ONTOLOGY-BASED ARTIFACTS FOR RESEARCH COLLABORATIVE NETWORKS

Cláudio de Lima[1]

Rogério Cid Bastos[2]

Gregório Varvakis[3]

Abstract: *The identification and understanding of patterns in Research Collaborative Networks has become increasingly important, however, the absence of a systemic vision about the process of codification, management and dissemination of knowledge implies a reduction in the capacity to institutionalize the actions of these networks. The purpose of this article was to analyze how ontology-based artifacts have been used to represent Research Collaborative Networks. For this, an integrative literature review was conducted. The analysis of this field shows that these artifacts can provide more assertive access to relevant information and knowledge. They have become increasingly important as the use of knowledge graphs and machine learning has grown. Context is a critical feature for the data that are implicit in the analysis to really be intelligence indicators.*

Keywords: *Research Collaborative Network*; *Ontology*; *Knowledge Representation*; *Knowledge Discovery*; *Social Network Analysis*.

Resumo: *A identificação e a compreensão de padrões das Redes Colaborativas de Pesquisa se tornaram cada vez mais importantes, entretanto, a ausência de uma visão sistêmica sobre o processo de codificação, gestão e disseminação do conhecimento implica na redução da capacidade de institucionalização das ações destas redes. Este artigo teve como objetivo analisar como os artefatos baseados em ontologia são utilizados para representar estas redes. Para tal, foi realizada uma revisão integrativa da literatura. A análise desse campo mostra que esses artefatos podem prover o acesso mais assertivo às informações e conhecimentos relevantes. Eles têm se tornado cada vez mais importantes à medida que o uso de grafos de conhecimento e aprendizagem de máquina tem crescido. O contexto é característica crítica para que os dados analisados sejam indicadores de inteligência.*

Palavras-chave: *Rede Colaborativa de Pesquisa; Ontologia; Representação do Conhecimento; Descoberta de Conhecimento; Análise de Rede Social.*

[1] PhD candidate of the Postgraduate Program in Knowledge Engineering and Management – Federal University of Santa Catarina (UFSC) – Florianópolis – Brazil. ORCID: https://orcid.org/0000-0002-4214-2547. e-mail: claudio.lima@ufsc.br

[2] Professor of the Postgraduate Program in Knowledge Engineering and Management – Federal University of Santa Catarina (UFSC) – Florianópolis – Brazil. ORCID: https://orcid.org/0000-0002-3775-1516. e-mail: rogerio.bastos@ufsc.br

[3] Professor of the Postgraduate Program in Knowledge Engineering and Management – Federal University of Santa Catarina (UFSC) – Florianópolis – Brazil. ORCID: https://orcid.org/0000-0003-2576-4835. e-mail: g.varvakis@ufsc.br

1 INTRODUCTION

The most developed societies live on information, and digital technologies keep them oxygenated. These societies are characterized as Knowledge Societies due to their large production and dissemination of data, information, and knowledge (Lima; Bastos & Bastos, 2020). Modern organizations use the power of data analysis to decide where to open a branch, how to increase investments, or when to launch a service according to their target customers (Lima & Bastos, 2019).

The typical organizational unit of science has changed from individuals to the establishment of Research Collaborative Networks (RCNs), in all areas of knowledge. Understanding and identifying operational patterns of networks becomes increasingly important for the formulation of science and technology policies (Haddad, Mena-Chalco & Sidone, 2017). However, the absence of records and of a systemic vision about the process of codification, management, and dissemination of knowledge implies a reduction in the capacity to institutionalize the actions of the RCNs.

The academic-scientific area is one of those that suffer from the lack of organization and structure of the data published on the Web. According to Hussain *et al*. (2020), the lack of semantic definition provides little data interoperability, accessibility, and information discovery from heterogeneous data sources. In accordance with Palacios-Callender and Roberts (2018), national science systems should include mechanisms to leverage the connectivity of their researchers abroad and ensure that the country is dynamically inserted internationally in the RCNs, without losing the scope of local social needs.

In this perspective, considering the exponential increase in the availability of data, information, and knowledge on the Web, ontologies have become popular for knowledge representation in Artificial Intelligence (AI), providing a methodology for the easier development of reusable and interoperable knowledge bases. According to Nazario, Dantas, and Todesco (2014), they are considered an important means of representing, formalizing, and sharing knowledge.

In this context, this work aims to analyze the characteristics of recent years on ontology-based artifacts for RCNs, to identify trends and thematic gaps in publications indexed in international databases. Based on this analysis, this study seeks to verify how ontology-based artifacts are used to represent RCNs.

