Semantic Representation of Information by Ontological Networks to Improve Knowledge Management in Higher Education

Roberto A. García-Vélez[1], Jorge Andrés Galán-Mena[1], Ahmed Dahroug[2], Vladimir E. Robles-Bykbaev[1] and Martín López-Nores[3]

[1]GI-IATa, Cátedra UNESCO Tecnologías de apoyo para la Inclusión Educativa, Universidad Politécnica Salesiana, Calle Vieja 12-30 y Elia Liut, Cuenca, Ecuador

[2]Arab Academy for Science, Technology and Maritime Transport, Cairo, Egypt

[3]AtlanTTic Research Center for Telecommunication Technologies, University of Vigo, Vigo, Spain

Keywords: Ontologies, Higher Education, Knowledge Management, Semantics.

Abstract: Institutions of Higher Education (IHE) seek to respond to new ways of conceiving and projecting higher education embodied in the profile of a professional. Faced with this challenge, educational models focus mainly on the theoretical-practical references of critical training, constructivism and collaborative learning. Although there are several knowledge management solutions in the field of higher education, which have managed to formalize the organizational structure of academic institutions in ontologies, so far none of these proposals divides and makes explicit the development that the actors of the ecosystem (students and educators) take over time. Our proposal is to construct a semantic representation of the academic ecosystem by implementing an ontological network that allows managing the knowledge generated in a more efficient way. This architecture is intended to be used as an instrument to support the academic body and for the centralization of information. To achieve our objective, we carry out a reengineering process of relevant institutional documents, such as the academic record of the students, in each of their facets of learning. We have interviewed specialists in the area and reuse academic domain ontologies to form a consistent knowledge base.

1 INTRODUCTION

Higher education institutions periodically generate large volumes of data associated with the academic field (Mora-Arciniegas et al., 2017). Unfortunately, the information that revolves around the university academic environment is usually distributed among multiple sources of subsystems that use different databases or digital repositories (Mora-Arciniegas et al., 2017; Aminah et al., 2017). Having a diversity of unstructured formats from different sources of information generates a restriction of interoperability among them. This particularity restricts the ability to efficiently manage and exploit the full potential of this information in the form of knowledge, which could be used in a timely and relevant manner by the different actors of the academic ecosystem. Although there are several knowledge management systems in the field of higher education focused on solving this difficulty (among them, transactional systems that allow associating information between different subsystems), the information retrieved does not have semantic relationships that would allow subsystems to identify even whether two pieces of information relate to the same topic or entity.

Based on this background, in order to help universities ensure an efficient management, sharing, search and reuse of information, we have proposed the creation of an ontological network that allows linking the information generated and take it to a higher level, making sure the right information gets to the right people at the right time to make the right decisions. In this paper, we describe an experiment in which these ontologies supported knowledge management within the Universidad Politécnica Salesiana (UPS), located in the city of Cuenca, Ecuador. From our findings, we analyze the strengths of knowledge administration in higher education supported by semantic technologies. We also suggest some avenues to support this activity within higher education institutions.

The idea of using semantic technology is to enable universities to improve the mechanisms for searching for information and knowledge. This is particularly necessary for academic institutions of higher educa-

García-Vélez, R., Galán-Mena, J., Dahroug, A., Robles-Bykbaev, V. and López-Nores, M.
Semantic Representation of Information by Ontological Networks to Improve Knowledge Management in Higher Education.
DOI: 10.5220/0006920801610168
In *Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2018) - Volume 3: KMIS*, pages 161-168
ISBN: 978-989-758-330-8

tion, where sharing, reusing and inferring new knowledge plays a central role.

In Section 2, some related works developed in the field of higher education institutions are described. In Section 3, the development of the proposal is detailed indicating the methodology used (which implies the specification of requirements for the ontology network), the reuse of resources, the reengineering carried out to reuse the ontological resources and the conceptualization of the ontological network proposal based on the UPS datasets. Continuing, in Section 4 we present the results of the experimentation, and Section 5 contains the conclusions of our work.

