Information Research - Vol. 29 No. 1 (2024)

Assessing the influence of bibliometric factors and organizational characteristics on the centrality degree of inter-university collaborative networks: a neural network approach

Juan David Reyes-Gómez and Efrén Romero
DOI: https://doi.org/10.47989/ir291427

Abstract

Introduction. The centrality degree of a university collaborative network indicates how many other universities the given university has active collaborations with. The study analyses the centrality of university-level collaboration networks and aim to assess the influence of organizational characteristics and bibliometric factors of universities on the centrality degree.

Method. This study used artificial neural networks, particularly a multilayer perceptron. The input variables included number of documents published, citations, size, type, and location of the university. Data was extracted from the census of institutions identified within the inter-university collaborative networks of Santander and Caldas in Colombia. A total of 154 universities comprises the dataset for the territory of Santander and 126 for Caldas.

Results. The results indicated that bibliometric factors had a significant influence on the centrality degree of the networks. Organizational characteristics also had an influence, but to a lesser extent than bibliometric factors.

Conclusion. The study found that the research output and impact are the most important factors in predicting the centrality degree of a university in a collaborative network. This suggests that policies to increase the research output and impact of a university are likely to result in a more central position in the network.

Introduction

The globalization of higher education and the rise of international cooperation in scientific research have made the collaborative networks of universities more and more important. Inter-university collaboration networks provide opportunities for institutions to share resources and expertise, and to develop new research and teaching partnerships. These networks can help to increase the visibility of a university and its programmes, and to build its reputation as a leading research institution. In addition, university collaborative networks can help to attract new faculty and students, and to secure funding for joint research projects (Edgar and Geare, 2013; Rotolo and Messeni Petruzzelli, 2013)

The centrality degree of a university collaborative network indicates how many other universities the given university has active collaborations with. A high centrality degree means that the university is highly central within the collaborative network, while a low centrality degree means that the university is less central within the network. The size, type, location, research output, and research impact of universities are all critical factors that affect the centrality degree of university collaborative networks. The centrality degree evidences the university's ability to collaborate with other universities. Universities that are more central within university collaborative networks are better able to form and maintain relationships with other universities. These relationships are important for the exchange of ideas, the sharing of resources, and the formation of partnerships. The centrality degree of a university therefore affects the university's ability to engage in research collaboration.

There are several reasons why it is relevant to study the centrality degree of inter-university collaborative networks. One reason is that centrality can be used as a measure of the importance or influence of a university within a network. This is important because it can help researchers identify which universities are more influential within the collaborative network, and this information can be used to study the network as a whole. Additionally, the centrality degree can also be used to identify potential leaders within the network, which can be helpful in identifying potential collaborators or in managing the network itself.

We believe that understanding the bibliometric factors and organizational characteristics that impact a university's centrality degree is valuable for policy-makers as they seek to effectively design academic collaboration strategies to improve the competitiveness of their institutions.

This study aims to delve into the interplay between organizational characteristics, including size, type, and location of universities, and bibliometric factors, such as research output and impact, on the centrality degree of university collaborative networks. The research question we address in this paper is: What are the bibliometric and organizational factors that influence the centrality degree of university collaborative networks more?

Moreover, this research will shed light on the relative importance of bibliometric factors compared to organizational characteristics in predicting a university's centrality degree. While organizational characteristics provide insights into the structural and cultural aspects of universities, bibliometric factors offer quantifiable measures of research productivity and impact (Bonaccorsi et al., 2021). Understanding the relative significance of these factors will enable stakeholders to devise targeted interventions aimed at enhancing a university's collaborative engagement and positioning within these networks.

To achieve these objectives, we employ a case study research design focusing on the interuniversity collaborative networks of Santander and Caldas in Colombia. These regions are selected due to their high concentration of academic institutions and research centres, making them pivotal hubs of academic activity within the country (Quintero-Quintero et al., 2021). By comparing the characteristics and centrality degrees of universities in these territories, we can elucidate the specific factors that influence a university's centrality within the collaborative networks.

This paper explores the potential of artificial neural networks in assessing the influence of bibliometric factors and organizational characteristics of the universities on the centrality degree of inter-university networks. We apply artificial neural networks to model complex relationships between inputs and outputs. They can be used to predict the centrality degree of academic networks. The advantage of using artificial neural networks is that it can handle nonlinear relationships between inputs and outputs. In addition, they can be used to identify patterns in data that are too complex for humans to identify. By leveraging the potential of artificial neural networks to model complex relationships, we seek to unravel the underlying dynamics that influence a university's centrality within these networks.

