ABSTRACT

Ontologies occupy a central position in any conceptual model of the Semantic Web, aiming fundamentally to add explicit semantics to the contents available in the Web. In the domain of Brazilian higher education, especially with regard to the National System for the Evaluation of Higher Education (SINAES), still lacks a reference ontology that can provide a basic conceptualization common to the different stakeholders of the domain. In this work we discuss the role of an ontology network developed with the purpose of improving the quality of the data produced by the different actors, institutions and systems within the domain of SINAES, in addition to facilitating their publication as open and connected data. In particular, we highlight the importance of the NeOn methodology to be followed during the development of an ontology network, demonstrating its role in defining the life cycle and in the articulation of the various processes and activities in the OntoSINAES ontological network development process.

Keywords: Ontologies, Semantic Web, Web Science, Ontology Networks, Linked Data, Institutional Evaluation, Evaluation of Higher Education, SINAES, NeOn methodology.
INTRODUCTION

Evaluation is the process by which individuals seek to know and understand the surrounding reality and guide their actions. In this sense, we can assume that no kind of evaluation is essentially neutral since it carries within itself the intention of the evaluator. Luckesi (2011) defined evaluation as a judgment of value over reality in the context of a decision-making process. The institutional evaluation is a systemic and global process that is carried out in an institution in all its dimensions. Belloni (2003) defined the institutional evaluation as “a systematic process of analysis of an activity, facts or things that allow to understand, in a contextualized way, all its dimensions and implications having the goal to stimulate their improvement” (Belloni, 2003, p. 15). For this author, the institutional evaluation must happen in a global way, in which all processes, actions and results are analyzed following a historical and social contextualized way. Following Pinto (2015, p. 28), the institutional assessment “involves multiple observations and the use of tools and consistent criteria having the object examined throughout the process, taking into account its characteristics and particularities”. According to Lima (2015), “the current role of evaluation in the context of educational policies does not refer only to its instrumental and control dimensions, at the service of new modalities of regulation and state meta-regulation of public policies. Educational evaluation, more than that, is one of the maximal substantive expressions of contemporary educational policies, both in terms a national and local scale or in terms of a transnational scale”.

The knowledge the society has about a specific domain is one of the pillars on which it stands and develops itself. Each day, knowledge defines new frontiers for global villages that sometimes overlap to traditional geographical, political, social or economic landmarks. The National System for Evaluating Higher Education in Brazil (SINAES) has produced an immense amount of data during the last few years. However, when we seek to extract some valuable knowledge from the data about the Brazilian higher education, we find that these are not available on the Web or presented in formats that facilitate their use, particularly when we are using some type of computational process over those data. Making the data produced within SINAES more accessible and reusable is very interesting and important for researchers working on different areas of knowledge (Alcantara et al., 2015; Dias Sobrinho, 2010; Firmino, 2013). According to Alcantara et al. (2015) to facilitate data interoperability between such systems, it is important that they are developed following the specification of an ontology-based system, which will enable a semantic enrichment of the corresponding knowledge base. However, based on literature review we made, as well as in the research we done so far, we did not identify any kind of ontology developed specially to facilitate data interoperability in the SINAES domain.

In recent years, the publication of Linked Open Data (LOD) has allowed a significant increase of structured data on the Web. Although the recommendations (and good practices) for the publication of connected open data do not impose the use of ontologies for structuring the data, this is a well-known practice and recommended. In this paper we will discuss the role of networks of ontologies for the publication of open and connected data on the Web based on empirical reality of SINAES. In particular, we are interested to discuss the methodology for the development of ontologies where we intend to contribute for structuring, publication, use and reuse of data, as macro resources of the Web. All this problematic emerged from a research work we are doing, aiming the implementation of an environment - OntoSINAES – for collaborative development of ontologies and vocabularies, and for promoting the sharing of data and representation models in the domain of SINAES (Silva & Belo, 2017). The rest of the paper is structured as follows: section 2 reports some research work related
to this research project, demonstrating the relevance of the explicit formalization of relationships that guarantee the interconnection of ontologies within a network; section 3, presents a network of ontologies for the domain of SINAES, discussing the design and implementation of a network of ontologies, exposing their rationale and revealing its most important aspects; and finally, section 4, reveals some brief conclusions and some guidelines for future work.

