

Construction of a Complex Network Using Technological Surveillance for the Strategic Management of Science, Technology, And Research in Higher Education Institutions in Colombia

Olga Lucía Ostos ^a, Rafael Rentería-Ramos ^b, Fabio Cala ^c

^a Universidad Santo Tomás, Research and Innovation Directorate.

olgaostos@usantotomas.edu.co

^b Universidad Nacional Abierta y a Distancia, Escuela de Ciencias Básicas,

rafael.renteria@unad.edu.co

^c Universidad Jorge Tadeo Lozano, Director Área Académica y Modelado,

fabio.cala@utadeo.edu.co

Abstract: This research proposes the construction of a model of technological surveillance for the strategic management of scientific research in higher education institutions in Colombia. This model is composed by an information system that includes scientific articles, intervention plans for the regions and the nation, and information from public policies related to Science, technology, and information (STI). From this data set, the construction of algorithms for information processing was carried out, using techniques of automatic text analysis and natural language processing with complex networks to detect patterns and the configuration of trends related to science, technology, research, and innovation, which are the guiding topics of the national science and technology system. The main results obtained in this research are that the construction of a complex system revealed the topics on which scientific production is focused in the country. In addition, the analysis of the documents related to STI referenced in public policies analyzed for this research revealed the pressing needs of different sectors and regions in the country. Finally, the research showed that the topics on which scientific research in Colombia is conducted by the educational institutions of the country does not correspond to the needs found.

Keywords: Technology Surveillance, Scientific Research, STI, Innovation, Management.

1. Introduction

The pressing requirements of various productive sectors in Colombia have revealed a growing need to increase the participation of science, technology, and information (STI), as well as research, development, and innovation (R+D+i) as key factors to find solutions to the country's problems. Faced with this urgent need, public and private organizations demand the construction of strategic approaches that allow directing the work of scientific research towards solving the issues identified by the most important stakeholders of the National System of Science, Technology, and Innovation of Colombia (SNCTI in Spanish).

In this context, the construction of tools and mechanisms that allow the early detection of elements that facilitate the construction of strategic approaches is a priority. Therefore, technological surveillance becomes an important tool for organizations. According to Leon (2006):

“Technological surveillance (TS) is a concept inherent to technology management (GT), which involves processes of planning, directing, controlling, and coordinating development and implementing information to understand and anticipate technological changes, making an early detection of events that represent potential opportunities or threats.” (p. 93)

According to the definition above, technological surveillance has a diversity of elements that can be synthesized in the following areas (Palop & Vicente, 1999):

- Passive surveillance (scanning): Detects information that is important to the objectives or goals of a company.
- Active surveillance (monitoring): Directs strategies to promote capacities and knowledge to the organization from the results of the surveillance related to topics that are objective for the organization.
- Watching. It is a system that articulates the two areas above, as well as the popularization of the information obtained from the surveillance.

In particular, the last two areas (active surveillance and watching) are important to produce improvement alternatives related to the competitive and strategic intelligence of the organization, generating value processes from decision-making.

From this perspective, the sources of information are key assets to fulfill the implicit objectives in the previously mentioned areas. Regarding technological surveillance for STI, repositories and other scientific, academic and research popularization media are essential resources. A large part of these volumes of information and popularization media have chosen to maintain open access policies, which facilitate their analysis, accessibility, transformation, and manipulation. Some of these works present relevant information on indicators of the dynamics of STI in the countries (Zhu et al., 2013) and the companies (Liu et al., 2013).