This study provides a comprehensive analysis of the centrality degree in collaborative networks, focusing on the influence of bibliometric factors and organizational characteristics on the centrality of universities. The theoretical framework, including previous research findings and theoretical perspectives, is discussed in detail. The methodology employed in this study, which utilizes artificial neural networks as a powerful analytical tool, is described thoroughly. The results and their implications are presented, followed by a critical discussion of the findings and their limitations. Finally, the paper concludes with recommendations for policy makers and university administrators to enhance the centrality of universities within collaborative networks.

Theoretical background

Collaborative networks centrality

The literature on university collaborative networks is still sparse, although it has increased in recent years. For the purposes of this paper, we apply a relational capital-based approach in the context of universities (e.g., Andrews, 2010; Bucheli et al., 2012; Cricelli et al., 2018; de Frutos-Belizón et al., 2019; Salinas-Ávila et al., 2020) on the literature that specifically discusses university collaborative networks as a means of promoting and facilitating collaboration between universities. Relational capital refers to the relationships between different universities that allow for the exchange of knowledge and resources (Bucheli et al., 2012; de Frutos-Belizón et al., 2019). These relationships can take the form of formal agreements, such as research collaborations, or informal relationships, such as shared resources or information.

Burris (2004) developed a model to explain how universities are organized into a hierarchical network. The model suggests that each university occupies a position in the network based on its relative size, prestige, and power. A more recent work that has looked at university collaborative networks is a study by Weber and Yuan (2019), which aimed to investigate the research networking systems that help users identify investigators with particular areas of expertise, affiliations, interests, resources, or other characteristics that would make them potential collaborators. The authors discovered that research networking systems are emerging as the predominant source of data about investigators. This shift is expected to alleviate the administrative burden on investigators and provide metrics that more accurately reflect their contributions to collaborations and team science activities. A relevant study by Zhang et al. (2021) discusses the use of a neural network to predict the citation count of a paper. The study found that the neural network was able to predict the citation count of a paper with a high accuracy. The study also found that the factors that had the most impact on the citation count were the novelty of the paper, the bibliometric factors, and the academic-network factors.

Finally, a third group of relevant studies include Bucheli et al. (2012); Cricelli et al. (2018) and Salinas-Ávila et al. (2020), who focus on the role of relational capital on the knowledge generation and the growth of scientific production in Colombian universities. Two relevant mechanisms that emerge from these studies are the capacity for knowledge generation, which depends largely on the availability of relational capital, and the capacity for cooperation and knowledge

sharing, which also depends on relational capital.

Past research found that the centrality degree of a university corelates positively with the number of research papers published by the university, with the number of research collaborations between the university and other institutions, and the number of international research collaborations involving the university (Huggins et al., 2020; Uddin et al., 2013). These studies suggest that university's research productivity and visibility are good predictors of the centrality degree of a university, which in turn appears to be a good predictor of the university's success in international research collaboration.

Bibliometric factors and organizational characteristics of universities

There are two types of factors that influence the productivity and impact of a university's research: bibliometric and organizational. Bibliometric factors are related to the content and structure of the scientific literature, and include aspects such as the number of documents produced by an institution and the number of citations a paper receives, the number of co-authors, and the journal in which it is published. Organizational factors are related to the institutions themselves, including the size and the type of the institution, and the geographic location. Organizational factors are important in understanding how universities are interconnected and how they can best collaborate with each other.

Bibliometrics factors and inter-university collaborative networks centrality: research output and impact

The research output of a scientific institution is typically measured by the number of documents published in peer-reviewed journals. This metric is used because it is generally accepted that peer-reviewed journal articles are the most reliable and valid source of scientific knowledge. However, there are some limitations to this metric. First, not all scientific knowledge is published in peer-reviewed journals. Second, the number of documents published does not necessarily reflect the quality or impact of the research. For example, a small number of high-quality papers can have a greater impact than a large number of lower-quality papers.

Despite these limitations, the number of documents published is still the most commonly used metric for research output. This is because it is the most objective and easily measurable metric. Additionally, the number of documents published is a good proxy for the amount of research being conducted at an institution (Liang et al., 2021).

The number of documents published is also affected by the number of authors on a paper. Typically, papers involving a greater number of authors tend to garner higher citation rates (Gazni & Didegah, 2011; Katz & Martin, 1997). Moreover, collaboratively written papers exhibit an increased likelihood of achieving exceptionally high citation levels, thereby enhancing the prospects of publication as well (Wuchty et al., 2007).

