- Represent other types of academic collaborations

Finally, the Net Promoter Score (NPS) was also introduced into the questionnaires that were sent for the second time to users, this is a standard question widely used to estimate users' attitude about a tested object [42]. Users were asked for a numerical answer between 0 and 10, (minimum and maximum value); the answers between [0,6], [7,8] and [9,10] intervals are associated with a detractor, indifferent or promoter attitude, respectively. The average of the answers is considered as the final value for the tested object. The final value of the NPS for Onto4AIR ontology was 8.9; this value indicates a positive users' perception.

6. Conclusions

This paper described the Onto4AIR2 ontology to represent theses from repositories. The ontology uses established vocabularies like `foaf`, `schema.org`, `skos`, ontologies, and SPARQL queries, this formally represents persons, theses and their properties in English and Spanish languages. Persons are modeled by using the foaf vocabulary, educational organizations and educational resources are represented by using schema.org while a SKOS was proposed to include knowledge areas to classify thesis according to CONACYT. The vocabulary of this ontology can be shared between web users.

The scope of Onto4AIR ontology was determined by a set of competency questions that can be expressed as SPARQL queries, the answers to some of these questions were illustrated with figures that showed the use of the Protégé ontology editor.

Relevant related works have showed the successful use of ontologies for knowledge acquisition, integration and reconciling information in different contexts; however, information about assessment is limited. In this paper, managers, developers and potential users participated in an ontological assessment method. On one hand, the managers and developers used reasoners for validation of logical consistency and followed a CQ-based approach, the experimental results allowed them to establish that the ontology was able to represent CQs and their answers using its terminology. On the other hand, an exploratory and self-management survey was done to gather the opinion of potential users; the results indicated a positive perception to promote the ontology.

Onto4AIR ontology enables knowledge sharing and acquisition from the repositories domain and the construction and management of machine-readable datasets, this represents an alternative to increase quality, reliability and reusability of data from repositories. The ontology is of utility as this gives institutional visibility, this can be adopted by other educational organizations with little adjustments such as the incorporation of new terms, object properties, rules or restrictions.

The paper is a preliminary guide that provides of foundations for systematic enrichment of data and knowledge-based applications. At present, we are working with the implementation of software modules to transform IR data into ontologies instances directly. We expect that our process foster the benefits of open access policies.

As future work, we plan to work in the implementation of new extensions that explore potential power of automated reasoning and support data integration for later analysis.

References

[1] CONACYT, Lineamientos generales para el repositorio nacional y los repositorios institucionales. Repositorio Nacional, Sistema Integrado de Información sobre Investigación Científica, Desarrollo Tecnológico e Innovación (SIICYT), 2014, Consejo Nacional de Ciencia y Tecnología. Accessed: 2017-02-25.

[2] CONACYT, Lineamientos técnicos para el repositorio nacional y los repositorios institucionales. Repositorio Nacional., Sistema Integrado de Información sobre Investigación Científica, Desarrollo Tecnológico e Innovación (SIICYT), 2014, Consejo Nacional de Ciencia y Tecnología. Accessed: 2017-02-25.

[3] D. Gayatri, Capturing intellectual capital with an institutional repository at a business school in India, *Library Hi Tech* **26** (2008), 110–125, Accessed: 2018-04-21. https://doi.org/10.1108/07378830810857843.

[4] CUDI, Red Mexicana de Repositorios Institucionales, Corporación Universitaria para el Desarrollo de Internet A. C. Internet 2 México, 2019, Accessed: 2019-08-06.

[5] NR, Repositorio Nacional, Gobierno de México. Consejo Nacional de Ciencia y Tecnología, 2019, Accessed: 2019-08-06.

[6] C. Lagoze and H. Van de Sompel, The Open Archives Initiative: Building a Low-barrier Interoperability Framework, in: *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '01, ACM, New York, NY, USA, 2001, pp. 54–62. ISBN 1-58113-345-6. doi:10.1145/379437.379449.

[7] DCMI, DCMI Metadata Terms, DCMI Recommendation, Dublin Core Metadata Initiative, 2001, CDMI Usage Board. https://www.dublincore.org/specifications/dublin-core/dcmi-terms/.

[8] T.R. Gruber, Toward principles for the design of ontologies used for knowledge sharing, *International Journal of Human-Computer Studies* **43**(5–6) (1995), 907–928. doi:10.1006/ijhc.1995.1081.

