Journal of Integrated Information Management electronic journal

Developing the infrastructure for managing institutional research archives

Anastasios Tsolakidis[1], Evangelia Triperina[1], Ioannis Triantafyllou[2], Christos Skourlas[1]

[1]University of West Attica Department of Informatics and Computer Engineering
[2]University of West Attica Department of Archival, Library and Information Studies
atsolakid@uniwa.gr [**ORCID**: 0000-0001-7364-4542], evatrip@uniwa.gr [**ORCID**: 0000-0003-4282-2259], triantafi@uniwa.gr [**ORCID**: 0000-0001-5273-0855], cskourlas@uniwa.gr [**ORCID**: 0000-0003-4464-5305]

Article Info

Article history:
Received April 21 2023
Received in revised form May 12 2023
Accepted June 01 2023

http://dx.doi.org/10.26265/jiim.v8i1.4515

Abstract:

Purpose *- Capturing research in modern academic institutions is of vital importance. A robust approach to systematically harvest, host, and update a university-wide repository is not only beneficial for the academics within the HEI, but it can also boost and promote the visibility of the University.*

Design/methodology/approach *- Although many decentralised approaches offer the possibility to showcase an individual's research, the vast and disparate sources cannot facilitate an institutional repository. Since research constitutes one of the main aims of a university, we follow a solution that records the whole research that occurs within an institution, which is VIVO. In this paper, we present the adaptation of VIVO by a Greek University, namely, the University of West Attica. In our approach, we have merged the existing systems and harvested research-related information from different sources on the web.*

Findings *- The proposed solution serves as an academic research repository. It provides consistency over the presented research outputs, making the research of the University wider visible.*

Originality/value *- The paper presents an ontology-based system for documenting the undertaken research within an institution. Moreover, it allows the personalisation of the ontology, thus making it possible to customise the repository to fit an institution's needs, the web interface, therefore presenting the perceived significant components for the University, and the information visualisation.*

Index Terms — institutional repository, linked open data, ontology, research archives, research management, semantic web.

I. Introduction

Nowadays, the highly competitive academic environment dictates the need to record, present and assess the activities and collaborations that take place in universities. Since academic institutions have manifold roles involving research, education, innovation, and industry collaborations, this information must be systematically captured in an institutional repository. Moreover, it is vital to examine the links between research and education and to inspect how they affect each other and to what degree. The conducted research, its outcomes and the involved people, the educational activities, collaboration and their results, and the collaboration within academia and with organisations correspond to valuable information for an institution and can contribute to quality assurance, strategic planning, dissemination of educational and research outputs and deduction of meaningful results. Apart from facilitating the various academic and managerial processes, this information can also enable academic networking. We present a system that aggregates research and education records and services. Our system aims to offer a solution for managing and manipulating academic information for our institution's and its departments' evaluation and quality assurance.

Our approach builds upon the VIVO ontology and introduces an elaborate research management information system called IREMA (Institutional REsearch MAnagement) **[1]**. Our system uses the sharable and reusable linked data produced by the VIVO semantic web application and implements web services. As a direct consequence, our solution is both sustainable and extendable. Within IREMA we have implemented and profiled VIVO for our academic institution, the University of West Attica, Greece. We have adopted the VIVO-ISF ontology and extended it to meet our University's requirements, which led to the creation of the AcademIS ontology **[2]**. All the information of our system is stored in the VIVO instance, and the web services are built upon VIVO to access the data and communicate the information to the other employed systems. VIVO also covers the academic networking requirements of the institution, while IREMA retrieves the institutional data stored in VIVO by executing SPARQL queries in the instance of the semantic web application, the outputs of which are acquired in .json format. IREMA then reuses the information

of VIVO, applies metrics to the data and visualises the results.

Moreover, based on the data retrieved by the before-mentioned installation, we implement a decision support mechanism aided by visualising the data. IREMA provides efficiency measure techniques, data mining methods (cluster analysis, association rules and Bayesian networks), and social network analysis (including community detection, identification of research hubs and discovery of the most important researchers based on specific characteristics). The motivation for developing the IREMA system and linking it to our VIVO instance is integrating the academic services in our institution to satisfy the various and diverse needs that exist in the academic setting by a single integrated information system. Therefore, it is possible to explore all the activities and collaborations in an academic institution and have insights about future collaborations. In this paper, there is a thorough analysis of the way we use the semantic application VIVO and its importance for our approach. Furthermore, we present the exploitation and the extension of the underlying VIVO-ISF ontology, which assist us in accumulating academic information, structuring and analysing it. Our system supports the further presentation of the data, including visualisations and visual analytics, IREMA **[1]**, **[3]**, which is also outlined.