A large body of research has shown that the number of documents published is positively associated with the centrality degree of university collaborative networks (e.g., Leydesdorff & Wagner, 2008). In other words, universities that publish more research output are tend to more central in collaborative networks. However, a study by Rotolo and Messeni Petruzzelli (2013) suggest that at the individual scientific level, it is observed that scientific productivity experiences a decrement once a specific threshold value of centrality is surpassed. This observation highlights the presence of an inverted U-shaped relationship, signifying that elevated levels of social capital linked to centrality carry an associated opportunity cost. This cost pertains to the potential hindrance in the knowledge creation process as heightened social capital may expose a scientist to uncooperative behaviors, reluctance, and potential sabotage from other individuals or groups within the same organizational or community.

The research impact of a paper is typically measured by the number of citations it has

received from other papers (Aksnes et al., 2019; Liao, 2011; Thelwall and Maflahi, 2020). The more citations a paper has, the greater its impact. Impact is an essential metric for assessing the quality of research (Zhao et al., 2019). It is used to compare the relative importance of different papers and to identify the most influential papers in a field. This metric is used because it is a good proxy for the paper's influence. However, there are some limitations to this metric. First, it only captures the impact of papers that have been published. Second, it does not consider the quality of the citations. A paper with many citations from low-quality journals is not necessarily more impactful than a paper with fewer citations from high-quality journals. Third, the number of citations a paper has received does not always reflect its quality. A paper that is highly cited may be of poor quality, and vice versa. Fourth, the number of citations a paper has received is not always a good predictor of its future impact. A paper that is highly cited in its first year may not be highly cited in subsequent years.

Finally, the number of citations a paper represent a good but not perfect impact measure (Aksnes et al., 2019). The citations received by a paper are not always a good measure of its impact on society. A paper that is highly cited may not have had a significant impact on society, and vice versa. Societal relevance is often considered to be something which is much harder to measure than scientific relevance or impact (Aksnes et al., 2019; Martin, 2011). Despite these limitations, the number of citations a paper has received is still the most commonly used metric for assessing its impact (Martin, 2011).

The literature on the centrality degree of university collaborative networks remains inconclusive about the most effective measurement method (Martin, 2011). Nevertheless, it does indicate that the number of citations received by each university serves as a reasonable proxy for its importance within the network as it can be easily calculated and compared across different networks (Tahamtan et al., 2016; Tahamtan & Bornmann, 2019).

For example, a study of the relationship between centrality and research impact in the context of the *world-wide web* found that the most central nodes in the web were also the most highly cited (Tian et al., 2021).

Citations may be a particularly important predictor of centrality degree in disciplines where collaboration is essential to progress, such as in the natural sciences. In such fields, papers resulting from collaboration among a large number of universities are more likely to be highly cited than those resulting from collaboration between a small number of universities.

One study found that the centrality degree of university collaborative networks in Colombia was positively correlated with the number of citations received by the universities in the network (Bucheli et al., 2012). This finding suggests that the more central a university is in the collaborative network, the more likely it is to be highly cited.

Organizational characteristics and collaborative networks centrality: size, type and location of the university

The centrality degree of university collaborative networks has been found to be positively associated with the size of the university, as measured by the number of students enrolled (Chen et al., 2020; Liao, 2011). In other words, universities with more students tend to have more collaborative relationships with other universities, as measured by the number of co-authored papers and joint research projects. This finding is consistent with the idea that larger universities are more likely to have the resources and the incentive to engage in collaborative research with other universities.

Chen et al. (2020) found that, when controlling for a variety of factors, the size of the university had the strongest positive association with the centrality degree of the university's collaborative network. This finding suggests that, other things being equal, a university with more students is more likely to be more central in the university collaborative network.

Theoretically, it is expected that larger universities have a higher centrality degree due to their greater ability to attract resources and generate knowledge. However, this relationship has not been extensively studied. The few studies that have been conducted suggest that university size is not a significant predictor of centrality in university collaborative networks. For example, a study of university-industry collaboration in Colombia found that university size was not a significant predictor of the number of collaborative agreements signed by universities (Cricelli et al., 2018). The absence of empirical evidence on the impact of university size on the centrality degree of university collaborative networks suggests the need for further research to investigate this relationship.