2 RELATED WORK

There is a broad field of work related to knowledge management in the academic domain using ontologies. Several authors have used them as a means to represent different perspectives of the academic domain. For example, the authors of (Laoufi et al., 2011) proposed an approach to establish an organizational Memory of the University System (MUS). The MUS is not only on pedagogical knowledge (topics, subjects, courses, etc.) but also works with other knowledge used in the context of a university system, whether administrative, educational or related to the intellectual capital of the university. For the construction of the ontology, the actors are trained to explain their knowledge, promoting in this way a collaborative culture of creation and indexation of knowledge. The captured knowledge provides a resource dedicated to the representation of ontological knowledge, which facilitates the search and navigation between concepts related to the academic system. The architecture of the system is able to integrate different ontologies, and enables different types of reasoning and intelligent information retrieval.

A similar approach was done by the Universidad Técnica Particular de Loja (UTPL) research university (Mora-Arciniegas et al., 2017). They propose the use of semantic technologies to represent academic knowledge through ontologies, creating a set of comprehensible and interoperable data. To fulfill their objective, they developed an ontological network called *Linked Academic Data* (LDA) for the representation of the organizational structure and teaching planning. LDA was comprised of ontological resources (FOAF, AIISO, VCARD, ORG, DCTERMS, VIVO), non-ontological resources that required reengineering to transform them into accessible resources, and the creation of a new vocabulary.

A more recent proposal can be found at (Melgar and Quilca, 2016), consisting in an architecture for organizational memory systems in higher education institutions. The proposed architecture for organizational memory is based on CESM model and CommonKADS methodology, concerned with the knowledge representation for semantic search of documents. This approach allows adding semantic content to the documents and to the information, enabling the construction of knowledge bases. The semantic content added to the documents allows the retrieval of information based on inference, unlike the relational database that is based only on the content of a field. To achieve this, they add semantic annotations to documents and link them to defined ontologies.

Finally, the authors of (Gasmi and Bouras, 2018) proposed a system that allows to improve vocational education, conducting an analysis of the gap between education and industry. To this aum, they used a system of semantic concordance based on ontological models that allows to analyze these two domains. In particular, the educational and industry profiles are represented as profiles of O*NET competency frameworks, which allows to analyze the gap according to the structure of higher education and the peripherals required by the industry in the social context.

3 DEVELOPMENT

In order to unify the scattered information in several database repositories that occurs around the different learning processes of higher education institutions, we propose the construction of a new formal model that captures the corresponding knowledge. We chose the NeOn framework (Suárez-Figueroa et al., 2012b) as the basic was chosen for our implementation, because it offers features that allow to manage the lifecycle of the ontological network very easily.

3.1 Elicitation of Requirements

In order to identify the purpose of creating the formal model, who the end users would be and the features it would enable, we went through a process of specifying ontological requirements (Suárez-Figueroa and Gómez-Pérez, 2012). To fulfill the different tasks specified by the NeOn methodology, a group of specialists was formed, composed of practitioners and specialists from the academic side of higher education institutions, who issued their criteria according to their role in each task. Table 1 presents the group of most frequent terms obtained from an interview for