The paper is structured as follows: the first section is the introduction, whereas the second section corresponds to the background information. Section 3 outlines the basic concepts of our research management infrastructure and the presentation of VIVO adoption in our academic institution, including the implementation of VIVO-ISF ontology and its extension. The fourth section demonstrates how the proposed approach can be used to guide academic decisions. Finally, in section 5, there is the conclusion and the future work.

II. Background

A. Capturing research

Universities have become a major source of novel knowledge and research due to their involvement in research activities **[4]**. Universities and research institutions must capture and disseminate research and research outputs, which are needed from various user groups, including scientists, policymakers, policy researchers, industry, and media **[5]** and should keep up with the emerging requirements **[6]**. Academic institutions utilise information systems dedicated to research to depict the research activities. A research information system is "any informational tool dedicated to providing access to and disseminating research information" **[7]**. A research management system is a subcategory of a research information system, which supports workflow for the entire research procedure **[8]**, and IREMA constitutes such a system.

B. Research Information Systems and data models

More and more academic institutions have made use of Research Information Management Systems (RIMS) and Current Research Information Systems (CRIS) **[9-13]**. According to the literature, CRISs have been implemented in many universities autonomously or integrated **[14]** with other web services. CRISs have been enhanced with Linked data and semantic web technologies **[15], [16],** blockchain technology **[17]**, visualisations **[15]** and other technologies to better support the hosting and the presentation of the conducted research. The imperative need for CRIS in the academic setting is underlined by creating and disseminating intelligent systems for training stakeholders to use the research information systems effectively and efficiently **[18]**. Other approaches have examined RIS over their ergonomic evaluation **[19]**. CRIS systems can be categorised as institutional, regional, national, or international.

Among the most used data models for capturing information about the conducted research is the CERIF (Common European Research Information Format), which is a standard for managing and exchanging research data **[20]**, **[21], [22]**. VIVO is an ontology that offers a unified, formal, and explicit specification, which depicts data about researchers, institutions, activities, outputs, and related relationships [23].

III. Research Management Infrastructure And Institutional Repository

The research management infrastructure includes the facilitation of academic records, and the analysis of the research information employing visualisations and their interconnections. The University's academic records are stored in VIVO, which serves as an academic repository, while the IREMA system analyses the involved information. The infrastructure includes functionalities such as academic archiving and networking, quality assurance, visualisations of the data based on specific metrics and decision support aided by visualisations. It also provides efficiency measure techniques, data mining methods and social network analysis. To further process the VIVO instance's data, we have developed a connection between VIVO and IREMA based on VIVO-ISF ontology. It is also compatible with the ontology we developed, the AcademIS ontology **[8], [9]**, which is based on the VIVO-ISF ontology. All the components and functionalities of our system (**Fig. 1**) will be thoroughly described in the subsequent sections, as well as how the components are linked to each other, how they interact, and how the institutional data from VIVO are transferred to IREMA.

A. VIVO

VIVO **[24, 25]** is the backbone of our system since it is the basis of the IREMA web services and facilitates academic networking needs. Within our VIVO instance, we store information related to research, education, and the quality of both research and education. Furthermore, we explore and accumulate information about how education and research interrelate within an academic institution. To accumulate all the previously mentioned academic

information, we have developed the AcademIS ontology.