The type of university significantly influences the centrality degree within collaborative networks, owing to distinct organizational structures and cultural disparities present in public and private universities (Cricelli et al., 2018; Salinas-Ávila et al., 2020). Public universities often adopt hierarchical and bureaucratic structures, fostering departmental silos that impede cross-departmental relationship building (Cricelli et al., 2018). This hierarchical framework can create difficulties for individuals seeking to establish connections beyond their specific department. Conversely, private universities tend to exhibit more flexibility and encourage interdisciplinary collaboration, facilitating easier networking across departments or faculties (Salinas-Ávila et al., 2020). Such cultural differences and organizational structures significantly shape the formation and sustainability of collaborative networks within academic settings.

The geographic location of a university has been found to influence the centrality degree of university collaborative networks, since institutional collaboration decreases exponentially as the geographic distance separating collaborative partners increases (Kenekayoro et al., 2014). In particular, universities located in the same country tend to have more centrality in their collaborative networks than universities located in different countries. This may be due to the fact that universities in the same country are more likely to share common research interests and to be familiar with each other's work. Additionally, the proximity of universities may make it easier for them to establish and maintain collaborative relationships.

The impact of ICT on the establishment of relations among universities has been particularly evident during the COVID-19 pandemic. The widespread use of video conferencing and other online communication tools has made it possible for universities to continue to collaborate despite the physical distance between them. In many cases, the pandemic has actually led to an increase in university collaboration, as universities have turned to each other for help in addressing the challenges posed by the pandemic (Kalmar et al., 2022).

Overall, the geographic location of a university does appear to influence the centrality degree of university collaborative networks. However, the impact of ICT has made it possible for universities to overcome the geographical barriers to collaboration, and in some cases, has actually led to an increase in collaboration among universities.

Artificial neural networks

According to Yaghi et al. (2020), artificial neural networks are structures that are inspired by the functioning of the brain. These networks can perform model function estimation and handle linear or nonlinear functions by learning from data relationships and generalizing to unobserved situations. A popular type of artificial neural network is a multilayer perceptron. Multilayer perceptrons are a kind of mathematical models inspired by biological neural networks which are used to approximate functions from many inputs. They can learn from data relationships and generalize to unseen situations. Multilayer perceptrons are a type of artificial neural network that is composed of an input layer, hidden layers, and an output layer. Information flows in one direction from the input layer to the output layer. Each connection between neurons has its own weight.

A neural network can approximate a wide range of statistical models without the need to hypothesize in advance certain relationships between the dependent and independent variables. This non-*a priori* model should be more accurate than the classical parametric methods of logistic or linear regression, as proposed for instance by Persson (2017) to evaluate co-author networks and by Yu et al. (2014) to predict citations impact. This theoretical superiority is presupposed from the high complexity, computational power and learning capability associated with nonparametric approaches (Heazlewood et al., 2016). Neural networks are good at capturing non-linear relationships. Finally, neural networks are good at generalizing from data. This is important because it allows the findings from the investigation to be applied to other university collaboration networks.

The relative importance output of a multilayer perceptron can be interpreted as the relative importance of each input variable in predicting the output variable. The output of a multilayer perceptron is a list of the input variables in order of importance, with the most important variable listed first. The importance of each input variable is quantified by a percentage, which indicates the percentage of the total variance in the output variable that is explained by that input variable. The output of a multilayer perceptron can be used to identify the most important input variables for a given output variable. This information can be used to simplify the input data set, or to focus on the most relevant input variables when building a predictive model.

Methods

Data collection and case selection

Our research design is a case study of two territories within Colombia (departments of Santander and Caldas). A case study research design is appropriate for this study because it enables an in-depth understanding of the development and operation of university collaborative networks among territories. Furthermore, a case study allows for comparisons between the two territories, facilitating a more accurate and comprehensive analysis of the data. Including both territories in the study helps control for any confounding factors that may exist between the two collaborative networks, providing a more accurate depiction of the true relationships between the variables of interest. Additionally, incorporating both territories in the analysis allows for an increased sample size, thereby enhancing the power of the study.

The departments of Santander and Caldas are home to a number of academic institutions and research centres, which makes them crucial centres of academic activity in Colombia. These organizations are valuable for the development of the region and the country. They contribute to the economic development and provide jobs. They also contribute to the knowledge base and help to improve the quality of life for the people of the region. As Santander and Caldas are two of the most populous regions in Colombia, their inclusion enhances the comprehensiveness of the country's university collaborative networks. Furthermore, these two territories have a high concentration of universities, rendering them prominent nodes in the network. The inclusion of these regions also facilitates a more fine-grained analysis of the organizational characteristics of universities regarding the centrality degree of university collaborative networks.