The following sections are organized as follows: section 2 presents the background. Section 3 describes the methodology used. Section 4 presents the results found, discusses the approaches and trends. Finally, the final considerations are presented in section 5.

2 BACKGROUND

Collaborative networks result from diversity, from the internationalization that integrates people, knowledge, skills, and methodologies around a shared cognition (Leite & Pinho, 2016). A collaborative network is defined by Camarinha-Matos and Afsarmanesh (2008) as “an alliance made up of a variety of entities that are largely autonomous, geographically distributed, and heterogeneous, in terms of their operating environments, culture, values, and goals, but that collaborate to better achieve common or compatible goals, and whose interactions are achieved through computer networks”.

The RCNs allow the existence of a dynamic of relationships between researchers, institutions, and countries that develop an academic work or a project, whose joint publications represent one of the clearest ways to register collaborative relationships in the scientific world. According to Leite and Pinho (2016), the development of RCNs is essential to achieve excellence in scientific dynamics. One way to capture the activities and relationships among academics is through the use of ontologies.

From the perspective of Computer Science, the term Ontology was initially used in Artificial Intelligence and nowadays it is also used in Knowledge Engineering (KE), a discipline dedicated to knowledge modeling and the creation and insertion of knowledge systems in organizations (Nazario, Dantas & Todesco, 2014). In AI, an area that deals with reasoning about models of the world, it is used with the aim of describing world concepts and their relationships.

Ontology is defined by Studer *et al*. (1998) as "a formal and explicit specification of a shared conceptualization". Formal because it must be machine-readable; explicit because the types of concepts used and the constraints on its use are explicitly defined; conceptualization because it refers to an abstract model that identifies the relevant concepts of some phenomenon in the world; and shared because it captures the consensual knowledge accepted by a group.

3 METHODOLOGY

This study is based on an integrative literature review about ontology-based artifacts for representing RCNs. In an integrative review, researchers objectively criticize, summarize, and draw conclusions about a subject. This occurs through systematic research, categorization and thematic analysis of previous qualitative and quantitative research on the subject (Lobiondo-Wood & Haber, 2017). The searches occurred in the Scopus, Web of Science (WoS) and Scielo databases in May 2021. The database searches used the keywords presented in Table 1. The search was realized by topic, that is, the title, the abstract and the keywords of the records were analyzed.

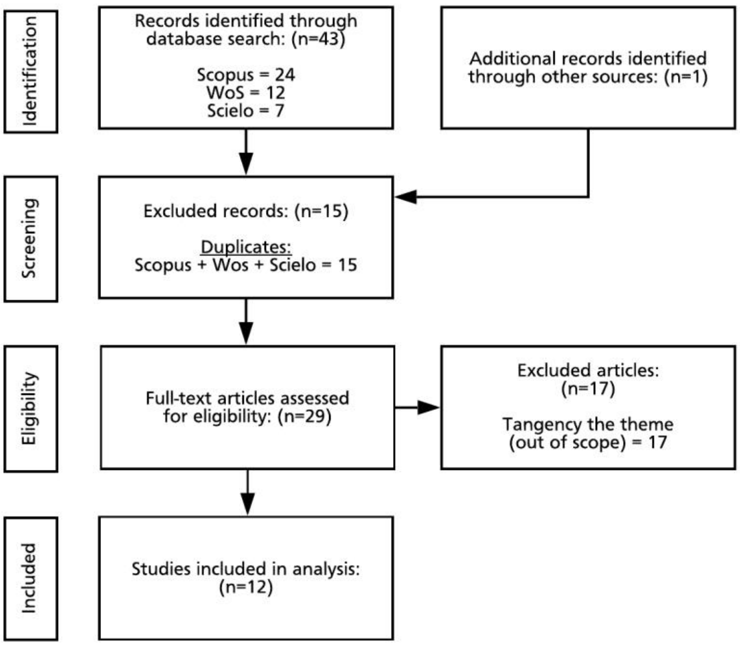
In the initial search 43 records were found in the databases. An article from a similar search has been added. Then, the articles were grouped by database and imported into the Mendeley software, where 15 duplicates were removed totaling 29 articles. Figure 1 presents the flowchart of the research process and article selection strategy. The abstracts of the articles were read and 17 studies were eliminated because they only tangency on the topic or were considered outside the scope of this review, remaining 12 articles for full reading, which represent the final sample of this work.