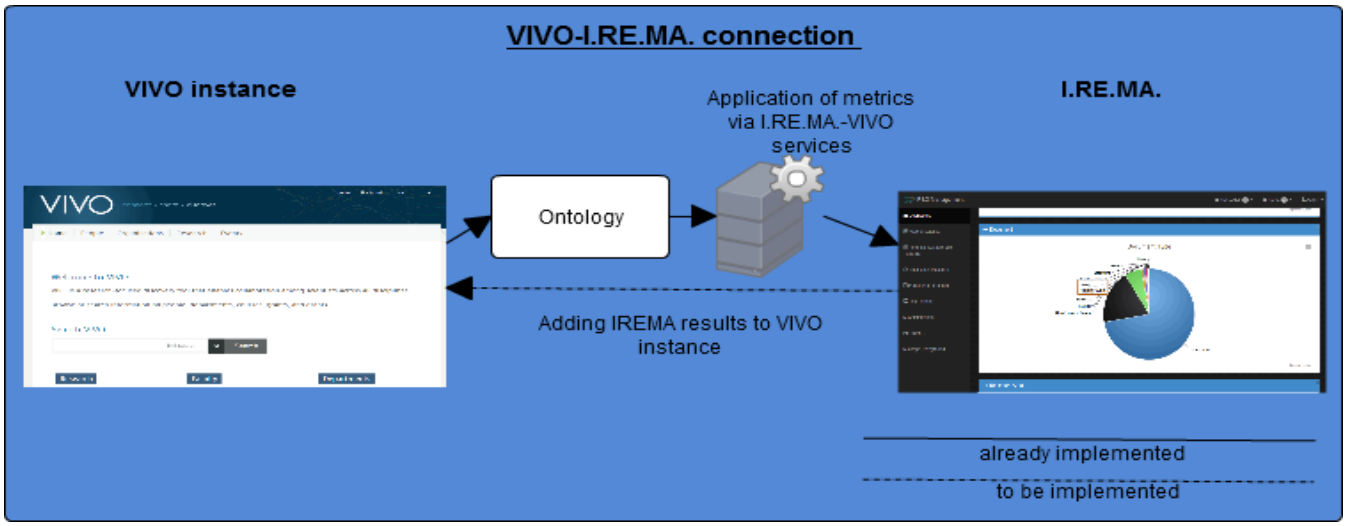

Fig. 1. *The VIVO-IREMA research management infrastructure*

In this section, we will refer to VIVO-ISF ontology and present thoroughly the AcademIS extension and the VIVO instance of the University of West Attica.

VIVO-ISF ontology models information about research and education in HEIs. However, our system needed more details about education, the points in which education and research correlate and the quality management of academic endeavor. Consequently, we have used the VIVO-ISF ontology, which covers most of our requirements and has extended it to comply with our additional needs. To maintain the interoperability of the VIVO ontology **[10]**, we designed the additions to be as general as possible so that the AcademIS could be applicable to a broad range of educational institutions.