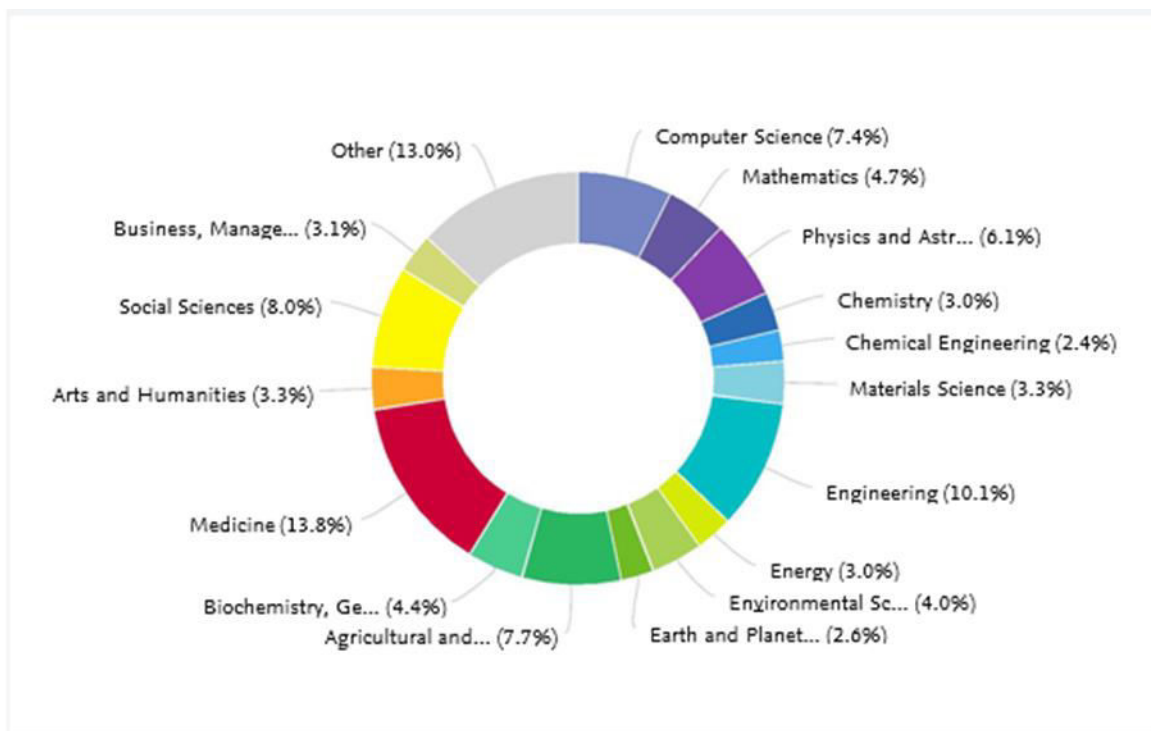
Scopus, and Web of Science (WOS) stand out among the most important STI and R+D+i repositories. They are bibliographic databases that have allowed researchers to create technology surveillance models (the scanning and monitoring paradigm specifically) related to scientific, technological, humanistic, and sociological aspects using bibliometric indicators obtained through analyzing complex networks. (Zhu et al., 2013; Liu et al., 2013; Liu, Yu, Guo, & Sun, 2014; Liu, Yu, Guo, Sun, & Gao, 2014). Building complex networks using sources related to STI has even made possible the generation of public policy instruments. Such is the case of Kash and Rycraft (2000), who, while representing complex systems using networks, analyzed innovation, development, and technology patterns and trends. In addition to their findings, Kash and Rycraft built indicators to evaluate the effectiveness of public policies on the stakeholders of the system, using the emergent self-organization of the interactions between science, technology, and innovation.

In this sense, the objective of this research is to build a model of technological surveillance to strategically direct research in higher education institutions in Colombia using bibliometric analysis techniques based on network theory. To fulfill this objective, this document presents two sections: in the first one, the construction of the information system is described, as well as the computational and statistical techniques selected to clean and pre-process the information to build the networks. In the second, the results of the construction of the information system are presented, including the respective units of analysis and technological surveillance that aims to reduce the existing gap between scientific research and the needs of the sectors, as well as strengthening scientific and research planning of higher education institutions in Colombia.

2. Materials

For this study a data set containing information related to the behavior of STI in articles developed in Colombia and published in high-impact journals indexed in Scopus was constructed. The information included in the data set responded to keywords and summaries found in the articles. Based on them, textual databases were generated. The scope of this research was limited by prioritizing publications made from 2016 to 2021. A total of 66,842 articles were retrieved and their areas of knowledge are presented in the portions shown in Figure 1.

Figure 1. Distribution by areas of knowledge of publications produced in Colombia between 2016-2020.



Source: own elaboration.