Table 1 – Search terms

Database	Keywords
Scopus	((collaborat* W/2 network*) W/3 (academic OR scientific OR research OR scholar* OR educat*)) AND (ontolog* OR "semantic web" OR (semantic W/3 technolog*))
WoS	((collaborat* NEAR/2 network*) NEAR/3 (academic OR scientific OR research OR scholar* OR educat*)) AND (ontolog* OR "semantic web" OR (semantic NEAR/3 technolog*))
Scielo	((collaborat* AND network*) AND (academic OR scientific OR research OR scholar* OR educat*)) AND (ontolog* OR "semantic web" OR (semantic AND technolog*))

Source: Elaborated by the authors.

Figure 1 – Flowchart of the research process



Source: Elaborated by the authors.

Previously, the articles were tabulated in a synthesis matrix (Garrard, 2016), using a spreadsheet to identify similarities based on the respective objectives, considering the artifacts applied in the context of RCNs. Finally, in order to group similar documents into clusters, the selected documents were analyzed by the Orange tool (Orange, 2021). For this task, document titles, abstracts and keywords were used. Cosine similarity was used to measure the similarity between the vectors, along with the Ward hierarchical clustering method, represented by a dendrogram.

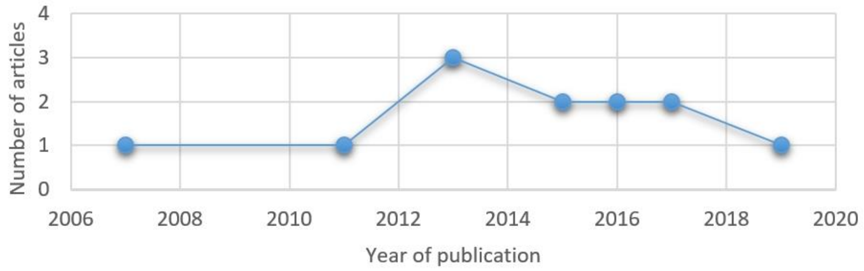
4 RESULTS AND DISCUSSION

This section presents and discusses the results found. The perspectives on ontology-based artifacts for RCNs are highlighted at the end of this section.

4.1 ANALYSIS OF RESULTS

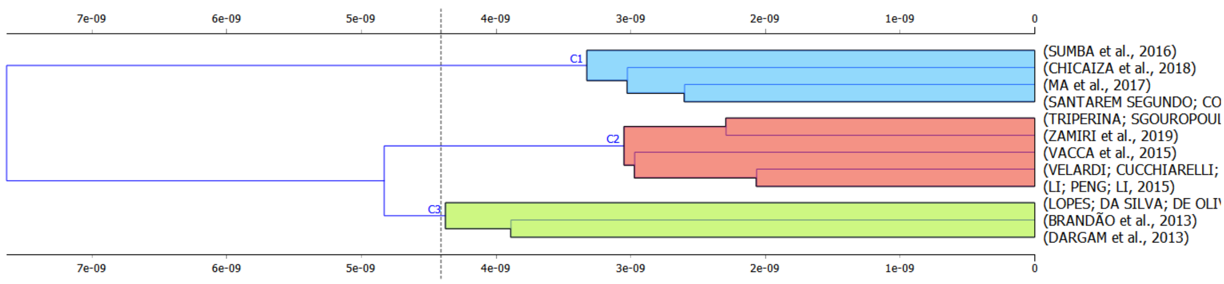
The selected publications took place between 2007 and 2019. About 60% of the sample was published in the last 6 years. Figure 2 shows the quantity of selected articles distributed by year of publication. Figure 3 shows the dendrogram resulting from the analysis of this grouping of documents. Considering the titles, abstracts, and keywords of the articles, with the support of the *Orange* tool, it was possible to obtain the word cloud of the selected publications, as highlighted in Figure 4.

Figure 2 – Selected articles by year of publication



Source: Elaborated by the authors.

Figure 3 – Dendrogram of selected documents



Source: Elaborated by the authors.

The most cited articles in the sample were written by Brandão *et al*. (2013), with 31 citations in *Scopus*, and Vacca *et al*. (2015), with 12 citations in *WoS* and 11 in *Scopus*. Table 3 shows the 3 clusters identified by the tool and the list of selected articles per cluster. The first cluster is formed by 4 studies based on Semantic Web concepts and technologies, with approaches dealing with semantic enrichment, using data mining, and also solutions that allow the management of researchers' academic and scientific information, such as the VIVO platform (Borner *et al*., 2012; VIVO, 2021). The second cluster presents 5 studies more focused on ontologies, with approaches dealing with recommendations, use of the VIVO platform or ontology, and also Knowledge Management (KM) aspects. Finally, the third cluster presents 3 studies that deal mainly with Social Network Analysis (SNA).