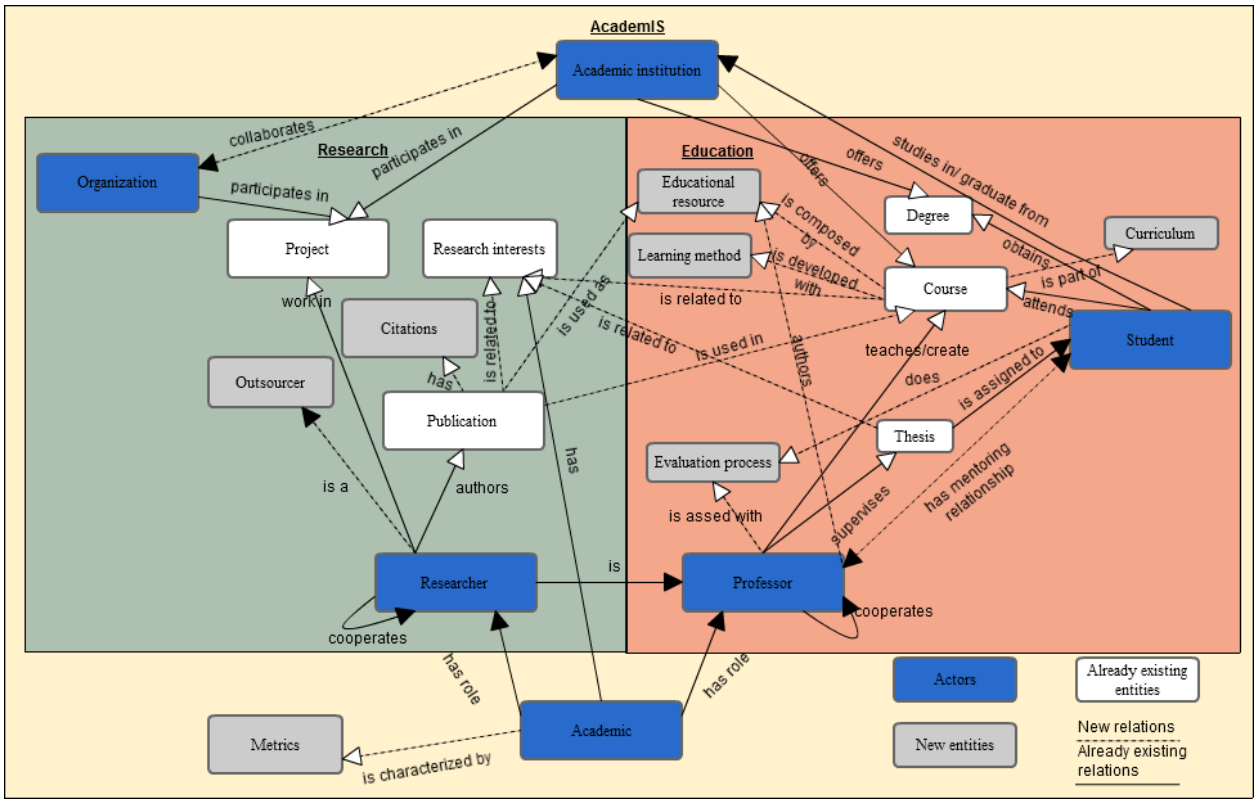

Fig. 2. *AcademIS ontology*

The AcademIS ontology is presented in **Fig. 2**. Our purpose for developing the AcademIS ontology is to provide a knowledge model of the academic domain. The AcademIS (Academic Information System) ontology is used as a basis for facilitating academic records management in our VIVO instance and the IREMA system.

The ontology has been implemented in OWL 2, in Protégé and then added in the VIVO University of West Attica instance. The domain of AcademIS is the academic activities and collaborations in HEIs. It encapsulates the research and education of an academic institution. The intended end users are the faculty, including the professors and the researchers, the policymakers of academic institutions, the quality assurance unit, and the VIVO and IREMA users of the HEI. The main indented uses of our ontology are the facilitation of research and education management, the quality management of an academic institution and academic data analysis via visualisations.

The AcademIS ontology introduces several new terms and relationships regarding academic institutions, the most important of which will be presented in this section. Regarding research quality, AcademIS ontology only adds information about citations of publications. AcademIS includes information about the educational process to measure the quality of education. The outsourcer entity regarding the research has been added, corresponding to the external research collaborator. In the context of education, we have appended the learning methods and the educational resources. The AcademIS ontology also measures how the research results and resources are used in or inspire the creation of educational resources and how the educational resources and outputs (such as thesis, etc.) are used in or inspire further development in research. The quality management aspects of the AcademIS are citation count, evaluation, etc.

B. VIVO instance – the case study of UNIWA

Initially, we implemented an instance of the VIVO for the Department of Informatics of the University of West Attica, and then we extended it at a larger scale. It now covers the whole academic institution. The VIVO instance of our academic institution includes research information, such as publications and research projects, as well as educational information, like courses. It uses the additional fields of AcademIS ontology by capturing learning outcomes, learning methods, educational resources, information about the academic institution's services and information about the interaction of education and research. It also provides quality management information.

The information was aggregated to our research repository from various sources, including institutional data (university records about research and education that happens in the premises of the University of West Attica in relational databases and other digital formats, etc.) and data from large online publications databases (Scopus, dblp). The VIVO instance was filled by employing the harvesting and ingestion options offered by VIVO. The data has been incorporated into the system, allowing users to edit, delete, or add information. To elaborate, the individuals curated the data to avoid invalid information and duplicate records to ensure the data's quality and integrity. Using the VIVO repository also enables the unification of all the academic information, increases its availability, and makes it available in Linked Data format **[11]**.