In addition to the sources described above, information from the Colombian context describing the main R+D needs in the different regions of the country were also included,

as well as the public policies generated by the central government to promote and prioritize State resources to fulfill those needs. Such policies are conditioned by the intention of generating new knowledge, but they also seek to facilitate the production of instruments and tools that allow solving the main problems faced by the most vulnerable population groups. That is why this research included the construction of a textual database with information related to documents related to STI referenced in public policies in Colombia. In the information included in this database the following topics stand out:

- Sustainable bioscience.
- Strengthening of the human talent training system.
- Policy for the commercial development of biotechnology based on the use of biodiversity.
- Green growth policy.
- Tax benefits.
- 2020 Science, Technology, and Innovation Policy.

3. Methods

At present, building models based on scientific discourse, as well as scientometrics, bibliometrics and other disciplines related to scientific popularization, uses quantitative-qualitative methodologies to recognize patterns, central themes, and feelings, among other morphological aspects of the information disclosed in different scientific repositories. In addition to these components, there are some conditioning factors related to the specificities of the language used (basic, technical, or specialized), the dynamics and the different relationships that can be built between its different units. For this reason, the analysis of scientific discourse must go beyond simple word, term, and concept counts. That is, building models requires multidimensional modeling approaches and formalisms that are useful to establish evaluations of the linguistic entities that are the components of the discourse as a unitary system. In this regard, complex networks become an ideal tool to achieve the aforementioned objectives.

Complex networks are graphs that are useful to build representations of systems (Estrada, 2021). The vertices or nodes of the graph are the entities of the system, and the edges are the relationships or interactions between these entities. According to Hu et al. (2008), most of the networks used for purposes similar to the one proposed in this research are called complex real-world networks, due to the number of interactions that their entities have. Typically, their interactions are not subject to random rules and, with this, properties such as emergence, evolution (an important indicator of dynamics), among others, can be quantified. In the last decade, this type of study has become an area with significant growth in different disciplines in which the works of Grabska-Gradzińska et al. (2012), Molontay and Nagy (2020), Dorogovtsev and Mendes (2013) and Pastor-Satorras and Vespignani (2007) stand out. They have presented approaches that use complex networks to provide solutions to STI issues that are analytical, less reductionist, more holistic or systemic. In

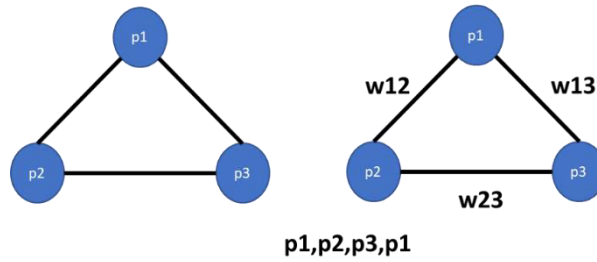
other words, the intrinsic elements of a complex network are useful to understand dynamics and processes. For example, the subgraphs of a network are subsystems or configurations of a complex system, and they generate quantitative evaluations useful for decision-making when they are complemented with physics and statistics tools.

Language is one of the most wonderful tools developed by humans (Pinker, 2003; Bickerton, 2009) and recognizing emerging elements when building a model is a challenge that must be undertaken through complex networks, given the quality and effectiveness of the results obtained by various studies (Markošová, 2008; Mehler, 2008; Choudhury & Mukherjee, 2009; Borge-Holthoefer & Arenas; Solé et al., 2010; Baronchelli et al., 2013; Mihalcea & Radev, 2011; Biemann, 2012, Ferrer-i-Cancho 2005). In addition to these contributions, it is also helpful to keep in mind the definition of language provided by De Saussure, considered the father of modern linguistics (Palop & Vicente, 1999). According to him language is a system in which each linguistic unit is defined by, and only by its relationships (Cong et al., 2014). This statement shows the usefulness of complex networks for the study of human language, especially when scientific results are popularized.

The definition of Beckner et al. (2009) is typically used to build complex networks using information collected from scientific popularization media: $N = (V, E)$, in which N is the information network from the scientific articles composed by the entities or vertices that belong to V , and E is the set of edges or interactions that account for the relationships between the entities that comprised the article or scientific work. For the analysis of STI that is required for this type of study, Hu et al. (2008) and Estrada (2012) argue that the linguistic units to be considered should be words under the principle of co-occurrence. Co-occurrence networks (also known as collocation networks) are pairs of linguistic units in certain windows in the text that have a relationship of some kind, such as hierarchy, temporality, sequentiality, as well as other non-real elements (named this way because they are interaction factors defined by rules or models external to the text such as statistical metrics, mathematics, among others).