BACKGROUND

The contemporary society produces, stores, organizes and retrieves information every day, in large quantities, making it through a set of technologies using the most up-to-date Web technologies. The importance of the Web to today’s society stems from the advances already made, which made electronic information less and less ambiguous for people and machines. From this combination of efforts, the expression “Web Science” was born. Basically, it represents a recent multidisciplinary area that is dedicated for advancing the knowledge about the Web, both from the point of view of the technologies involved and from the various aspects that involve its understanding as a phenomenon that affects different dimensions of the current society.

According to Hall and Tiropanis (2012), in its still short life, the Web Science encouraged the creation of a very interesting community of researchers, possessing today a diverse set of theories and methods, beginning to gather scientific evidences and allowing idealists and policymakers to ensure that the effects of the Web on society are largely beneficial. Meanwhile, O’Hara, Contractor, Hall, Hendler, & Shadbolt (2013) confirmed the expectation of Berners-Lee (2007), which expected that:

- The Web-based micro-actions, developed by individual actors, organizations and systems, would produce a growing set of Web resources (micro-resources) according to their own needs and available technologies.
- From the use and reuse of these micro-resources, on a large scale, new resources will emerge (macro-resources).
- The process of producing more and better resources at the macro level will occur through a cycle of reaction and innovation, characterized by the creative collaboration of the Web community itself.

Viewing the Web as a network of resources that has evolved over time, we can assess the challenges that are posed by Web Science, and in particular for Computer Science, based on the micro actions supported by local assumptions and requirements and the effect of scale of the actions performed by millions of users. In this sense, Dutton (2013) emphasized that this has a computational bias having a tendency to value the efforts done by engineers. According to O’Hara and Hall (2011), this tendency occurs due to the importance that must be given to the development of systems and tools that guarantee the expected sustainability at the macro level, since these may be extremely limited at the level of individuals actions (micro), when analyzed outside the scope of their creation.

The Semantic Web is a quite diverse set of technologies, analyzed and used in different ways by different communities and social groups. In recent years, Web Science has devoted considerable efforts for the development of the so-called Semantic Web (Berners-Lee, 2000; Berners-Lee, Hendler,
& Lassila, 2001; Berners-Lee et al., 2006), which does not replace the classic Web, as we know. On the contrary, it intends to make its own content (and everything that will come in the future) into suitable information for regular consumption, both by persons and by automated processing mechanisms.

Ontologies occupy a central position in the conceptual model of the Semantic Web. This is not by accident. For Horrocks (2002), ontologies are the basis of a technology capable of offering important contributions to achieve the goals proposed by the Semantic Web, mainly for providing an explicit semantic to the contents available on the Web, favoring its operation from sources with precisely defined terms. Ontologies have been object of study since the antiquity. Today, we can easily find many publications about them in different areas of knowledge. The work on which we support our arguments comes essentially from areas like Philosophy, Information Science and Computer Science. The different meanings of the term in these fields are vital to perceive the contributions ontologies can provide, as one of the major technologies of the Semantic Web.

In recent years, many concepts of ontologies have been presented and defended by different authors (Gruber, 1993; Guarino, 1995; Guarino & Giaretta, 1995; Noy & McGuinness, 2001; Guizzardi, 2005; Almeida, 2013). In most cases, the conceptual divergences that emerged are consequences of the specific context in which the definition was formulated. However, if we synthesize all the definitions presented for an ontology, we can conclude that in the context of Computer Science, an ontology describes in a formal, declarative and explicit way the concepts, properties and their relationships shared in a given domain. In general terms, ontologies are modeled to organize knowledge about a domain, facilitating its sharing and reuse. Reinforcing this latter aspect, several authors emphasized the importance of interconnecting ontologies in the context of the Semantic Web. For example, Allocca, D’Aquin, and Motta (2009) have already presented a network of ontologies organized in the form of a common ontology repository, managed by an ontology especially developed to provide a starting point for analyzing, representing, manipulating and reasoning about the adjacent structure of the network. On the other hand, Diaz, Motz, and Rohrer (2011) emphasized the potential of ontologies for integrating with other existing ontologies, favoring the modularization and reuse of their content, and D’Aquin (2012), reported that large ontologies are difficult to manipulate and maintain.