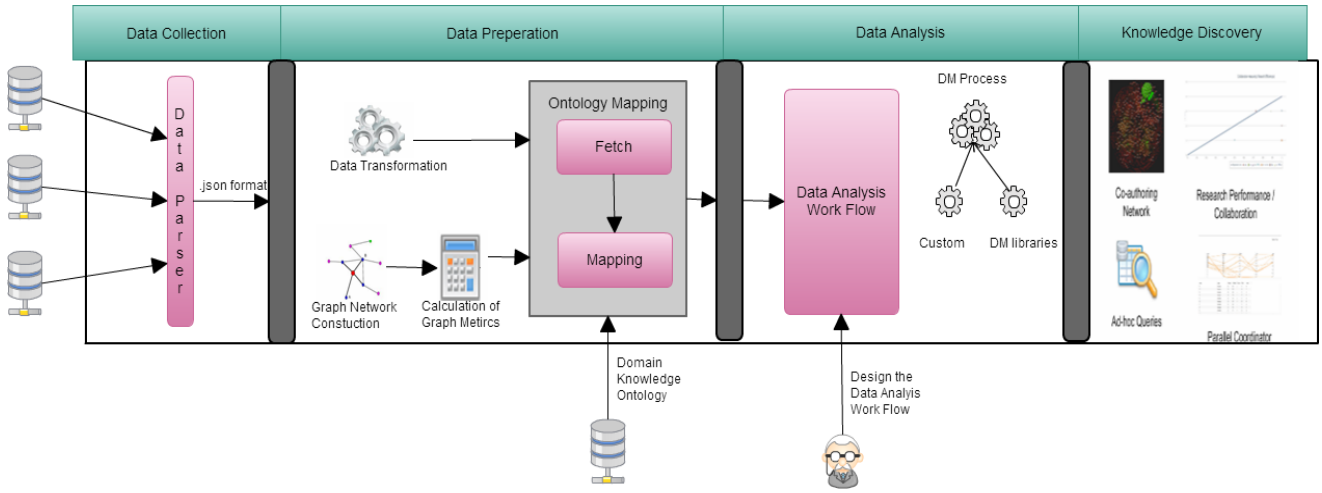

Fig. 3. *IREMA architecture*

C. IREMA - Institutional Research Management

IREMA stands for Institutional Research Management. The IREMA system collects data from multiple academic sources, mainly from the VIVO instance, followed by data transformation and ontology mapping using the domain ontology to prepare the data to be inputted into the system. The graph metrics are also calculated in that stage. The consequent step is the data analysis, while the final phase of the process is the knowledge discovery. To elaborate, the knowledge discovery is performed either with visualisation (co-authoring network, parallel coordinator, etc.), or with ad-hoc queries. In **Fig. 3**, we present the architecture of IREMA.

Apart from the process of our institutional data, we have implemented this process, which is applicable to other types of data that are available on the Internet (**Fig. 4**). For instance, it is possible to perform our interactive analysis on other VIVO instances, or even to use IREMA for social networks' analysis. It is also feasible to employ the IREMA

system in other domains by simply adjusting the ontology with the assistance of domain experts. Thus, IREMA can be utilised to carry out, for instance, environmental or financial analysis, depending on the input data and the underlying domain ontology.

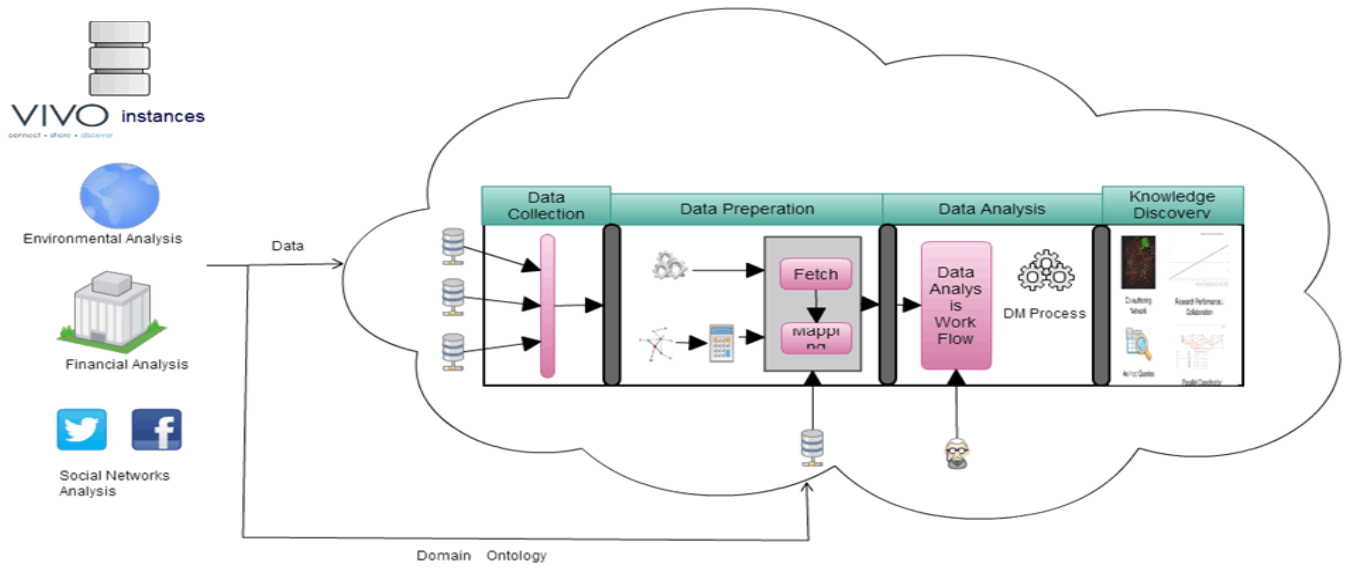

Fig. 4. *IREMA as a web service*

IREMA is a system that can be used to explore the co-authoring networks formed within an academic institution and gain insights into important scientific information. This system allows users to make better and more intuitive decisions by exploiting the information already in the VIVO repository. However, by applying metrics to the data, the information becomes more apparent and with the visualisations of the data, it is more efficiently perceived. Even large datasets can be explored easier and more interactively.