In this research, the connectivity of the words will be determined by the sequentiality they have in the documents included in the textual databases described in the materials section. This type of connectivity was chosen keeping in mind the results of multiple works (Ferrer-i-Cancho & Solé, 2001; Masucci & Rodgers, 2006; Zhou et al., 2008; Shi et al., 2008; Brede & Newth, 2008; Sheng & Li, 2009; Liang et al., 2009; Grabska-Gradzińska et al., 2012; Liang et al., 2012; Gao et al., 2014), in which it was possible to identify, using the topological measures of networks, key aspects of the discourse, such as central topics, research intentions, trends, among others. Those key aspects allowed the authors to learn the patterns of scientific publications in social communities around specific topics. For this reason, the co-occurrence network under the principle of sequentiality used for this research is presented in Figure 2.

Figure 2. Principle of co-occurrence network construction for the analysis of discourse and human language.



Source: own elaboration.

In figure 2, p1, p2 and p3 refer to the words obtained from the keywords found in each of the selected articles and the documents referenced in the public policies related to STI in Colombia (materials section), and the edges refer to the sequential co-occurrence of the words in those sources of information. The frequency in which these words are found in textual databases are represented by w_{12} , w_{13} and w_{23} .

Once these networks were created (some generated from the scientific articles and others from the documents related to STI referenced in public policies in Colombia), the next step was to measure their properties and the intrinsic aspects that facilitate the emergence of nodes, that is, the trends evidenced by the generation of new knowledge in Colombia. In accordance with the objectives of this research, the following metrics were selected:

- Density
- Degree
- Assortativity
- Modularity

3.1 Density (Δ)

$$\Delta = \frac{2L}{n(n-1)}$$

- Capacity for interaction among words, terms and concepts in the articles and documents related to STI referenced in public policies in Colombia.
- Its dimension is between zero and one. The closer it gets to one, the more interaction the terms or concepts used in the document or discourse have.
- L is the number of links and n is the number of nodes in the network, therefore $n(n-1)$ is the maximum number of links that can exist in the network.
- For this research, density will make it possible to assess the interaction or association of the words used by Colombian scientists in their academic production and the needs raised in the country by the Colombian science and technology regulatory bodies.

3.2 Degree

- Number of interactions the words or concepts related to STI in Colombian articles and policy documents have.
- The higher the degree, the more important or frequent the word is.

- The value of the degree in this research will allow evaluating the centrality of the word in the texts that belong to the analysis category. Thus, to define the fundamental pillars of each unit of analysis, the words with the highest degree in the network must be selected.