Figure 4 – Word cloud of selected documents

Source: Elaborated by the authors.

Table 3 – Clusters of selected documents

Cluster	Articles
01	(Sumba *et al.*, 2016; Ma *et al.*, 2017; Santarém Segundo, Coneglian & Lucas, 2017; Chicaiza *et al.*, 2018)
02	(Velardi, Cucchiarelli & Petit, 2007; Triperina, Sgouropoulou & Tsolakidis, 2013; Li, Peng & Li, 2015; Vacca *et al.*, 2015; Zamiri *et al.*, 2019)
03	(Lopes, Silva & Oliveira, 2011; Brandão *et al.*, 2013; Dargam *et al.*, 2013)

Source: Elaborated by the authors.

The next section analyzes the selected studies, considering their respective objectives, from the perspective of ontology-based artifacts for RCNs.

4.2 PREVIOUS CONSIDERATIONS

Based on the similarities found in the respective objectives and considering the ontology-based artifacts applied in RCN context, this section analyzes each category in more depth. Table 4 presents the studies by identified cluster.

Table 4 - Highlights of selected studies by identified cluster

#	Study highlights
Cluster 01	▪ In order to identify common research areas and potential RCNs, Sumba *et al*. (2016) proposed an architecture to join multiple bibliographic sources through a combination of ontologies, vocabularies, and Linked Data to enrich a database model. The architecture encompassed a process to extract, enrich, and represent bibliographic resources to discover patterns using Data Mining algorithms. A prototype was implemented to provide a centralized repository with bibliographic sources and find similar knowledge areas in the domain of the Ecuadorian research community. ▪ Chicaiza *et al*. (2018) presented a conceptual view for analyzing RCNs, in which researchers could use the model developed to identify critical areas of investment in research organizations, to find peers interested in common or related issues, or simply to learn about the topics researched by people from a particular sector. The model sought to facilitate interoperability and data sharing between different scientific applications or services. In addition, it allowed exploring the semantics of Linked Data and taking advantage of large collections of knowledge organized hierarchically by people and that are available in free format on the Web. The study presented as an application scenario the discovery of popular themes underlying the scientific production of some countries in the Andean region (Bolivia, Colombia, Ecuador, and Peru). ▪ In the context of the *Deep Carbon Observatory* (DCO), Ma *et al*. (2017) created a platform for the implementation of a knowledge network, the *Deep Carbon Virtual Observatory* (DCvO), aiming to promote collaboration among DCO participants, improve the openness and reproducibility of carbon-related research, facilitate accreditation for resource contributors, and potentially stimulate new ideas and discoveries in deep carbon-related studies. Using the network, underpinned by ontologies, DCvO is able to support various aspects of collaborative research in geoscience. The core of the DCO ontology is the VIVO ontology, which reuses and extends a list of ontologies to support information management in academia. A key feature of DCvO are the interconnections between various registered objects and resources, as well as the flexible ways to discover and access them. ▪ Finally, Semantic Web artifacts that contribute to building RCNs are discussed by Santarém Segundo, Coneglian, and Lucas (2017). To identify how semantic platforms behave in this task, the authors performed a proof of concept on the VIVO platform. To do so, they used a set of scientific publications from Brazilian journals. The tests showed a great adherence between Semantic Web technologies and the proposed organization performed in all interactions between the system and the user, even allowing the creation of other relationships.

