Assessing a country's scientific contribution towards sustainability from higher education: a methodology for measuring progress towards the Sustainable Development Goals (SDG)

Olga Lucia Ostos Ortiz[*]
Oscar Yecid Aparicio-Gómez[**]
Otto Federico von Feigenblatt[***]

Recibido: 02-08-2022
Aceptado: 11-09-2022

Citar como: Ostos Ortiz O. L., Aparicio-Gómez, O. Y. y von Feigenblatt, O. F. (2023). Assessing a country's scientific contribution towards sustainability from higher education: a methodology for measuring progress towards the Sustainable Development Goals (SDG). *Revista Interamericana de Investigación, Educación y Pedagogía, 16*(2), 343-361. https://doi.org/10.15332/25005421.8848

Abstract

Objective. To develop a methodology to evaluate the country's scientific production regarding the fulfillment of the Sustainable Development Goals. *Methodology.* A data set of national and international repositories of science and technology where the scientific production of Colombia is housed was configured; these repositories are recognized for housing products related to the topics of the Sustainable Development Goals. Complex network analysis and indicators such as

[*]Doctora en Modelado en Política y Gestión Pública. Directora de Investigación, Universidad Santo Tomás.
Correo electrónico: olgaostosortiz@gmail.com
ORCID: https://orcid.org/0000-0002-6477-9872
Google Scholar: https://scholar.google.com/citations?user=yCBpLUsAAAAJ&hl=es

[**]Doctor en Filosofía y Doctor en Educación. Director, Ed&TIC Research Center.
Correo electrónico: oaparicio@editic.net
ORCID: https://orcid.org/0000-0003-3535-6288
Google Scholar: https://scholar.google.com/citations?hl=es&user=kH00EsUAAAAJ

[***]Doctor en Filosofía y Doctor en Educación. Profesor, Keiser University - Florida, Estados Unidos de America.
Correo electrónico: vonFeigenblatt@hotmail.com
ORCID: https://orcid.org/0000-0001-6033-6495
Google Scholar: https://scholar.google.com/citations?hl=es&user=G84X8-sAAAAJ

perplexity and coherence were used as methodology. *Results.* Among the most important results, the synchronization of scientific products related to the «End Poverty», «Zero Hunger» and «Health and Well-being» goals stands out. The methodology used as a tool to analyze the scientific production of a country regarding the fulfillment of the sustainable development goals is highlighted.

Keywords: technology, scientific research, innovation, methodology, sustainable development goals.

Evaluar la contribución científica de un país hacia la sostenibilidad desde la educación superior: Una metodología para medir el progreso hacia los Objetivos de Desarrollo Sostenible (ODS)

Resumen

Objetivo. Desarrollar una metodología para evaluar la producción científica del país en relación con el cumplimiento de los Objetivos de Desarrollo Sostenible. *Metodología.* Se configuró un conjunto de datos de repositorios nacionales e internacionales de ciencia y tecnología donde se aloja la producción científica de Colombia; estos repositorios son reconocidos por alojar productos relacionados con las temáticas de los Objetivos de Desarrollo Sostenible. Como metodología se utilizó el análisis de redes complejas e indicadores como perplejidad y coherencia. *Resultados.* Entre los resultados más importantes destaca la sincronización de productos científicos relacionados con los objetivos «Fin de la Pobreza», «Hambre Cero» y «Salud y Bienestar». Se destaca la metodología utilizada como herramienta para analizar la producción científica de un país en relación con el cumplimiento de los objetivos de desarrollo sostenible.

Palabras clave: tecnología, investigación científica, innovación, metodología, objetivos de desarrollo sostenible.