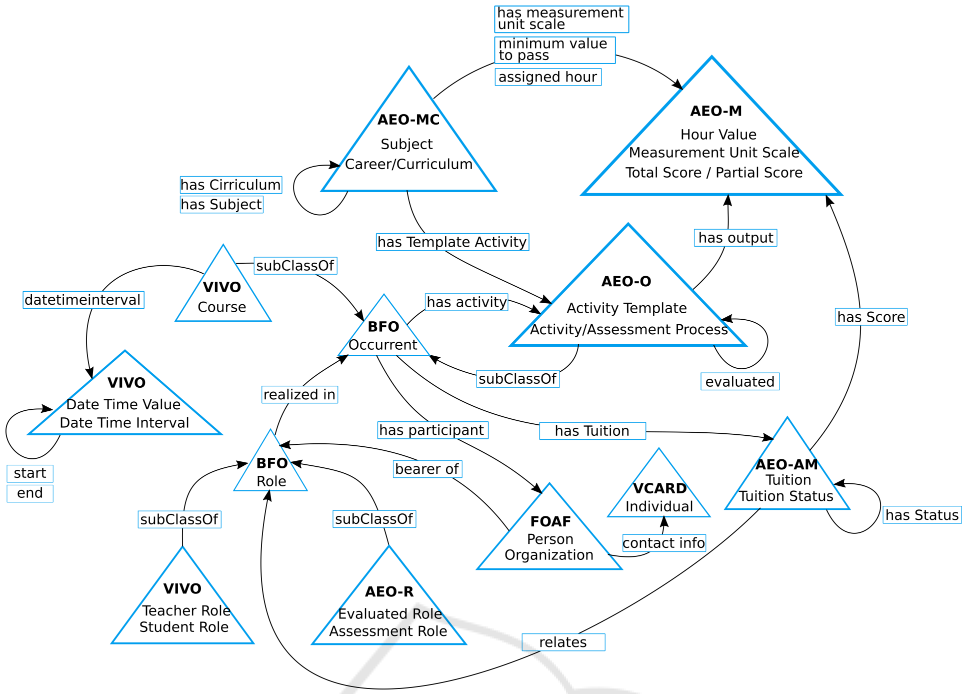

Figure 1: Modular scheme of the ontological network of the academic ecosystem.

the identification of functional requirements, which were already grouped and validated.

Table 1: Group of most frequent terms obtained.

Code	Item	Competence Questions
TG87	Micro curriculum	PC18, PC26
TG100	Class period	PC11
TG114	Task	PC7, PC22
TG41	Teacher	PC1, PC3, PC6

3.2 Formalization of the Scheme

Starting from the requirements elicitation process, we sought existing resources to reuse and reengineer (Suárez-Figueroa et al., 2012a), ending up with the following selection:

- **Friend of a Friend (FOAF)** (Kalemi and Martiri, 2011) which describes a basic social network of people and their relationships;
- **VCARD** (Iannella and McKinney, 2014) is an ontology that describes information of people and organizations
- **VIVO** has a focus for a community of research.
- **Basic Formal Ontology (BFO)** (Arp et al., 2015) is a top-level ontology that helps the integration of multiple ontology in a single formal framework.

In the construction of our ontological network it was decided to modularize the classes, both those built from the reengineering of non-ontological documents and the selected ontologies to be reused.

Figure 1 shows the scheme of our network, where we have reused VCARD, FOAF, BFO and VIVO ontologies by means of different links and selection of components according to the necessary domain through a process of alignment and modularization of ontologies. Our academic ecosystem ontology (AEO) is separated into 5 modules: Micro Curricular (MC), Roles (R), Academic Metrics (AM), Metrics (M) and Occurrent (O). These modules participate as links between the reused ontologies, which have different dependencies and cover specific domains within the same context.

The main classes defined in our ontology are the following:

- **Person:** It is a representation of a person, reused from the FOAF ontology.
- **Evaluated Role:** Inherent role to a person that fulfills the function of being evaluated within a process.
- **Subject:** Representation of a subject that is taught in a course within a study program.
- **Career:** Consensus of the name that is given to a study program that leads to the granting of a degree to practice a profession.
- **Curriculum:** Organization of the subjects and requirements of a career that determine the pro-

fessional profile of the student.

- **Tuition:** Representation of the document of record of notes of a student enrolled in a subject belonging to a curriculum.
- **Tuition Status:** status of a license plate, it can be: approved, failed or canceled.
- **Identity Card:** Identity document that belongs only to one person.

Each one of the modules of the ontological network belong to a specific context, allowing us to better manage the different components imported into our network. For the construction of the modules, the **Protégé**[1] tool was used as shown in Figure 2, where each one of the ontologies is developed separately taking into account the alignment of the components through links that are classes of other modules. To verify the operation of the network as a first step, all the modules were imported into the same project. Finally, we ran the **HermiT**[2] reasoner to verify that there were no inconsistencies in the axioms.