#	Study highlights
Cluster 02	▪ In the context of the characteristics of scientific and technological achievements, Li, Peng and Li (2015) built an ontology that represents latent collaborative relationships and detected clusters of RCNs. The authors present a hierarchical recommendation structure that enriches domain ontologies and retrieves the most relevant information resources. They conducted a case study to collect a dataset of research achievements in the field of electric vehicles and obtained the best cluster results. ▪ In a project supported by the *University of Florida (UF) Clinical and Translational Science Institute* (CTSA), Vacca *et al*. (2015) used the VIVO platform to extract the social network of scientific collaborations on publications and awarded grants across all UF colleges and departments. Based on the notion of network interventions, the authors designed a change program to add specific links to the RCN. According to the authors, the results provided information on the feasibility of intervention programs in RCNs, as well as suggestions on the implementation of such programs to assemble interdisciplinary scientific teams in CTSA institutions. ▪ Triperina, Sgouropoulou, and Tsolakidis (2013) extended the VIVO ontology to cover teaching activities and connections of academics to facilitate work on RCNs in Higher Education Institutions (HEIs). Based on this ontology, the approach also presented services and tools to support quality assessment processes of academic organizations, in particular by visualizing RCNs. The solution falls within the perspective of KM handling in *Linked Open Data* environments. The authors also presented a new service developed and tested for a specific academic department within a Greek HEI. ▪ From the perspective of the European INTEROP NoE network, Velardi, Cucchiarelli, and Petit (2007) presented an algorithm to automatically build a thesaurus and an application in which taxonomic ordering between terms is explored to improve the diagnostics of a knowledge map (KMap), intended to facilitate and support synergy and collaboration in a distributed research community. The KMap draws a picture of the research status, in addition to keeping this image updated in the future and periodically carrying out a diagnostics of the extent of collaboration and coordination of the survey among INTEROP partners. From the stored data, a set of research competencies was extracted for each INTEROP partner and a similarity measure capable of expressing the degree to which competencies are common for pairs of partners was defined. ▪ Finally, Zamiri *et al*. (2019) discussed how the creation of an RCN can positively influence KM. The role of *Living Labs* and *Digital Innovation Hubs* in creating different fields of research and projects was addressed. One of the results of this approach was improvements in the ontology proposed in the study and its related knowledge. The authors highlighted a strategic partnership of the CARELINK Project with a *Living Lab*, which obtained, as a direct result of the initial meetings, in addition to the dissemination of the project's visions and objectives, an informative communication on the state of development of the *Open Internet of Things* throughout the zone of the European Union and in Portugal in particular.

#	Study highlights
Cluster 03	▪ Using SNA to recommend collaborations in RCNs, Brandão *et al*. (2013) proposed two new metrics to recommend new collaborations or the intensification of existing ones. Each metric considers a social principle that is relevant within the academic context, in order to check how they influence the resulting recommendations: the institutional affiliation aspect (homophily) and the geographic location information (proximity) of all researchers in the network. In the study, the authors focus on networks where the social ties were given by research links, specifically co-authorship social networks. ▪ The weights of the relationships among the actors in the social network aim to measure the importance of the relational ties among the actors and are also important to consider in an SNA. From this perspective, Lopes, Da Silva, and De Oliveira (2011) proposed the application of the Gini coefficient to the SNA. The Gini coefficient, a measure of statistical dispersion proposed in 1912 by Italian statistician Corrado Gini, is commonly used to assess the inequality of income and wealth distributions, but can be used for other distributions. The authors applied the coefficient to a co-authorship social network, in which the Gini coefficient measured the level of homogeneity of collaboration, that is, if only a few researchers maintained a good level of collaboration or if all researchers in the network were actually contributing to the group. ▪ Finally, Dargam *et al*. (2013) presented the EWG-DSS Collab-Net Project of the EURO Working Group on Decision Support Systems (DSS), which serves as a means for a social network perspective of research collaboration within the DSS community in Europe. With co-authorship as the main focus of the analysis, the network is designed to show the dynamics of collaboration among researchers, members of the EWG-DSS group. The authors presented the specification of the ontological model used within the research collaborative network, highlighting its benefits for the project. They also stated that the project will provide the DSS community in Europe with more accurate and up-to-date information about research projects and co-authorships, leading to better collaboration opportunities in the future.

Source: Elaborated by the authors.

4.3 PERSPECTIVES AND TRENDS

Highlighting some selected environments, within Table 5, this section analyzes how ontology-based artifacts are used to represent RCNs and discusses perspectives and trends in current environments.