Avaliar a contribuição científica de um país para a sustentabilidade a partir do ensino superior: uma metodologia para medir o progresso em direcção aos Objectivos de Desenvolvimento Sustentável (ODS)

Resumo

Objectivo. Desenvolver uma metodologia para avaliar a produção científica do país em relação ao cumprimento dos Objectivos de Desenvolvimento Sustentável. *Metodologia*. Foi configurado um conjunto de dados de repositórios nacionais e internacionais de ciência e tecnologia onde se aloja a produção científica da Colômbia, reconhecidos por alojar produtos relacionados com os temas dos Objectivos de Desenvolvimento Sustentável. A metodologia utilizada foi a análise de redes complexas e indicadores como a perplexidade e a coerência. *Resultados*. Entre os resultados mais importantes, destaca-se a sincronização de produtos científicos relacionados com os objectivos «Acabar com a Pobreza», «Fome Zero» e «Saúde e Bem-estar». Destaca-se a metodologia utilizada como ferramenta para analisar a produção científica de um país em relação ao cumprimento dos Objectivos de Desenvolvimento Sustentável.

Palavras-chave: tecnologia, investigação científica, inovação, metodologia, objectivos de desenvolvimento sustentável.

DOI: https://doi.org/10.15332/25005421.8848

ISSN: 1657-107X
e-ISSN: 2500-5421

Introduction

The 2030 Agenda for Sustainable Development, through its 17 goals (SDGs) and 169 targets, has become a global commitment of a cross-cutting and multidimensional nature and constitutes the guide for the pursuit of sustainability, from economic, social, human, environmental and governance perspectives, in which science, technology and innovation play a fundamental role (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2015). Thus, a country's Science, Technology and Innovation (STI) plans and/or activities become one of the main strategies for the materialization of sustainable development objectives.

The analysis of the relationship between science, technology and innovation and the achievement of sustainable development objectives must consider the heterogeneity of countries, since their characteristics and conditions differ significantly and extend to the social and scientific-technical spheres. At the same time, the maturity of innovation systems and the degree of development of research processes are divergent (Instituto Complutense de Estudios Internacionales [ICEI], 2020).

The Observatory of Science, Technology and Society, coordinated by the Organization of Ibero-American States (OEI) and the Network of Science and Technology Indicators (RICYT), points out that 57 % of Ibero-American researchers carry out their activity in universities, 28 % in the business sector and 14 % in public institutions. Consequently, the axis of research and academic production continues to be concentrated in universities (ICEI, 2020).

Figure 1. Sustainable development goals

Source: https://www.un.org/development/desa/disabilities/about-us/sustainable-development-goals-sdgs-and-disability.html

Materials and methods

Materials

A set of data generated from databases and repositories of science, technology, research and innovation related to Colombian scientific productivity, case study, was constructed. Several factors were included in the evaluation of the impact of research products on the materialization of sustainable development objectives, such as:

- The generation of new knowledge published in scientific journals indexed in Scopus and Web of Science.

- Scientific products generated because of research calls from the Colombian Ministry of Science and Technology within the framework of the Horizon 2030 cooperation and financing fund.

DOI: https://doi.org/10.15332/25005421.8848

ISSN: 1657-107X
e-ISSN: 2500-5421

- Products with highly significant results and with scalability in the different territorial contexts of the country, according to the needs established by the National Council of Economic and Social Policy in the framework of the sustainable development objectives.

- References of the Science, Technology and Innovation policy and development plans of the Colombian departments and municipalities.

A seven-year dataset (2014-2021) was constructed with the following characteristics:

- 4568 books as research output.

- 2500 book chapters

- 9566 articles

- 350 documents related to policies, decrees, resolutions, national and regional development plans, as well as reference documents for research and funding calls from the Colombian Ministry of Science and Technology.

Methods

One of the most important characteristics of the dataset used for this research is that all the information is in text format, and we sought to guarantee the quality of the data and the construction of a robust analysis model that would allow us to obtain significant results. The methodology implemented used data in text format: content of books, chapters, journals, working papers, public policies, guiding documents of the Sustainable Development Goals (SDGs),

etc. The process of improving the quality of the data used in this research is based on a text format.

The process of improving the quality of the data consisted of carrying out activities to clean and select the most important information. For this purpose, the proposal of Rentería *et al.* (2022) will be used, where the following algorithmic sequence was established:

- Tokenization. This is a process by which the text is segmented into words, i.e., a textual corpus is transformed into a set or list of words. For example, in the text "National System of Science and Technology", its tokenization is: "System", "National", "Science" and "Technology".

- Lemmatization. It is used for the dimensional reduction of a word or phrase by transforming it to its base and root form.