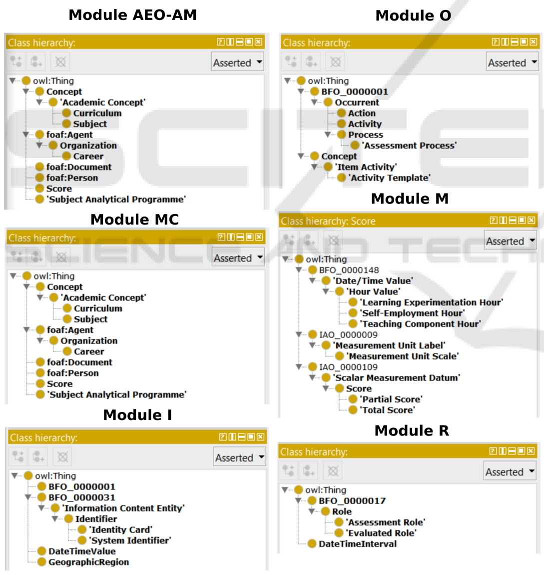

Figure 2: Protégé capture of the classes of the ontological network modules of the academic ecosystem.

3.3 Information Mapping

The construction of the ontological network implies both the formalization of the RDF (Klyne and Carroll, 2006) scheme. Starting from this referential semantic framework we start with the population of the ontological network on which information can be inferenced. Commonly, the transactional systems in which the data acquisition of the academic ecosystem is carried out in higher education institutions have two characteristics: the systems do not use a single repository of information but are found in different separate databases, and they use relational databases. The management of multiple entity-relational models implies that the data in each of these repositories must be interpreted to derive their meaning, but the interpretation lends itself to ambiguity. On the contrary, an expressive language of knowledge such as ontologies allow interpreting the knowledge unambiguously.

A strategy that can be used to transfer the information stored in a traditional database to a semantic network is through the data mapping procedure. Data mapping allows populating the ontological network with a semantic encoding of instances in RDF format in an ontological scheme. As can be seen in Figure 3, on the left side there is a fragment of the model entity relationship of an academic system of students of a higher education institution and on the right its equivalence in the model of the ontological ecosystem, in which we can map the information between schemes starting from their tables in the relational model:

- **Student:** The table contains the personal information of the student, including ID, name and surname, email and date of birth. The representation of a student in the ontological model separates concepts of the role as a student and the person itself.
- **Subject:** The table contains the information of the name of a subject offered. The mapping in the ontological model is similar to the table where there is only one data property of the name of the subject.
- **Tuition:** It contains a field with the offer date and a one-to-many relationship from Student and also from Subject. The ontological model for this section already implies that in order to enroll a student in a subject there must be a process that participates as an intermediary –in this case, the instance of the Course class must also be instantiated.

This first mapping exemplified from the information of a student named Peter Smith allows us to define in a logical way a first interpretation of the information that is stored in an entity relationship model and its equivalence in an ontological model as instances of the defined classes in the modular schemes of the ontological network of the academic ecosystem.

As there are several transactional systems in the same academic ecosystem in higher education institutions, we can show an example of an excerpt from the relationship model of the employee administration

[1] https://protege.stanford.edu/
[2] http://www.hermit-reasoner.com/

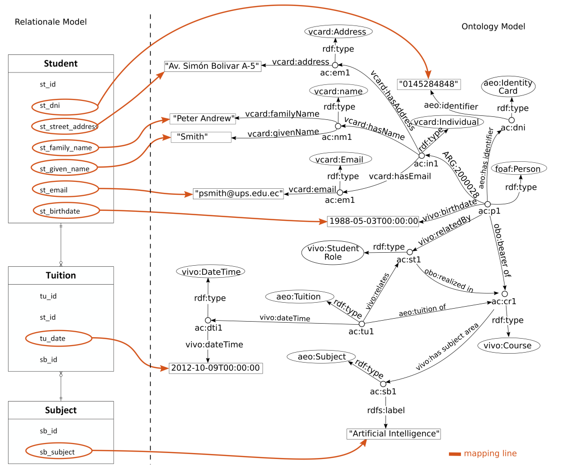

Figure 3: Mapping scheme between the model entity relationship of the academic system against its version in an ontological.