For this study, we collected data from agricultural innovation systems, which are networks of actors (universities, research centres, territorial entities, enterprises, NGOs) working collaboratively (Galeano-Barrera et al., 2022). Universities play a central role in the collaboration between actors in innovation systems at the global level, through functions and roles such as knowledge brokering, supported by their social capital (Romero-Riaño et al., 2019). The heterogeneity of scientific collaboration networks is a facilitator of innovation diffusion and the impact of research results.

We analyse the citation networks of authors and organizations in agricultural research in Colombia, as an academic contribution to the understanding of the productivity and impact of universities, as well as collaboration within this discipline. Data was extracted from the

study conducted by Romero-Riaño et al. (2019) corresponding to the census of institutions identified within the collaborative networks of Santander and Caldas for the period 2014-2018. A total of 154 universities comprises the dataset for the territory of Santander and 126 for Caldas.

Variables definition

The independent variables used in this study include bibliometric and organizational factors of the universities that influence the degree of centrality of the collaborative network (see Table 1). This model contains five input variables and one output variable. The variables chosen for this study were size, expressed as a scalar variable, and type and location of the university, which were expressed as dummy variables. The rationale for this decision is that they are commonly used in studies with similar methodological designs.

In the case of the variable type, a value of one was assigned to local universities, whose main campus is located in Caldas or Santander, and zero to foreign universities, whose main campus is located outside Colombia. The nature of the university was categorized as follows: public universities were assigned a value of one and private universities were assigned a value of zero. The values for the size of the university were extracted from the official pages of the institutions. The scores for number of documents and number of citations of the universities were extracted from Scopus database.

Variable	Type	Description	Measurement
Degree of centrality	Output	Active collaborations of a university	Number of direct collaborative links of a university
Size	Input	The size of the university in terms of enrolled students	Number of students enrolled students
Type	Input	The type of the university in terms of ownership	1 public, 0 private
Location	Input	The location of the university in terms of the site of the main campus	1 domestic, 0 foreign
Documents	Input	Number of documents published	Number of documents published by a university during the period
Citations	Input	Number of citations obtained by documents published	Number of citations obtained by documents published by a university during the period

Table 1. Description of the input or independent and output or dependent variables

Neural network model

In the context of our paper, artificial neural networks, particularly multilayer perceptron, are used to assess the importance of the input variables over the output variable. The artificial neural networks will be able to make predictions, in this case, about how bibliometric and organizational factor of university networks may impact the centrality degree of a university collaboration network.

The first step in building the model is to divide the data collected into two parts: training and validation data. The training data comprises 70% of the data, while the validation data comprises 30%. The neural network model was built using the software IBM SPSS v. 26. The multilayer perceptron typology implemented was as follows: the multilayer perceptron architecture is an input layer with two scalar factors and three categorical covariables as independent or, input variables, and seven

units or neurons; a hidden layer with five unit or neuron and hyperbolic tangent activation function; and an output layer with one dependent variable with one unit or neuron, and identity activation function and cross-entropy error function. The training data were used to obtain the best configuration of the multilayer perceptron architecture, while the validation data were used to validate the model.

The default error function is sum of squares error, which is the sum of differences between each observation group means, and it is used to indicate the variation within a cluster. Therefore, the smaller the sum of squares error, the less the variation between the cluster.

Results

The case processing summary for multilayer perceptron network is shown in Table 2. The dataset used in this study is divided into two groups which are training and testing comparing both cases analyzed. The training set for multilayer perceptron is 66.2% for Santander (102/154) and 69% for Caldas (87/126). The testing set is 33.8% (52/154). The overall data from the dataset is N = 154 and N = 126, no data were excluded.

		Santander		Caldas	
		n	Percentage	N	Percentage
Sample	Training	102	66,2%	82	65,1%
	Testing	52	33,8%	44	34,9%
Valid		154	100%	126	100%
Excluded		0	0%	0	0%
Total		156	100%	126	100%

Table 2. Case processing summary

		Santander	Caldas
Training	Sum of squares error	11,01	4,26
	Relative error	21,8%	0,9%
Testing	Sum of squares error	1,40	0,29
	Error relative	26,7%	32,6%

Table 3. Summary of the model

Regarding the normalized importance of the input variables, results indicate that the bibliometric factors, such as number of documents (100% and 100%) and citations (79,5% and 48,7%), are the most influential in determining the degree of centrality of both universities, as shown in Table 3.