- Recognition and debugging. This is a stage where the main and important words are selected, based on a heuristic algorithm called tabu list, which consists of discarding from the document those words that do not provide important information for the understanding of the text. Within this debugging, terms such as: "the", and "they", (also known as stop words) are eliminated.

Once the textual corpus was obtained, the process of evaluating the configurations of the most important terms in each of the scientific products began, determining their co-occurrence, sequentially and location.

Subsequently, a technique was selected from the existing ones for topic modeling using Bayesian statistics; among the available techniques, the following stand out: Latent Semantic Indexing (LSA) (Zhao *et al.*, 2015) the Probabilistic Latent Semantic Analysis (PLSA) (Hofmann 2001, 1999), and Latent Dirichlet Allocation (LDA) (Blei *et al.*,

DOI: https://doi.org/10.15332/25005421.8848
ISSN: 1657-107X
e-ISSN: 2500-5421

2015). According to Duran *et al.* (2015), LDA is one of the techniques that has generated more significant contributions to the study from different areas of knowledge: social sciences (Ramage *et al.*, 2009; Li and Lei, 2021; Jacobi *et al.*, 2016), economics (Hong *et al.*, 2016), health (Paul and Dredze 2013; 2014) among others.

For this research, the selected approach was the LDA because the topics constructed with this model are highly coherent (Stevens *et al.*, 2012) and, therefore, it is a useful tool to assess the relevance of the results obtained from Colombian scientific productivity, related to the challenges established in the SDGs.

For the construction of the LDA algorithm, the proposal of Duran *et al.* (2015) was used, where given a textual corpus D composed of M documents, with a document d with N_d words $(d \in \{1\ldots., M\})$, the construction of topics was done as follows:

1. Select a multinomial distribution ϕ_t for the topic $(t \in \{1\ldots., T\})$, from a Dirichlet distribution with parameter β.

2. Select a multinominal distribution θ_d for $(d \in \{1\ldots., M\})$, from a Dirichlet distribution with parameter α

3. For a word w_n $(n \in \{1\ldots., N_d\})$ in a document d,

i. Select a topic z_n from θ_d

ii. Select a word w_d from ϕ_{zn}

One of the most important aspects for this algorithm are the variables ϕ and θ, and the hyperparameters α and β that, as recommended by (Duran *et al.*, 2015) must be defined from the perspective of latent variables. Therefore, under this principle for the estimation of the parameters in the textual corpus D.

$$P(D|\alpha,\beta)=\prod_{d=1}^{M}\int P(\theta_d|\alpha)\left(\sum_{n=1}^{N_d}P(z_{dn}|\,\theta_d)\,P(w_{dn}|\,z_{dn},\varphi)P(\varphi|\beta)\right)d\theta_d d\varphi \quad (1)$$

As the parameters θ and ϕ are part of the integration, it is necessary to build simulation models to establish the inference of these values, according to (Blei *et al.*, 2003; Rogers *et al.*, 2005; Griffiths and Steyvers, 2004; Shivashankar *et al.*, 2011; Coelho *et al.*, 2010) the best computational alternative for this inference is Markov Chain Monte Carlo. Once the topics have been generated, it is necessary to establish a quality metric for the results obtained, which means, to define the optimal number of topics required to adequately classify all the words of the textual corpus D. In this sense, the most appropriate indicator is Perplexity, which is a measure of information theory, to evaluate how well a statistical model describes a data set (Duran *et al.*, 2015), therefore, the lower the quantification of Perplexity, the quality of the results obtained with the model is more reliable. The Perplexity equation (ς) that will be used in this research is the one proposed by (Duran *et al.*, 2015; Blei *et al.*, 2003):

In addition to Perplexity, the coherence indicator (ϑ) is adhered to evaluate the consistency of the words that make up each topic ϕ_i, ϕ_j (see equation 3), to later make a paired comparison based on the Kullback divergence metrics – Leiber (DKL) to establish the main similarities and differences from the distributions of the words in the topics and between the SDGs (Mei *et al.*, 2007).

$$\vartheta = log\frac{p(\phi_i,\phi_j)+\epsilon}{p(\phi_i)p(\phi_j)} \quad (3)$$

DOI: https://doi.org/10.15332/25005421.8848

ISSN: 1657-107X
e-ISSN: 2500-5421

Results

One of the first results that must be calculated in this research is the value of α and β, which are adjusted with a symmetric distribution $\left(\frac{1}{T}\right)$ therefore, it is important to define the optimal number of topics (T) of the LDA model, from perplexity using (2), for a randomly selected quantity t (see figure 2).