Table 5 - Ontology-based artifacts for research collaborative networks

#	Perspectives and Trends
VIVO platform	▪ In academia, there is a growing need for technologies that allow the expansion of digital collaboration, making it more accessible to a larger number of researchers. In this context, it stands out that selected studies make use of the VIVO project (Triperina, Sgouropoulou & Tsolakidis, 2013; Vacca *et al*., 2015; Ma *et al*., 2017; Santarém Segundo, Coneglian & Lucas, 2017). The VIVO platform presents itself as an innovative environment and is based on the structure of the VIVO ontology, which was built using several other ontologies and vocabularies that makes it a tool understandable by several other tools. ▪ The VIVO ontology is used to represent the expertise of the people involved in the creation, transmission, and preservation of knowledge. It describes the activities and accomplishments of individuals in terms of their relationships to particular artifacts of work, resources they use, institutions that employ them, and other indicators. Independent of knowledge, the VIVO ontology supports the identification, evaluation, and impact assessment of individuals and groups of individuals, as well as the identification and reuse of individuals' work (LYRASIS, 2020). ▪ The VIVO Project is materialized through an open source application representing academic and research communities, developed with Semantic Web technologies such as RDF, OWL, and SPARQL (Bordin, 2015). With the advent of the Internet, languages for creating ontologies that explore the characteristics of the Web have been created. Recommended by the *World Wide Web Consortium* (W3C), OWL, one of these most widely used languages, is an essential element for the Semantic Web and its main form of materialization, the *Linked Open Data*. ▪ The main objective of the Semantic Web is to enable computers to understand the meaning of the data, according to the specific domains where they are inserted, providing the user with refined information, with a greater aggregate of knowledge (Berners-Lee, Hendler & Lassila, 2001). In this context, ontologies are the main elements that support inference generation logics, since they provide elements containing axioms. They are responsible for organizing the knowledge of the Semantic Web. ▪ In a successful example of applying the VIVO ontology, Ma *et al*. (2017) implemented a portal that establishes a knowledge network, with annotation and linking as key features, and supports various stages of an open scientific process within and outside the DCO community. Through the knowledge network, community members are able to add publications and datasets that may be useful to others, find colleagues working on similar projects, discover methods and tools that can be used to analyze data in new ways, and create more and better research collaborations. ▪ Finally, unlike many HEIs and research institutes around the world, Brazil does not yet have a production installation of the VIVO platform. However, according to Rathke and Rocha (2019), the VIVO ontology is also capable of representing a Brazilian HEI, providing a well-defined semantic framework for Web data integration that expedites the process of interdisciplinary and inter-institutional collaboration for the formation of RCNs.

#	Perspectives and Trends
Knowledge Discovery	▪ Ontologies can be used to leverage KM and enable better decisions (Dargam *et al*., 2013). Transforming research data into useful information and useful information into knowledge is a critical factor for success (Lima & Bastos, 2019). From this perspective, the knowledge discovery process, an important KE technique, transforms data into information by interpretation, derives new information from existing information by elaboration, and acquires new knowledge by learning. It is noteworthy that selected studies use techniques such as *Data Mining* (Li, Peng & Li, 2015; Sumba *et al*., 2016; Chicaiza *et al*., 2018) and SNA (Lopes, Silva & Oliveira, 2011; Brandão *et al*., 2013; Vacca *et al*., 2015). ▪ The *Knowledge Discovery in Databases* (KDD) process seeks to identify and uncover implicit relationships between information stored in databases (Fayyad, Piatetsky-Shapiro & Smyth, 1996). *Data Mining* is one of the steps in this process. Similarly, text mining is one of the steps in the Knowledge Discovered in Texts (KDT) process. *Clustering*, an unsupervised learning approach, corresponds to the task of dividing the data set into groups called *clusters*. In *Text Mining*, documents can be represented numerically as feature vectors. Thus, the similarity in the text of documents can be compared by measuring the distance between these vectors. ▪ The first attempts to analyze the RCNs focused on networks based on scientific publications (Grba & Mestrovic, 2018). RCNs are a special case of social networks in which nodes represent actors who collaborate on certain projects or scientific publications (Newman, 2001a). According to Brandão *et al*. (2013), co-authorship social networks are formed by researchers and their connections are given by collaborations in publications and patents. The use of SNA can lead to the identification of groups that are actively researching in a given area of knowledge and the influence of different groups within the researched community (Newman, 2001b). ▪ Finally, recommending collaborations is useful for increasing a group's connections, and then boosting the group's research as a collateral advantage (Brandão *et al*., 2013). Social recommendation, one of the applications of SNA, seeks to recommend or predict new links to help a researcher to form new groups or teams, to seek collaboration when writing a subsidy proposal, to improve the quality of communication in the network, and to investigate different research communities. According to Li, Peng, and Li (2015), expanding the scope of sharing information resources using the relationship between users became the main objective of this field.

Source: Elaborated by the authors.

5 CONCLUSION

The growing need for access and interoperability between scientific digital repositories is a reality. As well as Research Collaborative Networks need access to increasingly diversified data collections with increasingly heterogeneous data sources, researchers also need to access and use the data repositories of other Research Collaborative Networks. The use of Semantic Web artifacts to organize scientific production, presenting and explaining relationships and existing productions, can help researchers and the community in general with more assertive access to relevant information and knowledge.