	Santander	Caldas
Type	5,1%	3,7%
Location	4,3%	3,0%
Documents	100,0%	100,0%
Citations	79,5%	48,7%
Size	20,8%	12,6%

Table 4. Relative importance of independent variables

The results indicate that the number of documents, number of citations, type, size, and location are important factors influencing the centrality degree of university collaborative networks. The most important factor is the number of documents, followed by the number of citations, the type of university, the size, and finally the location.

Discussion and conclusions

Artificial neural networks are a well-known and widely used machine learning technique. In the past few years, artificial neural networks have been applied to a variety of tasks in the field of information systems, with promising results. In this study, we applied a multilayer perceptron to assess the influence of bibliometrics and organizational factor on the degree of centrality of an inter-university network within two agricultural innovation systems in Colombia (Santander and Caldas).

We found that the main drivers of the centrality of a university in the network are the research output, measured by the number of documents published, and the research impact, measured by the number of citations. We also found that location, size, and type of the university have a residual importance.

There are two main explanations for these findings: first, the research output and impact reflect the research capabilities of a university. Therefore, it is not surprising that these two factors are the most important ones in predicting the centrality degree of a university in a collaborative network. That is, the more influential a university is, the more central it is. Central universities tend to have more resources and capabilities and usually are more visible and accessible to other universities, which makes them more attractive partners for collaboration, and thus they are more likely to be involved in collaborative research projects.

Second, the location, size, and type of the university are usually not indicative of the centrality degree of a university. For example, a small university located in a rural area may have a low research output, but a high research impact. Similarly, a large university located in an urban area may have a high research output, but a low research impact. In this sense, the location of a university is not a barrier to the collaborative behaviour, given that we can find a great amount of collaboration between universities that are located in different cities or countries. Finally, the size of a university is not a good predictor of the collaborative behaviour, because it is possible to find a great amount of collaboration between small and large institutions. Therefore, it is not surprising that these factors are not as important as the research output in predicting the centrality degree of a university in a collaborative network.

The findings of this study contribute to the understanding of the collaborative behaviour of universities in innovation systems. In particular, the results suggest that bibliometrics factors are the most important predictors of the centrality of a university in a collaborative network. This is a useful finding for policy makers and university

administrators, since it suggests that measures to improve the research output and impact of a university are likely to result in a more central position in the network. These factors should be considered when making decisions about research funding, university rankings, and collaborative partnerships. In particular, policies aimed at enhancing the research output and impact of universities should be considered to increase the centrality of universities in collaborative networks.

The study's results should be taken in light of their limitations. Due to the nature of the research, the paper is limited to using bibliometric data and organizational characteristics to assess the centrality degree of inter-university collaborative networks. The paper does not consider other factors that may influence the centrality degree, such as personal relationships or geographical proximity. Additionally, the paper is limited to using data from the Scopus database, which may not be representative of all inter-university collaborative networks. Data collected is specific to the task of assessing the influence of bibliometrics and organizational factors on the degree of centrality of a university network, and they're based on a specific dataset from two agricultural innovation systems in Colombia. These conclusions may not be generalizable to other types of machine learning tasks or other datasets.

Future research could be directed at assessing the influence of other factors on the centrality of university networks, such as the number of publications, the number of citations, the number of collaborations, or the number of external research funding. Additionally, future research could be directed at assessing the centrality of university networks in other contexts, such as in other countries.

Acknowledgements

The authors extend their gratitude for the support received during the postdoctoral stay for strengthening vocations and training in Science, Technology, and Innovation for economic reactivation in the post-pandemic context. This work was made possible through the 2020 Call 891 by the Ministry of Science, Technology, and Innovation of Colombia (MinCiencias). Special thanks to the Colombian Observatory of Science, Technology, and Innovation for their invaluable contributions.

About the authors

Dr. Juan David Reyes-Gómez holds a Ph.D. in Entrepreneurship and Management and serves as an Assistant Professor at University College of Cundinamarca in Colombia. Concurrently, he is a Postdoctoral Researcher at the Colombian Observatory of Science and Technology, leading a project that leverages data analytics to contribute to knowledge generation in science, technology, and innovation in Colombia. His research aims to provide inputs for policy recommendations in STI and enhance the national and international visibility of scientific activities.

Efren Romero is a Ph.D. candidate in Engineering at Autonomous University of Bucaramanga, focusing on developing innovative strategies derived from research findings through scientometric data analytics. Over the past six years, he has dedicated his efforts to advancing novel visualization models and fostering the utilization of open science data in business innovation. As the Deputy Director of Research at the Colombian Observatory of Science and Technology, he also conducts training sessions on visualization, scientometrics, and social network analysis at the University of Leiden, Netherlands.