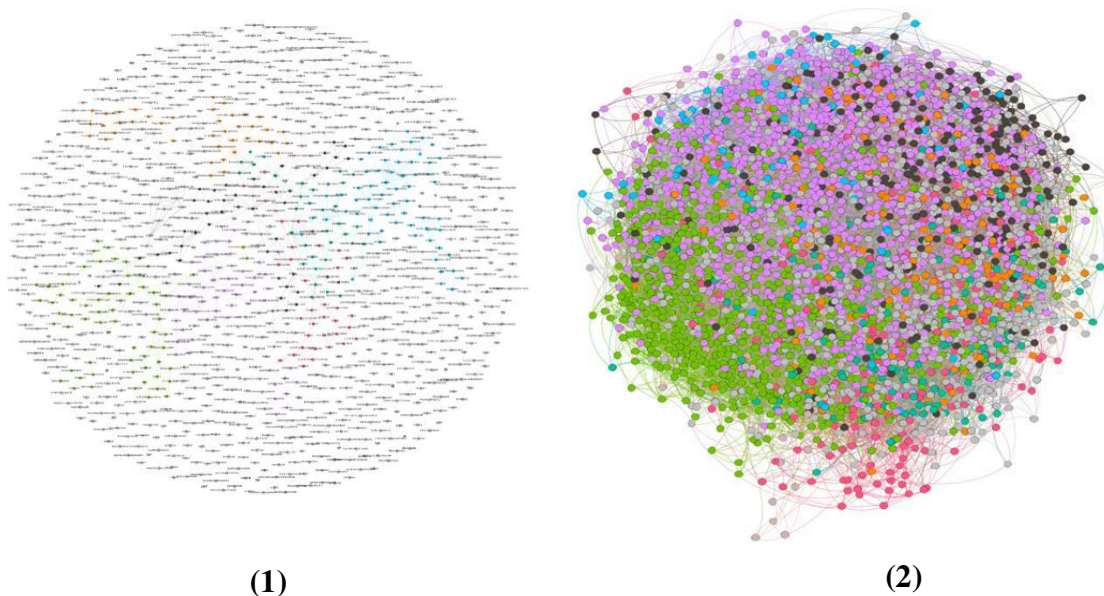
3.3 Modularity

Modularity consists of dividing the network into subgraphs or groups of nodes called modules (or communities) which have strong interactions between them. The technique used to modulate the network was the modularity maximization algorithm (Newman, 2006), which seeks to obtain the maximum and most significant number of modules into which a network can be divided. From the systemic perspective, these subgraphs or modules are subsystems formed in the articles and the documents related to STI referenced in public policies in Colombia. Therefore, these formations are the units by which the most important words will be detected, in accordance with the proportion of nodes of that module compared to the entire network.

4. Results and Discussion of Results

Complex networks were built using the scientific production of articles and the documents related to STI referenced in public policies in Colombia and are presented in Figure 3.

Figure 3. Co-occurrence networks of words and concepts found in Colombian scientific articles (1) and STI base documents (2).



Source: own elaboration.

One of the first differences between the two networks built was size. This topological metric revealed the number of words used in scientific discourse and by government entities that promote science resources and public policies. In the scientific articles 2623 words

were found (Figure 3 (1)), while 9681 (Figure 3 (2)) were found in the documents referenced in the public policies. Out of the total of terms found, 2927 are strongly connected. Despite this similarity, density (Δ) reflects something different: the density of the network of the scientific articles is $\Delta_1=0.048$ and the density of the network of the documents referenced in public policies is $\Delta_2=0.001$. This result reflects that the words used in the scientific articles have more interaction with each other, compared to the terms in the documents referenced in the public policies. This finding is relevant since, even though the policies reflect that there are topics that are decisive to direct science in favor of meeting the country's needs, their words have little articulation. However, Δ_1 has a low value if the density range is considered. Thus, despite its difference with Δ_2 (the issue of little interactivity or synchronization of scientific topics developed in the country), it is necessary to detail with local measures the identification of the main topics of the network, even though the interaction of the global network is low. Keeping in mind the above, the assessment of the nodal degree was conducted and presented in tables 1 and 2.

Table 1. Nodal degree of the network of scientific articles published in Scopus between 2016-2021.

Word	Degree
Development	560
Sectors	310
Resources	306
Knowledge	234
Sustainable	214
Growth	213
Innovation	202
Green	189
Business	178
Biodiversity	176
Products	174
Processes	173
Energy	168
Services	148
Capacity	122
Biotechnology	122
Productivity	108
Access	108
Technology	108
Transport	108

Source: own elaboration.

Table 2. Nodal degree of the network of documents referenced in STI public policies.

Word	Degree
Power	6077
Electric	6070
Distribution	4876
Networks, wind	3475
Power, electric	3475
Transmission	3475
Quantum	1966
Computer	1651
Health	1572
Converters	1536
Inverters, dc-dc	1536
Potential, electric	1536
Fuel	1266
Frequency	1251
Systems	1201
Supply	1196
Cells	1157
Phase	1104
Agents, infection, methicillin-resistant	1074
Aureus	1074