D. Features of the system

VIVO facilitated several functionalities, including academic networking and an institutional networking tool. Moreover, it has been employed during the data aggregation stage and as a management tool for academic information. In addition, our VIVO endpoint has also been employed to parse the academic information to the cooperating system, the IREMA. In this section, we will describe the architecture of our application, which consists of two main components, the VIVO-IREMA component, which includes all the data analysis methods and algorithms. All those methods are published using web services, and the clients consume (call) web services from the VIVO-IREMA. The layers of the VIVO-IREMA are the following:

i. Data collection and data cleansing: We use data from different data sources, for example, data from VIVO SPARQL endpoints or online publication databases, etc.
ii. Data transformation: During that process, a graph containing the academic researchers, based on the co-authorship of research papers, is created. In this stage, we calculate five (5) graph measures: degree centrality, closeness centrality, betweenness centrality, eigenvector and clustering co-efficient.
iii. Data Mining (DM): is used to extract hidden predictive information. The DM method falls into the categories of clustering, classification, and association analysis.
iv. Knowledge discovery via data visualisations: The proposed framework integrates interactive visual interfaces to support Knowledge Discovery (KD), thus providing the user with enhanced assistance throughout the decision-making (DM) process.

E. Presentation and visualisations of the data

The proposed framework supports the following visual representation techniques:

i. Co-authoring Graph is created based on the collaboration of faculty members in research papers.
ii. Efficiency Line is used to represent the correlation among the indicators.
iii. Parallel Coordinators is an interactive representation where the users can dynamically apply a set of criteria (dynamically) depending on their objectives.
iv. Map of Science, where each research area is represented in pie charts.

The visualisations are built based on specific metrics used to measure the research efficiency of academics and give insights into the trends regarding research activities and collaborations in the academic setting. In our application, we have implemented a graph analysis interface which calls web services from the VIVO-IREMA framework and displays the data in real-time using interactive graph networks. Moreover, we can make valuable deductions about the creation of research communities, the operation of

researchers as research hubs, the speed of the dissemination of information from an author to all the other authors in succession, the significance of an author in a network and the likelihood of an author and their connected authors to be a group. To examine the impact of collaboration patterns on the research activities of the faculty, we use the following network metrics. The graph density measures the connectivity among authors and their ability to collaborate. The average distance is the shortest path length between two connected nodes. The degree centrality is the number of arcs at each node and measures the 'activity' of the node. The betweenness centrality measures the ability of a node to connect nodes that do not have any direct connection. These nodes are called hubs because they can transfer information from one researcher to another.

The closeness centrality is defined as the inverted sum of the shortest distances between each node and every other node. It is interpreted as the ability to access information through the "grapevine" of network members. The eccentricity centrality of a node k is the largest geodesic distance from every other vertex. Therefore, it reflects how far each node is from every other node at most. The eigenvector centrality measures the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections with high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. The clustering coefficient quantifies how close the node and its neighbours are to be a clique. It determines whether a network is a small-world and calculates and displays all nodes' clustering coefficients. The HITS calculates the "hubs-and-authorities" importance measures for each node. These measures are defined recursively as follows: the hub is the degree to which a node links to other important authorities, and authority is the degree to which important hubs point to a node.

F. Connecting VIVO & IREMA

IREMA is a web service that receives academic data from VIVO installations or other sources and returns interactive visualisations and efficiency metrics. For the IREMA process to occur, the VIVO data must be imported or connected to IREMA. The connection can be described as follows: the data from VIVO are processed through IREMA with the aid of the ontology. Then metrics are applied to the data. Finally, the interface of IREMA presents the data in the form of visualisations. The clients can interact with the data collection layer to i) send the appropriate data, ii) call specific data analysis/preparation services, and iii) call specific data visualisations.

IV. Guiding Academic Decisions

IREMA – VIVO infrastructure is utilised to guide academic decisions. The system can produce meaningful insights about academic activity and collaborations and assist in decision-making about research. Its interface is user-friendly and intuitive. The visualisations utilise the user's perception abilities and make the process of exploring the results easier. In this section, we will present an example of the way the system can assist the process of institutional decision-making by indicating a significant question about academic activity that can be answered by our system, as well as the outline of the process that our system is following to answer those questions. The user wants to find the most qualified researchers and explore if they could collaborate. To find the most active researchers using the IREMA system, we would follow the following actions: to find the most connective researchers, to examine whether they have any common research area and which is the development of their common area due to the time.