References

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1). https://doi.org/10.1177/2158244019829575

Andrews, R. (2010). Organizational social capital, structure and performance. *Human Relations*, 63(5), 583–608. https://doi.org/10.1177/0018726709342931

Bonaccorsi, A., Belingheri, P., & Secondi, L. (2021). The research productivity of universities. A multilevel and multidisciplinary analysis on European institutions. *Journal of Informetrics*, 15(2), 101129. https://doi.org/10.1016/j.joi.2020.101129

Bucheli, V., Díaz, A., Calderón, J. P., Lemoine, P., Valdivia, J. A., Villaveces, J. L., & Zarama, R. (2012). Growth of scientific production in Colombian universities: An intellectual capital-based approach. *Scientometrics*, 91(2), 369–382. https://doi.org/10.1007/s11192-012-0627-7

Burris, V. (2004). The academic caste system: Prestige hierarchies in PhD exchange networks. *American Sociological Review*, 69(2), 239–264. https://doi.org/10.1177/000312240406900205

Chen, K., Zhang, Y., Zhu, G., & Mu, R. (2020). Do research institutes benefit from their network positions in research collaboration networks with industries or/and universities? *Technovation*, 94–95(January 2016), 102002. https://doi.org/10.1016/j.technovation.2017.10.005

Cricelli, L., Greco, M., Grimaldi, M., & Llanes Dueñas, L. P. (2018). Intellectual capital and university performance in emerging countries: Evidence from Colombian public universities. *Journal of Intellectual Capital*, 19(1), 71–95. https://doi.org/10.1108/JIC-02-2017-0037

de Frutos-Belizón, J., Martín-Alcázar, F., & Sánchez-Gardey, G. (2019). Conceptualizing academic intellectual capital: definition and proposal of a measurement scale. *Journal of Intellectual Capital*, 20(3), 306–334. https://doi.org/10.1108/JIC-09-2018-0152

Edgar, F., & Geare, A. (2013). Factors influencing university research performance. *Studies in Higher Education*, 38(5), 774–792. https://doi.org/10.1080/03075079.2011.601811

Galeano-Barrera, C. J., Mendoza-García, E. M., Martínez-Amariz, A. D., & Romero-Riaño, E. (2022). Theoretical model of territorial agro-industrial development through multi-focus research analytics. *Journal of Rural Studies*, 94, 295–304. https://doi.org/10.1016/j.jrurstud.2022.06.014

Gazni, A., & Didegah, F. (2011). Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics*, 87(2), 251–265. https://doi.org/10.1007/s11192-011-0343-8

Heazlewood, I., Walsh, J., Climstein, M., Kettunen, J., Adams, K., & DeBeliso, M. (2016). A comparison of classification accuracy for gender using neural networks multilayer perceptron (MLP), radial basis function (RBF) procedures compared to discriminant function analysis and logistic regression based on nine sports psychological constructs t. *Advances in Intelligent Systems and Computing*, 392, 93–101. https://doi.org/10.1007/978-3-319-24560-7_12

Huggins, R., Prokop, D., & Thompson, P. (2020). Universities and open innovation: the determinants of network centrality. *Journal of Technology Transfer*, 45(3), 718–757. https://doi.org/10.1007/s10961-019-09720-5

Kalmar, E., Aarts, T., Bosman, E., Ford, C., de Kluijver, L., Beets, J., Veldkamp, L., Timmers, P., Besseling, D., Koopman, J., Fan, C., Berrevoets, E., Trotsenburg, M., Maton, L., van Remundt, J., Sari, E., Omar, L. W., Beinema, E., Winkel, R., & van der Sanden, M. (2022). The COVID-19 paradox of online collaborative education: when you cannot physically meet, you need more social interactions. *Heliyon*, 8(1), e08823. https://doi.org/10.1016/j.heliyon.2022.e08823

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26(1), 1–18. https://doi.org/10.1016/S0048-7333(96)00917-1

Kenekayoro, P., Buckley, K., & Thelwall, M. (2014). Hyperlinks as inter-university collaboration indicators. *Journal of Information Science*, 40(4), 514–522. https://doi.org/10.1177/0165551514534141

Leydesdorff, L., & Wagner, C. S. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, 2(4), 317–325. https://doi.org/10.1016/j.joi.2008.07.003