[9] M.A. Medina, A. Sánchez, O. Cervantes, R.C. Medina, J. De la Calleja and A. Benitez, Representación semántica de conocimiento operativo y de dominio para repositorios institucionales. Registro público del derecho de autor. Number: 03-2017-042511235500-01. 2017., 2017, Accessed: 2017-04-26.

[10] G. van Heijst, A.T. Schreiber and B.J. Wielinga, Using explicit ontologies in KBS development, *International Journal of Human-Computer Studies* **46**(2–3) (1997), 183–292. doi:10.1006/ijhc.1996.0090.

[11] A. Seaborne and E. Prud'hommeaux, SPARQL Query Language for RDF, W3C Recommendation, W3C, 2008, http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/.

[12] Eprints, Eprints services, University of Southampton. Electronics & Computer Science, 2019, Accessed: 2018-08-16.

[13] Digital Commons bepress, Digital Commons, 2016, Accessed: 2019-08-06.

[14] K. P., Exist db: the open source native xml database, 2019, Accessed: 2019-04-21.

[15] Greenstone, Greenstone Digital Library Software, University of Waikato, New Zealand. New Zealand Digital Library Project, 2019, Accessed: 2010-03-17.

[16] D. Wilcox, Supporting digital preservation and access with Fedora, *IFLA World Library and Information Congress (WLIC 2017)* (2017), 5. http://library.ifla.org/1758/1/150-wilcox-en.pdf.

[17] S. Mackenzie, M. Barton, M. Bass, M. Branschofsky, G. McClellan, D. Stuve and J. Walker, Dspace: an open source dynamic digital repository, *D-Lib magazine* **9**(1) (2003), ISSN1082-9873. DOI: 10.1045/january2003-smith.

[18] Jisc open access services, OpenDOAR: Directory of open access repositories, 2019, Accessed: 2018-04-21.

[19] Fari H., Khan S., *Exchanging data from institutional repositories to the semantic web*, Lap Lambert Academic Publishing, 2015, ISSN: 978-3659780271.

[20] B. Arsić, M. Đ*okić-Petrović, P. Spalević, I. Milentijević and D.R.M. Živanović, SpecINT: a framework for data integration over cheminformatics and bioinformatics RDF

repositories, *Semantic web - interoperability, usability, applicability* **10** (2018), 1–19, DOI 10.3233/SW-180327. Accesed: 2020-03-09. http://www.semantic-web-journal.net/content/specint-framework-data-integration-over-cheminformatics-\and-bioinformatics-rdf-1.

[21] G. Schreiber and F. Gandon, RDF 1.1 XML Syntax, W3C Recommendation, W3C, 2014, http://www.w3.org/TR/2014/REC-rdf-syntax-grammar-20140225/.

[22] K. Atanas, O. Damyan and K. Vesselin, An ontology representation and data integration (ORDI) framework (2008).

[23] O. use for semantic e-science, Boyan Brodaric and Mark Gahegan, *Semantic web - interoperability, usability, applicability* **1** (2010), 149–153, DOI 10.3233/SW-2010-0021. Accesed: 2020-03-09. http://www.semantic-web-journal.net/content/ontology-use-semantic-e-science.

[24] B. Bishop, A. Kiryakov, D. Ognyanoff, I. Peikov, Z. Tashev and R. Velkov, OWLIM: a family of scalable semantic repositories, *Semantic web - interoperability, usability, applicability* **2** (2011), 33–42, DOI 10.3233/SW-2011-0026. Accesed: 2020-03-09. http://www.semantic-web-journal.net/content/owlim-family-scalable-semantic-repositories.

[25] G. Guizzardi, Theoretical foundations and engineering tools for building ontologies as reference conceptual models, *Semantic web - interoperability, usability, applicability* **1** (2010), 3–10, DOI 10.3233/SW-2010-0015. Accesed: 2020-03-09. http://www.semantic-web-journal.net/content/theoretical-foundations-and-engineering-tools\-building-ontologies-reference-conceptual-model.

[26] J. Estevão and F. Strauhs, Information retrieval in institutional repositories: proposal of an ontology for historical organizational memory, *Qualitative and quantitative methods in libraries* **4**(4) (2017), 983–992. http://78.46.229.148/ojs/index.php/qqml/article/view/297.

[27] F.J. Cantú and H.G. Ceballos, A multiagent knowledge and information network approach for managing research assets, *Expert Systems with Applications* **37**(7) (2010), 5272–5284. doi:10.1016/j.eswa.2010.01.012.