system in which a data mapping is applied in order to transform it into instances to populate our ontological network as presented in Figure 4, where we can interpret tables of the relational database as follows:

- **Employee:** The table contains the personal information of an employee, as for the Student in Figure 3. In the ontological model as well as in the previous model, the appointment of professor is taken as a position within an organization and the person as something totally different, where personal information can be mapped in a similar way as in Student.
- **Institutional Position:** It contains the information of the positions that exist in the institution. It is mapped in a similar way with a data property with the name of the position, but we make explicit that there is an organization to which that position is related, which in this case would be the institution of higher education.
- **Employee_Institutional_Position:** It is an intermediate table between Employee and Institutional Position where the start and end date fields of the position are also counted. In the ontological model, a time interval is assigned to the position node by defining the start date as the end date as instances.

The mapping of this second fragment of the relational database of the employee management system demonstrates the integration of information from another source of information under the same schema of the ontological network, taking into account that the employee for reasons of exemplification was called Peter Smith as in the mapping of the academic system.

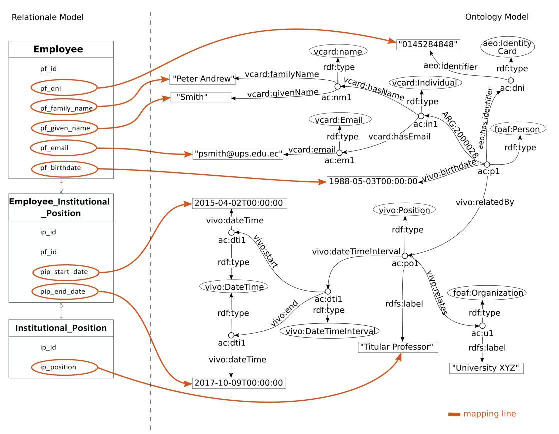

Figure 4: Mapping scheme between the model entity relationship of the employee management system against its version in an ontological model.

3.4 Axioms and Rules of Unification of Persons

The AEO-I module helps us to identify people who have unique identifiers by means of axioms about the classes and their relationships, taking advantage of the fact that OWL allows us to enrich the meaning of the properties of the objects ought certain axioms, one of these axioms are the functional properties that are those that, given an individual, by means of this property the individual can be at most related to a single individual (Horridge et al., 2004), for the case of our module we have made functional the object property "identifier of" which helps us to discern if an individual is the same as another, as we can see in Figure 5, an example is presented where, through the object property "identifier of", a node of the Identity Card class is related to two people: Andrew J. M. and A. Jone, which implies that Andrew JM and A. Jone are the same individual, this logically implies that the two nodes are merged through the "sameAs" relationship of OWL.

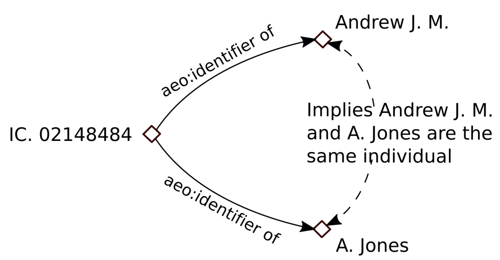

Figure 5: Example of the “identifier of” functional property of the AEO-I module.

The functional property would be a good option to detect similarity of nodes in our ontological network but this scenario is possible only if our instances of the

```
CONSTRUCT {
  ?nodeDNI1 owl:sameAs ?nodeDNI2 .
}
WHERE {
  ?nodeDNI1 a aeo:AC0000008 .
  ?nodeDNI1 aeo:identifier ?identifier1 .
  ?nodeDNI2 a aeo:AC0000008 .
  ?nodeDNI2 aeo:identifier ?identifier2 .
  FILTER (?identifier1 = ?identifier2) .
  FILTER (?nodeDNI1 != ?nodeDNI2) .
}
```

Figure 6: SPIN rule capture to find individuals of Identity Card class that are similar.