Ontologies are an adequate solution to the lack of interoperability between platforms that store and publish scientific data on the Web, since they play a central role in the representation, modularization, and distribution of knowledge. They have become increasingly important as the use of knowledge graphs, machine learning, natural language processing, and the volume of data generated daily has increased exponentially.

The merging of the areas of data mining and knowledge representation through the use of ontologies represents a great opportunity to face the challenge of dealing with the large volume of data, information, and knowledge generated in the academic environment. In this context, it is highlighted that an intelligent system means contributing to this problem of data, information and knowledge overload, creating an environment for sharing information and knowledge, in addition to maximizing the implementation of Knowledge Management and innovation in services.

Including semantics to the Research Collaborative Networks representation allows additional characteristics to be analyzed. Social Network Analysis can provide insight into social influences within Research Collaborative Networks, although the analysis itself focuses on connectivity, on how individuals collaborate. Many challenges are faced when trying to perform meaningful analysis on Research Collaborative Networks for decision support purposes. In current systems, analyzes seem to be performed independently of the issue of data context, a critical characteristic to ensure that the data that are implicit in the analysis are really indicators of the intelligence that is sought.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American Magazine*, 284.

Bordin, A. S. (2015). Framework Baseado em Conhecimento para Análise de Rede de Colaboração Científica. Tese de doutorado, Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, SC, Brazil.

Borner, K., Conlon, M., Corson-Rikert, J., & Ding, Y. (2012). Vivo a semantic approach to scholarly networking and discovery. 2, 1–180. doi.org/10.2200/S00428ED1V01Y201207WEB002

*Brandão, M. A., Moro, M. M., Lopes, G. R., & Oliveira, J. P. M. (2013). Using link semantics to recommend collaborations in academic social networks. WWW 2013 Companion - Proceedings of the 22nd International Conference on World Wide Web, 833–840. doi.org/10.1145/2487788.2488058

Camarinha-Matos, L., & Afsarmanesh, H. (2008). On reference models for collaborative networked organizations. *International Journal of Production Research*, 46, 2453–2469. doi.org/10.1080/00207540701737666

*Chicaiza, J., Piedra, N., Lopez-Vargas, J., & Tovar-Caro, E. (2018). Discovery of potential collaboration networks from open knowledge sources. IEEE Global Engineering Education Conference, EDUCON, 2018-April, 1320–1325. doi.org/10.1109/EDUCON.2018.8363382

*Dargam, F., Linden, I., Liu, S., Ribeiro, R. A., & Zaraté, P. (2013). The development roadmap of the EWG-DSS collab-net project: A social network perspective of DSS research collaboration in Europe. *Lecture Notes in Business Information Processing*, 164, 1–18. doi.org/10.1007/978-3-642-41077-2_1

Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 17(3).

Garrard, J. (2016). *Health Sciences Literature Review Made Easy: The Matrix Method*. (5a. ed). Jones & Bartlett Learning.

Grba, B., & Mestrovic, A. (2018). Tracking the evolution of scientific collaboration networks. 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings (pp. 503–508). Institute of Electrical and Electronics Engineers Inc. doi.org/10.23919/MIPRO.2018.8400096

Haddad, E. A., Mena-Chalco, J. P., & Sidone, O. J. G. (2017). Scholarly Collaboration in Regional Science in Developing Countries: The Case of the Brazilian REAL Network. *International Regional Science Review*, 40(5), 500–529. doi.org/10.1177/0160017615614898

Hussain, N., Wang, H. H., Buckingham, C. D., & Zhang, X. (2020). Software Agent-Centric Semantic Social Network for Cyber-Physical Interaction and Collaboration. International Journal of Software Engineering and Knowledge Engineering, 30(6), 859–893. doi.org/10.1142/S0218194020400100

Leite, D., & Pinho, I. (2016). *Evaluating collaboration networks in higher education research: Drivers of excellence*. Springer International Publishing.

Lima, C., & Bastos, R. (2019). Knowledge Creation supported by Data Governance. Proceedings of the 9th International Conference on Knowledge and Innovation, CiKi 2019. Porto Alegre, Brazil.

Lima, C., Bastos, R., & Bastos, L. (2020). The Information Quality impact on Learning Platforms. *RENOTE*, 18, 59–68. doi.org/10.22456/1679-1916.110204

*Li, X., Peng, J., & Li, S. (2015). Achievements Recommendation Framework based on Scientific Collaboration Network. 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 929–934. doi.org/10.1109/ICMLA.2015.182

Lobiondo-Wood, G., & Haber, J. (2017). *Nursing research: methods and critical appraisal for evidence-based practice*. (9th ed.). Mosby.