[28] A. Casas-Bayona and H.G. Ceballos, Integrating semi-structured information using semantic technologies - an evaluation of tools and a case study on university rankings data, in: *Proceedings of 3rd International Conference on Data Management Technologies and Applications - Volume 1: DATA,*, Science and Technology Publications, 2014, pp. 357–364, INSTICC. ISBN 978-989-758-035-2. doi:10.5220/0005004203570364.

[29] G.G. Pinilla, Aplicación de la lógica de predicados para la realización de inferencias en documentos de pasantía descritos a nivel ontológico, *Ingenierías USBMed* **7**(1) (2016), 11–19. doi:10.21500/20275846.1802. https://revistas.usb.edu.co/index.php/IngUSBmed/article/view/1802.

[30] Y. Guo, Z. Pan and J. Heflin, LUBM: a benchmark for OWL knowledge base systems, *Journal of Web Semantics First Look* **3** (2005), 158–182, DOI: 10.1016/j.websem.2005.06.005. Accesed: 2020-03-09. http://dx.doi.org/10.2139/ssrn.3199255.

[31] D. McGuinness and F. van Harmelen, OWL web ontology language overview, W3C Recommendation, W3C, 2004, http://www.w3.org/TR/2004/REC-owl-features-20040210/.

[32] G. Demartini, I. Enchev, J. Gapany and P. Cudré-Mauroux, The Bowlogna ontology: fostering open curricula and agile knowledge bases for Europe's higher education landscape, *Semantic web - interoperability, usability, applicability* **4** (2013), 53–63, DOI 10.3233/SW-2012-0064. Accesed: 2020-03-09. http://www.semantic-web-journal.net/sites/default/files/swj186_1.pdf.

[33] L. Ivanovic, D. Ivanovic and B.D. Surla, *Improving dissemination of human knowledge by exporting data from research information systems*, in *Proceedings of the Management, Knowledge and Learning International Conference 2014*, ToKnowPress, 2014, pp. 711–718, Chapter Human capital without borders: knowledge and learning for quality of Life;. https://ideas.repec.org/h/tkp/mklp14/711-718.html.

[34] H. Farid, S. Khan and M.Y. Javed, DSont: DSpace to ontology transformation, *Journal of Information Science* **42**(2) (2016), 179–199. doi:10.1177/0165551515591406.

[35] T. Tudorache, Ontology engineering: current state, challenges, and future directions, *Semantic web - interoperability, usability, applicability* **1** (2010), 1–14, Accepted. Accesed: 2020-03-09. http://www.semantic-web-journal.net/system/files/swj2313.pdf.

[36] M. Fernández López, Overview of methodologies for building ontologies, in *Proceedings of the IJCAI-99 workshop on ontologies and problem-solving methods (KRR5)*, 1999. http://www.lsi.upc.es/~bejar/aia/aia-web/4-fernandez.pdf.

[37] N.F. Noy and D.L. McGuinness, Ontology development 101: a guide to creating your first ontology, Technical Report, Standford University, 2001. http://www-ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html.

[38] M.A. Musen, The Protégé project: a look back and a look forward, *AI Matters* **1**(4) (2015), 4–12. doi:10.1145/2757001.2757003.

[39] N.F. Noy and C.D. Hafner, The state of the art in ontology design: a survey and comparative review, *AI Magazine* **18**(3) (1997), 53–74. http://www.aaai.org/Library/Magazine/vol18.php#Fall.

[40] C. Bezerra, F. Freitas and F. Santana, Evaluating ontologies with competency questions, in: *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 03*, WI-IAT '13, IEEE Computer Society, Washington, DC, USA, 2013, pp. 284–285. ISBN 978-0-7695-5145-6. doi:10.1109/WI-IAT.2013.199.

[41] M.A. Medina Nieto, D.A. Díaz, J. de la Calleja Mora, L. Zacatzontetl Hernández and M. Zacatelco Pérez, An ontology-based approach to describe collaborative work by reusing and enriching data from an institutional repository, *Electronic Notes in Theoretical Computer Science* **354** (2020), 129–139, Proceedings of the Eleventh and Twelfth Latin American Workshop on Logic/Languages, Algorithms and New Methods of Reasoning (LANMR). doi:https://doi.org/10.1016/j.entcs.2020.10.010. http://www.sciencedirect.com/science/article/pii/S1571066120300864.

[42] J.S. Erin Bradner, User experience and likelihood to recommend: linking UX and net promoter score, U.E.P.A.I. Conference, ed., 2012.