"Identity Card" class have already gone through a process to find their similarity. As we could see in the mapping process with the example of Peter Smith this person is represented in the system as two instances p1 and p2 corresponding to the Person class, but in turn both p1 and p2 have different IDs instantiated as *dni1* and *dni2*, so by not pointing to the same ID could not operate the functional property.

To solve this problem, we must first unify the instances of the **"Identity Card"** class by means of rules using the **SPIN** language as presented in Figure 6, where we compare whether the data properties in the instances of the **"Identity Card"** class.

When we establish the rules and define the axioms we will use, we can discover information between the instances of people who claim that one individual is similar to another. Returning to our example where we had two nodes p1 and p2 that represented Peter Smith, as a first step we analyzed their identifications dni1 and dni2 by means of the rule defined in SPIN where they would be unified forming a logical level as a single node and later through the Functional property axiom of "identifier of" will infer that p1 and p2 are the same individual. From a general perspective we have integrated the information of two different transactional systems in the same semantic context, but at the same time we also integrate the information under the same context, in our case an actor can have the information of his academic process, but also coexist in the same ecosystem as an employee of the university.

4 EXPERIMENTATION AND RESULTS

After checking the correct functioning of our schemes in the ontological network we have migrated the information corresponding to two systems of a higher education institution, the first system contains information corresponding to the students since 1993 and the second corresponding to the personnel management, each of these systems uses a traditional database of Univerisdad Politécnica Salesiana. The **Karma tool**[3] helps us to exploit what is exposed in the data mapping, as we can see in Figure 7. Based on the views of the two transactional systems, we were able to transform the information stored in the tables into an RDF data structure that is consistent with our ontological network.

Once the mapping process was finished we obtained two sets of instances in the form of RDF structures:

- The first set, coming from the migration of the transaction management system of employees, generated around 30,000 triples, which are instances of different classes of our ontological network detailed in Table 2 where it can be seen that there are about 2,000 employees in the system.
- The second set of data, corresponding to students, has 99879 records of people as can be seen in the breakdown of classes in Table 3. This set generated more than 2 million triples in our exported file.

Table 2: Results of the Mapping of the Transactional Employee System.

Class	Instances
foaf:Person	2236
vcard:Email	2089
aeo:Identity Card	2236
vcard:Name	2236
vcard:Individual	2236

Table 3: Results of the Mapping of the Transactional Students System.

Class	Instances
foaf:Person	99879
vcard:Email	78043
aeo:Identity Card	99879
vcard:Name	99879
vcard:Individual	99879

These RDF triplet files and the modules of our ontological network were charged to the **Allegro Graph**[4] triplet database, in order to be able to manage the amount of information that was generated from the mapping and design processes of the ontological network. To comply with the unification of our information through SPIN rules, we use a script of the **Top Braid Composer**[5] tool called Motion SPARQL,

[3]http://usc-isi-i2.github.io/karma/
[4]https://franz.com/agraph/allegrograph/
[5]https://www.topquadrant.com