V. Conclusion

Our approach captures the academic activities, co-operations, and complex connections generated in academia. Furthermore, it analyses academic information using metrics and employs interactive visualisations to give answers to significant questions related to institutional management. The approach we propose can be applied in various cases, the most significant of which is to any VIVO instance. Our future work comprises the following steps. We prioritise making the connection between the IREMA and the VIVO bidirectional. To be more precise, we want to return the results from IREMA to our VIVO instance and comprehensively present them. We have already examined and appended to our ontology all the required fields. Another aspect is the integration of educational activities and collaborations, as well as the connections between education and research in the IREMA. We have already enriched our ontology with the required education fields and those that reveal connections between education and research. We have also accumulated the educational records from our institution in our VIVO instance. The next step is to adjust IREMA and select what visualisations would be helpful for the analysis of education and which will be applied to explore the intersections of education and research in HEIs. Finally, we want to populate a data analysis endpoint (API) from which it will be possible to perform analysis in any VIVO instance.

VI. References

[1] Tsolakidis, A., Sgouropoulou, C., Papageorgiou, E., Terraz, O., & Miaoulis, G. (2013). Using Visual Representation for Decision Support in Institutional Research Evaluation. In Intelligent Computer Graphics 2012 (pp. 41-57). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31745-3.

[2] Triperina, E., Sgouropoulou, C., Xydas, I., Terraz, O., & Miaoulis, G. (2015). Creating the context for exploiting linked open data in multidimensional academic ranking. International Journal of Recent Contributions from Engineering, Science & IT (iJES), 3(3), 33-43. http://dx.doi.org/10.3991/ijes.v3i3.5023.

[3] Anastasios, T., Sgouropoulou, C., Xydas, I., Terraz, O., & Miaoulis, G. (2011, September). Academic research policy-

making and evaluation using graph visualisation. In Informatics (PCI), 2011 15th Panhellenic Conference on (pp. 28-32). IEEE. https://doi.org/10.1109/PCI.2011.38.

[4] Zimmerman, E. (2002, August). CRIS-Cross: current research information systems at a crossroads. In Proceedings of the 6th International Conference on Current Research Information Systems, University of Kassel (pp. 11-20).

[5] Koopmans, N. I. (2002). What's your question? The need for research information from the perspective of different user groups.

[6] Conlon, M., & Corson-Rikert, J. (2012). VIVO: A Semantic Approach to Scholarly Networking and Discovery (Vol. 2). Morgan & Claypool Publishers. https://doi.org/10.2200/S00428ED1V01Y201207WEB002.

[7] EuroCRIS, Current Research Information Systems, www.eurocris.org

[8] Ebeling, B., Klages, M., & Breitner, M. H. (2011). IT-gestütztes Management von Drittmitteln im Rahmen des Forschungsmanagements an deutschen Hochschulen. Informatik.

[9] Azeroual, O., & Abuosba, M. (2019). Improving the data quality in the research information systems. arXiv preprint arXiv:1901.07388.

[10] Bevan, S., & Harrington, J. (2011). Managing research publications: lessons learned from the implementation of a Current Research Information System. Serials, 24(1). https://doi.org/10.1629/2426.

[11] Zelepukhina, V. A., Danilova, T. S., Burmistrov, A. S., & Tarasevich, Y. Y. (2014). Particular experience in design and implementation of a Current Research Information System in Russia: national specificity. Procedia Computer Science, 33, 168-173. https://doi.org/10.1016/j.procs.2014.06.028.

[12] Johansson, Å., & Ottosson, M. O. (2012). A national current research information system for Sweden.

[13] Biesenbender, S., Petersohn, S., & Thiedig, C. (2019). Using Current Research Information Systems (CRIS) to showcase national and institutional research (potential): research information systems in the context of Open Science. Procedia Computer Science, 146, 142-155. https://doi.org/10.1016/j.procs.2019.01.089.

[14] Scholze, F., & Maier, J. (2012). Establishing a research information system as part of an integrated approach to information management: Best practice at the Karlsruhe Institute of Technology (KIT). Liber Quarterly, 21(2), 201-212. https://doi.org/10.18352/lq.8019.

[15] Dimou, A., De Vocht, L., Van Grootel, G., Van Campe, L., Latour, J., Mannens, E., & Van de Walle, R. (2014). Visualising the information of a linked open data enabled research information system. Procedia Computer Science, 33, 245-252. https://doi.org/10.1016/j.procs.2014.06.039.