A NETWORK OF ONTOLOGIES FOR THE DOMAIN OF SINAES

The domain of Brazilian higher education, especially regarding to SINAES, through individual actors, institutions and systems, produces, uses and reuses an increasing amount of data. However, as we have already mentioned, all these data produced and published through individual actions are not necessarily adequate for general consumption at the macro level of Web resources. Knowing the diversity of interests that characterize higher education in Brazil, we are proposing the development of a network of ontologies with the ability to provide a basic conceptual approach to all the stakeholders of the domain.

The SINAES and the Evaluation of Brazilian IES

SINAES was established by the law no. 10 861 of April 14, 2004 (Brasil, 2004), which is known as the “SINAES Law”. After its implementation, some authors acknowledged that SINAES provided a new perspective for the process of evaluating the Brazilian higher education, which is now more formative and centered on the institution as a whole, recovering the more complex concept of higher education whose purpose is the formation of citizens with capacity for critical reflection, having the
Methodology for The Development of an Ontology Network on The Brazilian National System... 5

society as the central reference (Barreyro & Rothen, 2006; Dias Sobrinho, 2010). The SINAES Law defined an evaluation process supported by three distinct but complementary processes:

- The Evaluation of Higher Education Institutions (AVALIES).
- The Evaluation of Undergraduate Courses (ACG).
- The Assessment of Student Performance (ENADE).

The evaluation processes are oriented to provide quality indicators at different levels and contexts, being their results analyzed in a systematic way in order to give effective information about the performance of higher education and the institutions that support the higher education system. The SINAES law deals specifically with the evaluation process of higher education in Brazil and establishes that the evaluation will occur in two distinct and complementary moments: self-assessment, under the responsibility of a specific evaluation committee (CPA), and the on-site external evaluation, which is carried out by commissions of evaluators, under the responsibility of the National Institute of Studies and Educational Research Anísio Teixeira (INEP). The same law also defined a set of ten distinct dimensions that must be considered, both in terms of self-evaluation as in terms of external evaluation. The self-assessment should be coordinated by the CPA regimentally constituted within each higher education institution (HEI).

Once the work of an assessment cycle is completed, the CPA should send through the e-MEC system a self-assessment report, which will be incorporated posteriorly in the set of documents of the overall process of regulation and evaluation of higher education (Brasil, 2014b). In turn, the external evaluation should be carried out by the evaluation committees organized by the INEP, having as reference the quality standards for higher education expressed in the evaluation instruments that takes into account the diversity of the higher education system and respects the identity of the various institutions that integrate it, considering the specificities of the different academic organizations, based on the focus defined in the Institutional Development Plan (IDP) and in the processes of institutional evaluation (internal and external). The instrument is organized in five axes, contemplating the ten dimensions of SINAES (Brasil, 2004; Brasil, 2014a). Lastly, according to INEP, the evaluation processes should defined as a system that allows for the integration of the various dimensions of the case evaluated, ensuring a conceptual, epistemological and practical coherence, as well as the achievement of the objectives of the various instruments and modalities.

The Network of Ontologies OntoSINAES

In general, ontologies are not studied as isolated systems. Based on the good practices of the ontology engineering, any ontology would be at least able to be integrated in a network of ontologies. For Diaz et al. (2011), usually ontology networks are interconnections of existing ontologies that enable modularization, reusing and reengineering of knowledge. The OntoSINAES ontology is being developed with the aim to contribute for improving the quality of the data produced by the different actors, institutions and systems of the SINAES domain, facilitating its publication as open and connected data (Silva & Belo, 2017). Besides the actors, institutions and systems in the domain of SINAES, the users of OntoSINAES could be any person that has interest in the results of the Brazilian higher education evaluation.
In the review of the literature, we have identified several papers (Allocca et al., 2009; D’Aquín, 2012; Diaz et al., 2011; Suárez-Figueroa, 2010) that demonstrate the advantages of a network of ontologies - when compared to solutions based on large monolithic ontologies - for projects in which the development of reference ontologies is desired in domains with similar characteristics to SINAES. It is important to note that a network of ontologies is not a new paradigm for the engineering of ontologies. However, its constitution implies that each of the ontologies that integrate the network share both its content and metadata, which must contain information about its relations with other ontologies in the network.