Liang, Z., Mao, J., Lu, K., & Li, G. (2021). Finding citations for PubMed: a large-scale comparison between five freely available bibliographic data sources. *Scientometrics*, 126(12), 9519–9542. https://doi.org/10.1007/s11192-021-04191-8

Liao, C. H. (2011). How to improve research quality? Examining the impacts of collaboration intensity and member diversity in collaboration networks. *Scientometrics*, 86(3), 747–761. https://doi.org/10.1007/s11192-010-0309-2

Martin, B. R. (2011). The research excellence framework and the "impact agenda": Are we creating a Frankenstein monster? *Research Evaluation*, 20(3), 247–254. https://doi.org/10.3152/095820211X13118583635693

Persson, R. A. X. (2017). Bibliometric author evaluation through linear regression on the coauthor network. *Journal of Informetrics*, 11(1), 299–306. https://doi.org/10.1016/j.joi.2017.01.003

Quintero-Quintero, W., Blanco-Ariza, A. B., & Garzón-Castrillón, M. A. (2021). Intellectual capital: A review and bibliometric analysis. *Publications*, 9(4), 46. https://doi.org/10.3390/publications9040046

Romero-Riaño, E., Guarín-Manrique, L., Dueñas-Gómez, M., & Becerra-Ardila, L. (2019). Reference framework for capabilities development in agricultural innovation system. *DYNA (Colombia)*, 86(210), 23–34. https://doi.org/10.15446/dyna.v86n210.74475

Rotolo, D., & Messeni Petruzzelli, A. (2013). When does centrality matter? Scientific productivity and the moderating role of research specialization and cross-community ties. *Journal of Organizational Behavior*, 34(5), 648–670. https://doi.org/10.1002/job.1822

Salinas-Ávila, J., Abreu-Ledón, R., & Tamayo-Arias, J. (2020). Intellectual capital and knowledge generation: an empirical study from Colombian public universities. *Journal of Intellectual Capital*, 21(6), 1053–1084. https://doi.org/10.1108/JIC-09-2019-0223

Tahamtan, I., & Bornmann, L. (2019). What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics*, 121(3), 1635–1684. https://doi.org/10.1007/s11192-019-03243-4

Tahamtan, I., Safipour Afshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics*, 107(3), 1195–1225. https://doi.org/10.1007/s11192-016-1889-2

Thelwall, M., & Maflahi, N. (2020). Academic collaboration rates and citation associations vary substantially between countries and fields. *Journal of the Association for Information Science and Technology*, 71(8), 968–978. https://doi.org/10.1002/asi.24315

Tian, S., Xu, X., & Li, P. (2021). Acknowledgement network and citation count: the moderating role of collaboration network. *Scientometrics*, 126(9), 7837–7857. https://doi.org/10.1007/s11192-021-04090-y

Uddin, S., Hossain, L., & Rasmussen, K. (2013). Network effects on scientific collaborations. *PLoS ONE*, 8(2). https://doi.org/10.1371/journal.pone.0057546

Weber, G. M., & Yuan, L. A. (2019). The power of research networking systems to find experts and facilitate collaboration. In Kara L. Hall, Amanda L. Vogel & Robert T. Croyle (eds.) *Strategies for team science success* (pp. 541–562). Springer International Publishing. https://doi.org/10.1007/978-3-030-20992-6_42

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827), 1036–1039. https://doi.org/10.1126/science.1136099

Yaghi, R. I., Faris, H., Aljarah, I., Al-Zoubi, A. M., Heidari, A. A., & Mirjalili, S. (2020). Link prediction using evolutionary neural network models. In Mirjalili, S., Faris, H., Aljarah, I. (eds) *Evolutionary machine learning techniques. Algorithms for intelligent systems* (pp. 85–111). Springer. https://doi.org/10.1007/978-981-32-9990-0_6

Yu, T., Yu, G., Li, P. Y., & Wang, L. (2014). Citation impact prediction for scientific papers using stepwise regression analysis. *Scientometrics*, 101(2), 1233–1252. https://doi.org/10.1007/s11192-014-1279-6

Zhang, X., Xie, Q., & Song, M. (2021). Measuring the impact of novelty, bibliometric, and academic-network factors on citation count using a neural network. *Journal of Informetrics*, 15(2), 101140. https://doi.org/10.1016/j.joi.2021.101140

Zhao, F., Zhang, Y., Lu, J., & Shai, O. (2019). Measuring academic influence using heterogeneous author-citation networks. *Scientometrics*, 118(3), 1119–1140. https://doi.org/10.1007/s11192-019-03010-5