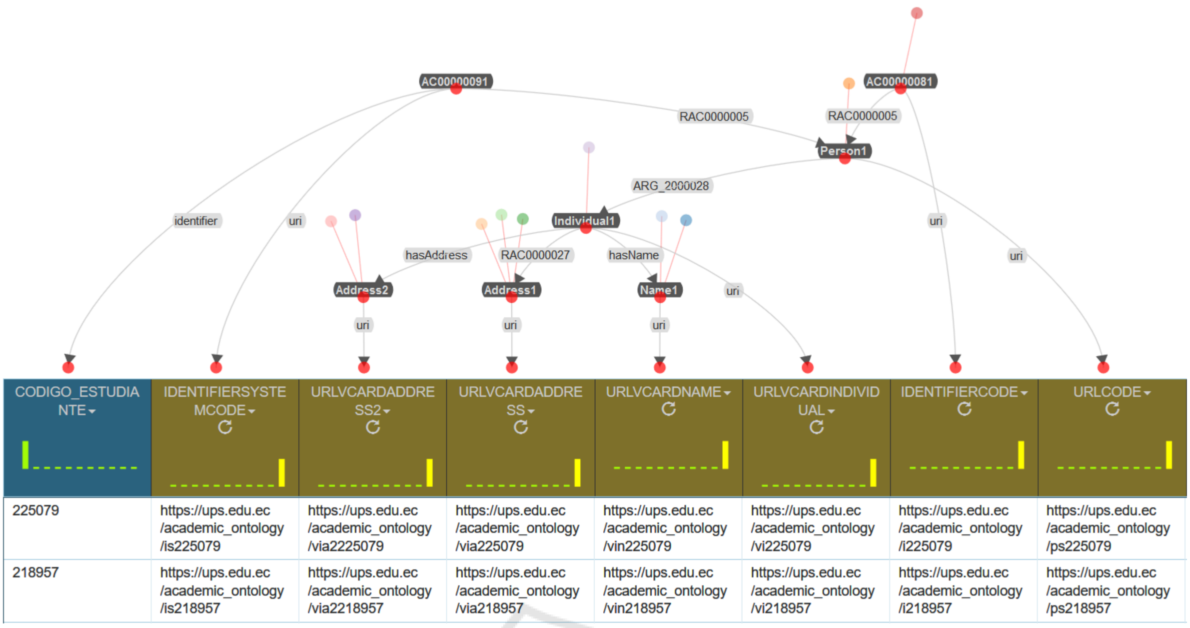

Figure 7: Capture of data mapping using the Karma tool.

which allows us to fulfill the role of an ETL tool. Figure 8 shows the diagram of the script where the triples are first imported through a connection to the triplet databases; then, through a process of building triples from rules with **SPIN**, we generate the similarity relationships through the **OWL sameAs relationship** to unify the information of the IDs and subsequently export the triples in a document in OWL format.

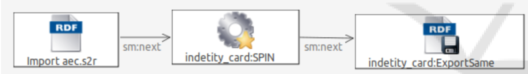

Figure 8: Capture of the SPARQL Motion script for generation of similar triples using SPIN.

Based on the rules with SPIN, we obtained that out of a total of 102,115 people between employees and students of the institution, 373 similarities of individuals belonging to the Identity Card class were created, which means that there are this number of people who are in the two systems, both that of students and that of employees, but they were unified in our ontological network, having the same context.

We loaded the set of pre-processed triples with SPIN rules to the triplet database to populate our ontological network, where by means of a SPARQL query as shown in Figure 9, we were able to verify the unification of the people who had the same ID in our instances of the triples:

The result of the SPARQL query is presented in Table 4, where two columns are presented of which each row is a similarity correspondence of instances of the **foaf class: Person**, individuals with an IRI that

```
SELECT ?nodePerson1 ?nodePerson2
WHERE{
 ?nodeDNI1
      owl:sameAs ?nodeDNI2;
      a aeo:AC0000008;
      aeo:RAC0000005 ?nodePerson1.
?nodeDNI2 aeo:RAC0000005 ?nodePerson2.
}
```

Figure 9: SPARQL query capture in AllegroGraph.

Table 4: Results of SPARQL consultation of people similarity.

Node Person 1	Node Person 2
aeo:psd69	aeo:ps999
aeo:psd7	aeo:ps633
aeo:psd70	aeo:ps102219
aeo:psd71663	aeo:ps8889528
aeo:ps76	aeo:ps105536

starts with **“psd”** followed of a number are corresponding to the instances generated from the employee administration system and the instances that start with ”ps” followed by a number correspond to the student system.