From the beginning, we had to take into account the foundations of Web Science. According to Hall & Tiropanis (2012), this has already developed a community and a very diverse and interesting set of theories and methods, starting to gather scientific evidences and working for enabling those who needs to define and ground any policy for ensuring that the effects of the Web are largely beneficial on our society, as well as allows for preserving the foundations that guarantee the “well-being” of the Web itself. In Figure 1 we can see an illustration of how the data produced and published by different actors in the domain of SINAES arouse social interest, gain relevance and become used and reused on a large scale. In these multiple social interactions, the data of SINAES is used and reused in very distinctive ways, some of them in a very different way that the one predicted by those who produced the original data. In the midst of all that, at the macro level of the Web, new data, information, knowledge and technologies are emerging and enhanced by the cycles of reaction and innovation provided by the collaboration between domain experts, scientists and engineers.

Figure 1. Data produced at the micro level of the Web. (Adapted from Berners-Lee, 2007; O’Hara et al., 2013).

According to Berners-Lee (2007), it is in this aspect where “magic” of Web Science begins, that is, the multiplication of data and the collaboration among stakeholders, scientists and engineers, proposing solutions and contributing to increase the quality and improve the processes of sharing
information and knowledge. These are the so-called macro features, which are derived from the scale effect and collaboration provided by the Web. Thus, taking into account the Web Science dynamics and its implications in the domain of SINAES, the OntoSINAES project has been formally developed as a network of interconnected ontologies, implying that each ontology incorporated in the network share the various knowledge elements of the others ontologies. The work carried out by Allocca et al. (2009) reveals in a very concrete way some theoretical contributions that guarantee the logic and the consistency of the model of an ontology network. However, our proposal differs from the one presented by these authors, since we have the goal to build a network of ontologies taking advantage of “ontological and non-ontological knowledge” (Suárez-Figueroa, 2010), not simply integrating a group of ontologies that already exists. In order to formalize and document all the relevant relationships of OntoSINAES network, we had to deal with some pertinent issues, such as:

- The identification of high level relationships, such as inclusion, similarity, incompatibility or version.
- The specification of the identified relationships, verifying the most relevant aspects for their characterization.
- The characterization of each relation through its algebraic properties - for example, the previous inclusion can be defined as reflective and transitive.
- The establishment of links between relationships.

In Figure 2 we can see a fragment of the conceptual model of the OntoSINAES network of ontologies, in which it is possible to observe some of the relationships that we considered important to formalize the integration of the ontologies in the network.

\[\text{Figure 2. A fragment of the conceptual schema of the OntoSINAES ontology network.}\]
Once the fundamental characteristics of the ontology network are defined, we pass the definition of the methodology for the development of the ontologies in the OntoSINAES project.

**A METHODOLOGY FOR THE CONSTRUCTION OF THE ONTOSINAES ONTOLOGIES**

Similarly to software engineering, ontology engineering involves the use of a set of processes, activities, languages, tools and life cycles, usually articulated using a methodology. For choosing a methodology we must take into account the scope and the basic requirements established for the ontology to be developed. In addition to clarifying the objectives of the ontology to be developed and the identification of its main users, during the initial activities for requirement elicitation we defined some non-functional requirements. The first one was: “RNF-00: The environment for collaborative development of ontologies and vocabularies for the domain of SINAES (OntoSINAES) will be constituted by a network of ontologies formally interconnected”. Consequently, we have always taken into account the specificities of creating an ontology network as a core of knowledge that could be used and reused by all those interested in the results of SINAES.