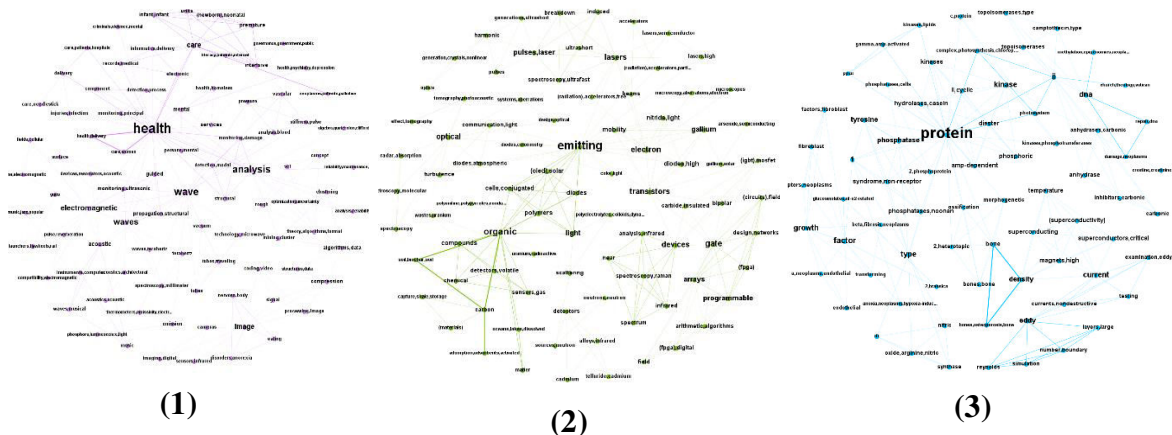
Source: own elaboration.

Tables 1 and 2 show the limited convergence of the main topics contained in the published research articles and the topics of the documents referenced in public policies for the promotion of science in the country. This divergence between the scientific research conducted by the stakeholders of the country's National Research System and the challenges demanded by the regions in terms of strategies to improve competitiveness has a profound impact on the operationalization of science to face such challenges. In other words, the research conducted in Colombia is focused on trying to address external needs that are more frequently visible in the high-impact popularization media, but do not cover the real challenges of the country. This approach evidences a systemic failure of the Government to articulate stakeholders in scientific research, resources, and policies.

Finally, this divergence is not only evident in the central topics of the reviewed sources, but also found in the driving factors of each one of them. In other words, although words such as development and growth are mentioned, the connectors are much more isolated than the strategic claims of the local and central government of the country. For this reason,

the analysis of larger communities in the network was performed using the modularity method defined in section 3.3. This analysis is presented in Figure 4.

Figure 4. Communities detected in the network of scientific articles published in Scopus between 2016-2021.



Source: own elaboration.

The main node of the first community is health since the word interacts with specific components that are important for its materialization. These relationships make the design of instruments for health care and the monitoring of variables evident through different quantitative modeling and simulation techniques. Despite the importance of such relationships, none of them is related to the diseases with the highest morbidity and mortality in the most vulnerable regions of the country. The main topic of the second community is particle physics, as well as other terms related to solid state physics, which evidences the focus of research on important topics for academia and the basic sciences. However, such topics are not relevant in the development plans of the country and its most vulnerable regions. The topics of the third community are related to biology, in particular to the “omics” areas (genomics, proteomics, metabolic), and have a very academic approach, little related to the needs identified in the documents referenced in the policies for STI developed in Colombia selected for this research.

5. Conclusion

Despite the fact that in recent years scientific production in Colombia has notably increased in almost all areas of knowledge, its topics and results have no impact on the country's main challenges. Even the relationships of the main nodes identified in the government's plans and agendas do not have significant depth, which makes the lack of synchronization between scientific research and public policies greater and less useful when it comes to solidifying a model of STI for the country.

On the other hand, building complex networks allowed us to model and represent complex systems mixing different sources of information related to scientific popularization. Likewise, it allowed us to discover patterns in the interactions and topological metrics of these information sources. In them, the predominant topics are basic

sciences, computing, and statistical analysis, which are in accordance with academic trends but are not useful to solve the challenges and problems of the different territories of the country.