Figure 2. Assessment of perplexity according to the number of topics

Source: self-made

For this case, the number of categories or topics with which it can be modeled in a robust and statistically significant way is six because it is where ς is minimum, a value that is repeated in all the SDGs analyzed in this research. Once this value is obtained, the next procedure is to calculate the value of ϑ, for each ODS, as presented in Table 1.

Table 1. Levels of consistency of the topics for each SDG

SDG	Topics	Coherence	DKL
1	6	0.48	5.10
2	6	0.47	6.10
3	6	0.45	4.98
4	6	0.14	10.01
5	6	0.11	20.02
6	6	0.09	25.05
7	6	0.08	35.02
8	6	0.09	10.05
9	6	0.11	8.02
10	6	0.12	7.85

Source: self-made.

According to the results presented in Table 1, the first three SDGs, in addition to having the highest coherence values, show the lowest values of divergence with the fundamental principles of the United Nations declaration, and for such reason, they will be the prioritized elements to evaluate in this investigation.

SDG 1. No to poverty

Regarding this SDG, it was important to consider aspects included in the multidimensional poverty models proposed by Alkire *et al.* (2015), who calculated poverty from a synthetic indicator, using a weighting system based on the importance of each dimension in the configuration of deprivations antagonistic to well-being.

The findings of this SDG highlight the presence of important topics to address a set of solutions to the problems underlying the term, including terms such as "development", "economic" and "health", which are decisive aspects to solve the needs of poverty.

Therefore, in addition to incorporating the aforementioned factors, it is necessary to involve aspects related to habitat, and other

DOI: https://doi.org/10.15332/25005421.8848
ISSN: 1657-107X
e-ISSN: 2500-5421

factors that promote social wellbeing, such as access to quality public services, health services, among others.

This lack is reflected even in the smaller topics (smaller number of terms), demonstrating that the country's scientific development is not generating the evidence to solve poverty problems as established by Ramírez *et al.* (2017), Pinilla-Roncancio (2018) and Manzano-Núñez *et al.* (2022). Finally, Cuesta and Pico (2020), mention that these poverty and inequality gaps have become more acute due to the pandemic.

SDG 2. Zero hunger

In the case analyzed, in the sustainable development goal "Zero Hunger", scientific development has generated important results in aspects such as: "Agriculture", "Climate Change", "Economy", "Social Capital" and "Bioeconomy", which according to Khanal *et al.* (2021), are the keys to promote strategies to reduce hunger. "Social Capital" and "Bioeconomy" that according to Khanal *et al.* (2021), are the keys to promoting strategies to reduce hunger because the alterations generated because of climate change affect the availability and quality of the product, reducing the presence of some nutrients, proteins and vitamins required for health welfare.

Blesh *et al.* (2019), mentions that to ensure the construction of a sustainable model of zero hunger, in addition to a good food security system in the country, the support of a good health system, highly inclusive and based on the approach of prioritization of care from the social determinants of well-being, should be considered.

On the other hand, Sunderland *et al.* (2019), highlights that factors such as adaptation and resilience to climate change should be on the government's agenda to achieve the expected results and impacts driven by the guarantee of zero hunger in the population. Although these aspects are part of the most important issues (based on the

proportion of terms and main components since the creation of the objective) and coherent, the intra-theme analysis shows the lack of a homogeneous model that guarantees its adaptability to the territorial needs of the country, since most of the strategies implemented are focused on the country's capital cities, and not on isolated rural contexts and territories.

SDG 3. Good health and well-being

Health and well-being is another of the objectives that has a series of initiatives that have been developed in the country involving plans and programs supported by the Model of Comprehensive Health Care (Hernández *et al.*, 2019), and whose key aspects are found in the topics with the highest proportion of objective terms ("Health", "Quality-Health", "Care-Differentiated", "Coverage-Integral").