The literature review we made revealed a number of papers oriented for analyzing and comparing the main methodologies that exist to guide the development of ontologies (Bortolato, 2014; Corcho, Fernández-López, & Gómez-Pérez, 2003; Silva, 2009; Suárez-Figueroa, 2010). According to (Corcho et al. 2003), there are many methodologies with different levels of dependence concerning their final application, varying between the ones that are totally dependent, partially dependent or totally independent, highlighting among the latter Methontology (López, Gómez-Pérez, Sierra, & Sierra, 1999) as the approach more mature at the moment. More recently, in (Bortolato, 2014) it was presented a comparative analysis between different methodologies, in which the author confronts the methodology for building ontology networks (NeOn) (Suárez-Figueroa, 2010) with other methodologies, concluding that the NeOn provides a construction strategy for ontologies independent from the application and from a core ontology, supplies consistent documentation about the methods, techniques and tools related to the activities of an ontology engineering process, and disposes the necessary support for the construction of a network of ontologies. Therefore, based on the results of the comparative studies, and considering that we intend to build an environment for collaborative development of ontologies in a domain with a national dimension, we chose the NeOn methodology (Suárez-Figueroa, 2010) to guide the continuity of the Project OntoSINAES and the Methontology (López et al., 1999) to provide a method for conceptualization, once the NeOn methodology leaves this activity in an open state for allowing the use of different methods.

NeOn is a scenario-based methodology that was designed specifically for the development of ontology networks. Some of its characteristics reinforce its alignment with the goals we set previously for the OntoSINAES project, namely:

- The NeOn methodology does not impose a rigid workflow. On the contrary. It suggests alternative ways for the development of the ontology we pretend.
- The NeOn framework provides a glossary of processes and activities that are usually involved with the development of ontologies – e.g., nine scenarios covering the most common ontology engineering situations, two development lifecycle models, and a set of guidelines of how to carry out different processes and activities (Bortolato, 2014).
The methodology has detailed guidelines about how the processes and activities it proposes can be carried out in accordance with the various principles of software engineering.

The NeOn methodology takes into consideration two other issues that we consider very important whenever an ontology network has as a goal to contribute for the sustainable development of the Web macro level. The first issue is related to the explicit reuse of ontological and non-ontological knowledge resources that represent some conceptual consensus already established within the domain beyond to the consequent relationships with other resources, many of them already published on the Web. The second issue is about the possibility of using and developing Ontology Design Patterns (ODP) – (Suárez-Figueroa, Gómez-Pérez, Motta, & Gangemi, 2012) provide us a set of guidelines for the development of standards-based ontologies. In our case, we consider that the creation of a repository of ontological patterns will contribute for accelerating the development of new ontologies for the domain of SINAES and improving the quality of the ontologies.

THE DEVELOPMENT OF THE ONTOSINAES ONTOLOGIES

The NeOn methodology provides two lifecycle models for the development of ontologies. The cascade model, in which we wait for the completion of a phase to start a new one, and the incremental iterative model in which the project is developed in successive iterations, each one organized as a cascade model with a limit set of requirements well defined.

For the work we are developing, the incremental iterative model is more attractive and convenient, since we are dealing with a project that involves multiple domains and besides this it is expected a collaborative and continuous development of ontologies for the domain of SINAES. Each one of the iterations will be a cycle of collaboration and evolution of ontologies and vocabularies, where participants will have at their disposal several functional prototypes that they can use, observe their evolution and participate in their improvement.

It is expected that some punctual, agile and successful collaborations may contribute to the consolidation of a culture of collaboration within the domain. It should also be noted that the incremental iterative model provides greater flexibility for the inclusion or modification of requirements. In Figure 3 we can see the schema of the life cycle we followed in the development of the first version of the network of the OntoSINAES ontology.

![Figure 3. The development life cycle of the OntoSINAES ontology network. (Adapted from Suárez-Figueroa, Gómez-Pérez, & Fernández-López, 2015).](image-url)
The incremental and iterative model has the ability to predict a general initialization phase of the project. At this stage a global development plan and a basic set of requirements for the ontology network are elaborated. All the iterations should be planned following one of the cascading models. In each iteration can be used a different version of the cascade model, according to the characteristics of the ontology that is intended to develop.