6. Referencias

- Baronchelli, A., Ferrer-i-Cancho, R., Pastor-Satorras, R., Chater, N., & Christiansen, M. H. (2013). Networks in Cognitive Science. *Trends in Cognitive Science*, 17(7), 348-360. <https://doi.org/10.1016/j.tics.2013.04.010>
- Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., Ellis, N. C., Holland, J., Ke, J., Larsen-Freeman, D., Schoenemann, T., & The Five Graces Group. (2009). Language is a Complex Adaptive System: Position Paper. *Language Learning*, 59(Supl. 1), 1-26. <https://doi.org/10.1111/j.1467-9922.2009.00533.x>
- Bickerton, D. (2009). *Adam's Tongue: How Humans Made Language, How Language Made Humans*. Macmillan.
- Biemann, C. (2012). *Structure Discovery in Natural Language*. Springer. <http://doi.org/10.1007/978-3-642-25923-4>
- Borge-Holthoefer, J., & Arenas, A. (2010). Semantic Networks: Structure and Dynamics. *Entropy*, 12, 1264-1302. <https://doi.org/10.3390/e12051264>
- Brede, M., & Newth, D. (2008). Patterns in Syntactic Dependency Networks from Authored and Randomised Texts. *Complexity International*, 12, msid23. <https://core.ac.uk/download/pdf/1512368.pdf>
- Choudhury, M., & Mukherjee, A. (2009). The Structure and Dynamics of Linguistic Networks. En N. Ganguly, A. Deutsch & A. Mukherjee (Eds.), *Dynamics on and of Complex Networks: Applications to Biology, Computer Science, and the Social Sciences* (pp. 145-166). Birkhäuser. <http://doi.org/10.1007/978-0-8176-4751-3>
- Cong, J., & Liu, H. (2014). Approaching human language with complex networks. *Physics of life reviews*, 11(4), 598-618.
- Dorogovtsev, S. N., & Mendes, J. F. (2013). *Evolution of Networks: From Biological Nets to the Internet and WWW*. OUP Oxford. <http://doi.org/10.1093/acprof:oso/9780198515906.001.0001>
- Estrada, E. (2012). *The Structure of Complex Networks: Theory and Applications*. Oxford University Press.
- Ferrer-i-Cancho, R. F. (2005). The Structure of Syntactic Dependency Networks: Insights from Recent Advances in Network Theory. *Problems of Quantitative Linguistics*, 2005, 60-75.
- Ferrer-i-Cancho, R., & Solé, R. V. (2001). The Small World of Human Language. *Proceedings of the Royal Society, Biological Sciences*, 268(1482), 2261-2265. <https://dx.doi.org/10.1098/rspb.2001.1800>

- Gao, Y., Liang, W., Shi, Y., & Huang, Q. (2014). Comparison of Directed and Weighted Co-occurrence Networks of Six Languages. *Physica A: Statistical Mechanics and its Applications*, 393, 578-589. <https://doi.org/10.1016/j.physa.2013.08.075>
- Grabska-Gradzińska, I., Kulig, A., Kwapien, J., & Drożdż, S. (2012). Complex Network Analysis of Literary and Scientific Texts. *International Journal of Modern Physics C*, 23(07), 1250051. <https://doi.org/10.1142/S0129183112500519>
- Grabska-Gradzińska, I., Kulig, A., Kwapien, J., Drożdż, S. (2012). Complex Network Analysis of Literary and Scientific Texts. *International Journal of Modern Physics C*, 23(7), 1250051. <https://doi.org/10.1142/S0129183112500519>
- Hu, X., Yang, J. M., & Li, D. R. (2008). The Complex Network Analysis of the Enterprise Competitive Relationships Evolution-Take Software Industry in Guangdong Province as the Example. *Soft Science*, 6, 52-56.
- Kash, D. E., & Rycroft, R. W. (2000). Patterns of Innovating Complex Technologies: A Framework for Adaptive Network Strategies. *Research Policy*, 29(7), 819-831. [https://doi.org/10.1016/S0048-7333\(00\)00107-4](https://doi.org/10.1016/S0048-7333(00)00107-4)
- León, A. M. (2006). Valoración, selección y pertinencia de herramientas de software utilizadas en vigilancia tecnológica. *Ingeniería e Investigación*, 26(1), 92-102. <http://www.scielo.org.co/pdf/iei/v26n1/v26n1a12.pdf>
- Liang, W., Shi, Y., Tse, C. K., & Wang, Y. (2012). Study on Co-occurrence Character Networks from Chinese Essays in Different Periods. *Science China Information Sciences*, 55(11), 2417-2427. <https://doi.org/10.1007/s11432-011-4438-x>
- Liang, W., Shi, Y., Tse, C. K., Liu, J., Wang, Y., & Cui, X. (2009). Comparison of Co-occurrence Networks of the Chinese and English Languages. *Physica A: Statistical Mechanics and its Applications*, 388(23), 4901-4909. <https://doi.org/10.1016/j.physa.2009.07.047>
- Liu, X., Yu, Y., Guo, C., & Sun, Y. (2014). Meta-Path-Based Ranking with Pseudo Relevance Feedback on Heterogeneous Graph for Citation Recommendation. *CIKM 2014 - Proceedings of the 23rd ACM International Conference on Information and Knowledge Management*, 121-130. <https://doi.org/10.1145/2661829.2661965>
- Liu, X., Yu, Y., Guo, C., Sun, Y., & Gao, L. (2014). Full-Text Based Context-Rich Heterogeneous Network Mining Approach for Citation Recommendation. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, September, 361-370. <https://doi.org/10.1109/JCDL.2014.6970191>
- Liu, X., Zhang, J., & Guo, C. (2013). Full-Text Citation Analysis : A New Method to Enhance. *Journal of the American Society for Information Science and Technology*, 64(9), 1852-1863. <https://doi.org/10.1002/asi.22883>
- Markošová, M. (2008). Network Model of Human Language. *Physica A: Statistical Mechanics and its Applications*, 387(2-3), 661-666. <https://doi.org/10.1016/j.physa.2007.09.027>