5 CONCLUSIONS

The present investigation addressed the construction of an ontology that formalizes the knowledge of the processes produced from the chair in institutions of higher education, addressing its main actors, both teachers and students. Based on the **NeOn methodology**, we were able to cover the processes involved in the life cycle of our formal model, based on an interview and competency questions for the acquisition of our requirements, which were the starting point for the other processes of both resource reuse semantics as well as the reengineering of non-ontological resources, plus an alignment process to form our ontological network.

We have tackled some problems of the information distributed in several databases of transactional systems, where the interpretation of that information lends itself to ambiguity, where by means of mapping processes plus axioms and SPIN rules we can integrate and unify more than twenty thousand records in a set of instances that populated our ontological network with 2 million triples that more than allows the coexistence of information from different sources helped us to discern what 373 of employees and students were the same people.

As future works it is proposed to be able to use this ontological scheme in two directions: First, discover relevant information that allows to manage in a more efficient way both the human and physical resources of the University. Second to find through the ontological system characteristics that allow us to predict the performance of a student. In this context it is also very important to identify patterns related to students that could lead to a state of student loss or desertion, and be able to propose possible mitigating actions.

ACKNOWLEDGEMENTS

This work has been supported by the European Regional Development Fund (ERDF) and the Galician Regional Government under agreement for funding the AtlantTIC Research Center for Information and Communication Technologies, as well as the Ministerio de Educación y Ciencia (Gobierno de España) research project TIN2017-87604-R.

REFERENCES

Aminah, S., Afriyanti, I., and Krisnadhi, A. (2017). Ontology-based approach for academic evaluation system. In *2017 IEEE 33rd International Conference on Data Engineering (ICDE)*, pages 1569–1574.

Arp, R., Smith, B., and Spear, A. D. (2015). *Building Ontologies With Basic Formal Ontology*. The MIT Press.

Gasmi, H. and Bouras, A. (2018). Ontology-Based Education / Industry Collaboration System. *IEEE Access*, 6:1362–1371.

Horridge, M., Knublauch, H., Rector, A., Stevens, R., and Wroe, C. (2004). Copyright c The University Of Manchester August 27, 2004. *World*, pages 0–117.

Iannella, R. and McKinney, J. (2014). vCard Ontology-for describing People and Organizations. *W3C Group Note NOTE-vcard-rdf-20140522*.

Kalemi, E. and Martiri, E. (2011). FOAF-Academic ontology: A Vocabulary for the academic community. *Proceedings - 3rd IEEE International Conference on Intelligent Networking and Collaborative Systems, INCoS 2011*, pages 440–445.

Klyne, G. and Carroll, J. J. (2006). Resource description framework (RDF): Concepts and abstract syntax.

Laoufi, A., Mouhim, S., Megder, E. H., and Cherkaoui, C. (2011). An Ontology Based Architecture To Support The Knowledge Management In Higher Education. In *Multimedia Computing and Systems (ICMCS), 2011 International Conference on*, pages 1–6.

Melgar, A. and Quilca, A. (2016). An architecture for organizational memory systems in institutions of higher education. *2016 11th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6.

Mora-Arciniegas, M. B., Piedra, N., and Tenesaca-Luna, G. A. (2017). Semantic Representation of Teaching Planning , Pilot Experience at UTPL. In *2017 IEEE 37th Central America and Panama Convention (CONCAPAN XXXVII)*, pages 1–6.

Suárez-Figueroa, M., Gómez-Pérez, A., Motta, E., and Gangemi, A. (2012a). *Ontology Engineering in a Networked World*. Springer.

Suárez-Figueroa, M. C. and Gómez-Pérez, A. (2012). *Ontology Requirements Specification*, pages 93–106. Springer Berlin Heidelberg, Berlin, Heidelberg.

Suárez-Figueroa, M. C., Gómez-Pérez, A., and Fernández-López, M. (2012b). *The NeOn Methodology for Ontology Engineering*, pages 9–34. Springer Berlin Heidelberg, Berlin, Heidelberg.