*Lopes, G. R., Da Silva, R., & De Oliveira, J. P. M. (2011). Applying Gini coefficient to quantify scientific collaboration in researchers network. ACM International Conference Proceeding Series. doi.org/10.1145/1988688.1988767

LYRASIS. (2020). VIVO Ontology Domain Definition. Retrieved on April 12, 2021 from wiki.lyrasis.org/display/VIVODOC111x/VIVO+Ontology+Domain+Definition

*Ma, X., West, P., Zednik, S., Erickson, J., Eleish, A., Chen, Y., Wang, H., Zhong, H., & Fox, P. (2017). Weaving a knowledge network for deep carbon science. *Frontiers in Earth Science*, 5. doi.org/10.3389/feart.2017.00036

Nazário, D., Dantas, M., & Todesco, J. L. (2014). Knowledge Engineering: Survey of Methodologies, Techniques and Tools. *Latin America Transactions, IEEE (IEEE Latin America Magazine)*, 12, 1553–1559. doi.org/10.1109/TLA.2014.7014527

Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Phys. Rev*. E, 64(1), 16131. doi.org/10.1103/PhysRevE.64.016131

Newman, M. E. J. (2001b). The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences of the United States of America, 98, 404–409. doi.org/10.1073/pnas.021544898

Orange. (2021). Data Mining Fruitful and Fun: Open source machine learning and data visualization. Version 3.28. Retrieved on May 20, 2021 from orangedatamining.com

Palacios-Callender, M., & Roberts, S. A. (2018). Scientific collaboration of Cuban researchers working in Europe: understanding relations between origin and destination countries. *Scientometrics*, 117(2), 745–769. doi.org/10.1007/s11192-018-2888-2

Rathke, S. B., & Rocha, R. P. da. (2019). Sistema de informação de pesquisa: Uso da ontologia de VIVO no contexto das instituições brasileiras. *Brazilian Journal of Information Science: Research Trends*, 13(4), 132–151. doi.org/10.36311/1981-1640.2019.v13n4.08.p132

*Santarém Segundo, J. E., Coneglian, C. S., & Lucas, E. R. O. (2017). Conceitos e tecnologias da Web semântica no contexto da colaboração acadêmico-científica: um estudo da plataforma Vivo. *Transinformação*, 29, 297–309. doi.org/10.1590/2318-08892017000300007

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge engineering: principles and methods. *Data & knowledge engineering*, 25(1), 161-197.

*Sumba, X., Sumba, F., Tello, A., Baculima, F., Espinoza, M., & Saquicela, V. (2016). Detecting Similar Areas of Knowledge Using Semantic and Data Mining Technologies. *Electronic Notes in Theoretical Computer Science*, 329, 149–167. doi.org/10.1016/j.entcs.2016.12.009

*Triperina, E., Sgouropoulou, C., & Tsolakidis, A. (2013). AcademIS: An ontology for representing academic activity and collaborations within HEIs. ACM International Conference Proceeding Series, 264–271. doi.org/10.1145/2491845.2491884

*Vacca, R., Mccarty, C., Conlon, M., & Nelson, D. R. (2015). Designing a CTSA-Based Social Network Intervention to Foster Cross-Disciplinary Team Science. *Clinical and Translational Science*, 8(4), 281–289. doi.org/10.1111/cts.12267

*Velardi, P., Cucchiarelli, A., & Petit, M. (2007). Supporting scientific collaboration in a network of excellence through a semantically indexed knowledge map. In Doumeingts, G and Muller, J and Morel, G and Vallespir, B (Ed.), Enterprise Interoperability: New Challenges and Approaches (P. 231+). Springer-Verlag London Ltd. doi.org/10.1007/978-1-84628-714-5_22

VIVO. (2021). VIVO connect share discover. Retrieved on March 05, 2021 from duraspace.org/vivo/

*Zamiri, M., Marcelino-Jesus, E., Calado, J., Sarraipa, J., & Goncalves, R. J. (2019). Knowledge Management in Research Collaboration Networks. In L. M. Zheng F. Chu F. (Ed.), Proceedings of the 2019 International Conference on Industrial Engineering and Systems Management, IESM 2019. Institute of Electrical and Electronics Engineers Inc. doi.org/10.1109/IESM45758.2019.8948162