- Masucci, A. P., & Rodgers, G. J. (2006). Network Properties of Written Human Language. *Physical Review E, Covering Statistical, Nonlinear, Biological, and Soft Matter Physics*, 74(2), 026102. <https://doi.org/10.1103/PhysRevE.74.026102>
- Mehler, A. (2008). Large Text Networks as an Object of Corpus Linguistic Studies. En A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (vol. 1, pp. 328-382). Walter de Gruyter.
- Mihalcea, R., & Radev, D. (2011). *Graph-Based Natural Language Processing and Information Retrieval*. Cambridge University.
- Molontay, R., & Nagy, M. (2020). Twenty Years of Network Science: A Bibliographic and Co-authorship Network Analysis. arXiv preprint arXiv:2001.09006. <https://arxiv.org/abs/2001.09006>
- Newman, M. (2006). Modularity and Community Structure in Networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), 8577-8696. <https://doi.org/10.1073/pnas.0601602103>
- Palop, F., & Vicente, J. M. (1999). *Vigilancia tecnológica*. Fundación COTEC para la innovación tecnológica.
- Pastor-Satorras, R., & Vespignani, A. (2007). *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511610905>
- Pinker, S. (2003). *The Language Instinct: How the Mind Creates Language*. Penguin UK.
- Sheng, L., & Li, C. (2009). English and Chinese Languages as Weighted Complex Networks. *Physica A: Statistical Mechanics and its Applications*, 388(12), 2561-2570. <https://doi.org/10.1016/j.physa.2009.02.043>
- Shi, Y., Liang, W., Liu, J., & Tse, C. K. (2008). Structural Equivalence between Co-occurrences of Characters and Words in the Chinese Language. *International Symposium on Nonlinear Theory and its Applications* (pp. 94-97). <https://doi.org/10.34385/proc.42.A2L-B5>
- Solé, R. V., Corominas-Murtra, B., Valverde, S., & Steels, L. (2010). Language Networks: Their Structure, Function and Evolution. *Complexity*, 15(6), 20-26. <https://doi.org/10.1002/cplx.20305>
- Zhou, S., Hu, G., Zhang, Z., & Guan, J. (2008). An Empirical Study of Chinese Language Networks. *Physica A: Statistical Mechanics and its Applications*, 387(12), 3039-3047. <https://doi.org/10.1016/j.physa.2008.01.024>
- Zhu, W., & Guan, J. (2013). A bibliometric study of service innovation research: based on complex network analysis. *Scientometrics*, 94(3), 1195-1216. <https://doi.org/10.1007/s11192-012-0888-1>