[16] Parinov, S., Lyapunov, V., Puzyrev, R., & Kogalovsky, M. (2015). Semantically enrichable research information system SocioNet. In Knowledge Engineering and Semantic Web: 6th International Conference, KESW 2015, Moscow, Russia, September 30-October 2, 2015, Proceedings 6 (pp. 147-157). Springer International Publishing. https://doi.org/10.1007/978-3-319-24543-0_11.

[17] Cao, H., He, H., & Tian, J. (2022). A Scientific Research Information System via Intelligent Blockchain Technology for the Applications in University Management. Mobile Information Systems, 2022. https://doi.org/10.1155/2022/7512692.

[18] Benmoussa, K., Laaziri, M., Khoulji, S., & Kerkeb, M. L. (2018). Intelligent system for the use of the scientific research information system. International Journal of Advanced Computer Science and Applications, 9(6). https://doi.org/10.14569/IJACSA.2018.090619.

[19] Benmoussa, K., Laaziri, M., Khoulji, S., Larbi, K. M., & El Yamami, A. (2019). Enhanced model for ergonomic evaluation of information systems: application to scientific research information system. International Journal of Electrical and Computer Engineering, 9(1), 683. https://doi.org/10.11591/ijece.v9i1.pp683-694.

[20] Asserson, A., Jeffery, K. G., & Lopatenko, A. (2002). CERIF: past, present and future: an overview.

[21] Pinto, C. S., Simões, C., & Amaral, L. (2014). CERIF–Is the standard helping to improve CRIS?. Procedia Computer Science, 33, 80-85. https://doi.org/10.1016/j.procs.2014.06.013.

[22] Parinov, S. (2014). Towards an open data on how the research data are used: CRIS-CERIF based approach. Procedia Computer Science, 33, 55-59. https://doi.org/10.1016/j.procs.2014.06.009.

[23] Corson-Rikert, J., Mitchell, S., Lowe, B., Rejack, N., Ding, Y., & Guo, C. (2012). The VIVO ontology. In VIVO: A Semantic Approach to Scholarly Networking and Discovery (pp. 15-33). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-79435-3_2.

[24] Krafft, D. B., Cappadona, N. A., Caruso, B., Corson-Rikert, J., Devare, M., & Lowe, B. J. (2010, April). Vivo: Enabling national networking of scientists. InProceedings of the Web Science Conference (Vol. 2010, pp. 1310-1313).

[25] Albert, P., Holmes, K. L., Borner, K., & Conlon, M. (2012, June). Research discovery through linked open data. In Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries (pp. 429-430). ACM. https://doi.org/10.1145/2232817.2232926.

VII. AUTHORS

Dr. Anastasios Tsolakidis received his PhD degree in computer science from the University of Limoges, France, in 2015. His research interests lie in the fields of Visual Analytics, Decision Support Systems, Business Intelligence and E-health. During his PhD studies, he has been collaborating with the Quality Assurance Unit of the Technological Educational Institute of Athens, as Data Scientist and since July 2017 he has been working as Business Intelligent Analyst at "e-Government Center for Social Security (IDIKA SA)" at the sector of E-Health

Evangelia Triperina holds a PhD in Computer Science from the University of Limoges (France), with a thesis entitled "Visual interactive knowledge management for multicriteria decision making and ranking in linked open data environments". She is a Department of Computer Engineering graduate of TEI of Athens. She holds an MSc in Information Technology, Image Synthesis and Computer Graphics from the University of Limoges (France). She worked in European research projects at GRNET, Agro-Know Technologies and the University of West Attica.

Ioannis D. Triantafyllou holds a PhD from the National Technical University of Athens, Department of Electrical & Computer Engineering, and is currently an Associate Professor at the Department of Archives, Library and Information Studies at the University of West Attica. He has previously worked as a research associate in many European and Greek

research/projects at the Institute of Language & Speech Processing (ILSP /Athena RC). Recently, he participated in the CrossCult research program (Horizon2020) as a research team member. He specialises in Digital Libraries, Data & Text Mining, Text Classification & Clustering, Ontologies & Metadata, Linked Data, Information Extraction, Text & Information Retrieval, Automated Summary & Text Synthesis, Translation Memories, etc.

Christos Skourlas is an emeritus professor at the Department of Informatics and Computer Science of the University of West Attica. He was an analyst-programmer and head of the systems with the National Documentation Centre of Greece (1983- 89) and a research assistant with the Nuclear Research Centre "Demokritos" (1977-82). He was head of the research lab "Data, Information and Knowledge Management (InfoDat_KM)". He participates as a coordinator and/or key researcher in European and nationally funded research and development projects. His research work has been published in international journals and conference proceedings.