In addition to the typical iteration planning activities, during the initialization phase of the different iterations the global development plan and the general requirements should be reviewed and updated. The processes and activities involved with the execution of the first iteration were defined observing the various conceptual categories of the SINAES domain, which are grouped into subdomains representing distinct modules of the ontology in development. In Figure 4 we present the conceptual model of the SINAES domain.

![Figure 4. The conceptual model of the SINAES domain layer.](image)

The first ontologies created for the OntoSINAES network deal with essential concepts describing the organizations and the normative acts in the domain of SINAES and indicating the main relationships with other ontologies to be developed. According to (Silva & Belo, 2018), COrg is a core ontology that integrates the OntoSINAES ontology network, with the objective of providing a semantic model to formally represent the IES organizational structure, encompassing all the general concepts referenced in the legislation that regulates the SINAES.

For each ontology, there were created specific conceptual and formal models using the OntoUML language stereotypes (Guizzardi, 2005; Guizzardi & Wagner, 2005). The OntoUML language is a UML profile ontologically based on the Unified Foundational Ontology (UFO) proposed in (Guizzardi, 2005).

In all the formal models, concepts were stereotyped as classes, the relationships between concepts were defined through object properties, and the attributes of the instances were defined using data properties – in Figure 5 it is presented a small fragment of the conceptual model of the COrg ontology and its formal representation (Silva & Belo, 2018).
Finally, in the implementation phase, we chose to use Protégé (<http://protege.stanford.edu/>) for transforming the conceptual model into a computable artifact coded in OWL2 DL language. In Figure 6 we present the class hierarchy of the initial prototype of the COrg ontology.

Figure 5. A conceptual model fragment of the COrg ontology (Silva & Belo, 2018).

Figure 6. Classes of the COrg ontology (Silva & Belo, 2018).
CONCLUSIONS AND FUTURE WORK

In this article we presented the work we done in order to understand how the engineering of ontologies has been evolving in the last years, especially from the point where we noted an increase of the data published in the LOD format. In fact, this scenario was predicted earlier by Berners-Lee (2007), when he revealed that the production of Web micro resources would happen through personal initiatives, companies or groups, taking into account their particular objectives, i.e., they would not care about the state of the art but only with the fulfillment of their own needs.

Also in this work, Berners-Lee continues his reasoning by saying that macro features should be improved by thinking in a general user, being the collaboration between engineers and scientists quite important to promote the evolution of the state of the art, driven by relevance factors emerging from the Web scale effect. In our point of view, this is the state where today’s ontology engineering is. However, it is clear that any initiative that aims to achieve new contributions to the Web macro resource level must be well-founded and contextualized in a way that it not only solves problems but also present better alternatives for those who will continue to produce micro resources.

Regarding the inquiries about ontology networks, considering what has been presented so far, in our opinion, this is an unavoidable topic when one intends to use ontologies for environments where it is expected to produce or improve Web macro resources. To build a network of ontologies and an environment that contributes to the collaborative and disciplined evolution of the network, the NeOn methodology presents an added value to take into account (Suárez-Figueroa et al., 2012). Using a single methodology for all stages of the project will allow the documentation of the ontology development process to be used as a tutorial for future users of the environment, constituting a positive evaluation of the ontology a proof of concept for validation of the development environment.

In a near future, we intend to clarify some specific questions related to what has been done so far, especially some aspects about the expressiveness and foundations of ontologies, design and develop a core and an intermediate ontology, this last located between the top-level ontologies and the ontology of the domain, understand the most specific concepts of the SINAES domain, deploy in an iterative way the environment for the collaborative development of the OntoSINAES network ontologies, and finally, propose a set of good practices for the development of ontology networks in order to reach new contributes for the Web macro resources level. In our opinion, the emergent and unexpected results together with the impossibility for identifying users in a clear way, which occurs at the Web macro level, imposes new reflections and eventually the adoption of new paradigms for the development of ontologies on the context of this work.
REFERENCES