The study strongly suggested that the model established prioritizes the insurance of the population in the General Social Security Health System (SGSS), from an inclusive and differentiated approach, considering the socioeconomic gaps of the different population groups in the territory. In this way, the country has achieved health coverage of over 90% of the population and has established a population screening model considering the vulnerability approaches proposed by Diderichsen *et al.* (2019), which facilitates access to specialized health services.

Such contribution is reflected in the identification of the most relevant topics: "health", "care", "sanitary", "humanization", "risk", "inclusion", "social class" and "stratification". However, researchers such as Roncancio *et al.* (2020), mention that the development of this SDG in the country is not only achieved by guaranteeing the health system for the entire population, but it is also necessary to improve the incorporation of the social determinants of health, since there are sociodemographic and geographic factors that have had an important

DOI: https://doi.org/10.15332/25005421.8848
ISSN: 1657-107X
e-ISSN: 2500-5421

impact on the configuration of vulnerability that exacerbates the complications of some pathologies in population groups (Roncancio *et al.*, 2020 and Diderichsen *et al.*, 2019).

Discussion

One of the greatest challenges in monitoring compliance with the 2030 agenda of the SDGs is to develop robust and rigorous methodologies to measure the contributions of countries in their fulfillment. This work provides a methodology that allows monitoring the contribution of researchers to the SDGs from the analysis of scientific production represented in articles, books, book chapters, working papers, science and technology policy documents, among others.

Thus, the work developed proposes a methodology that allows analyzing the scientific contribution of a country in a given period of time using a technique that uses Bayesian statistics: Latent Direct Allocation (LDA) and text analysis by determining the co-occurrence, sequence and location of terms, identifying the preponderance of a term within a discourse.

The measure of perplexity was calculated, which determines the number of topics best suited to represent a source of information, the degree of coherence that identifies the relationship between terms and the degree of divergence of terms (DKL) between two analyzed discourses, in the case of study; research products, science, technology and innovation policies and the guiding documents of the sustainable development objectives.

The methodology employed functions as a systemic model for evaluating the fulfillment of the SDGs for a country. This dynamic is not only relevant for the Colombian scenario, but the evidence shows that this same behavior is observed in other countries with a

higher level of scientific and technological development. In the case analyzed, which corresponded to the Colombian scientific production of seven years, a greater contribution of researchers to the first 10 sustainable development goals was evidenced. However, the analysis shows that scientific production is more focused on complying with international standards than on the contribution of knowledge and the solution of the country's regional needs. This is explained by the incentives promoted by the National Ministry of Science, Technology and Innovation that focus on integrated production in Scopus and Wos and the creation of financing funds oriented especially to the development of the needs that are part of the first ten SDGs.

Even in the contribution to the SDGs in which there are fewer divergences (evaluated from the DKL), a significant bias towards the country's large cities was evident (research is directed to those territories), excluding the demands of rural contexts and settlements of population groups that have historically been excluded from the main state agendas and plans, whose situation has worsened considerably with the declaration of pandemic generated by COVID-19.

Therefore, it is necessary to promote new indicators of the impact of the scientific production of the SDGs at the country level, based not only on bibliometric compliance or based on international repositories and bibliographic databases, but also on territorial needs and the promotion of instruments for decision-making by local and national governments.

DOI: https://doi.org/10.15332/25005421.8848

ISSN: 1657-107X
e-ISSN: 2500-5421

References

Sabina, A., Foster, J., Seth, S., Santos, A., Roche, J., and Ballon, P. (2015). *Multidimensional poverty measurement and analysis*. Oxford: Oxford University Press, USA. https://ophi.org.uk/mpma-book-2015/

Blei, D., and Jordan, M. (2003). "Modeling annotated data". In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, 127-134. Association for Computing Machinery, New York, USA. https://doi.org/10.1145/860435.860460

Blei, D., Ng, A., and Jordan, M. (2003). "Latent Dirichlet Allocation". *The Journal of Machine Learning Research* 3: 993-1022. https://dl.acm.org/doi/10.5555/944919.944937

Blesh, J., Hoey, L., Jones, A., Friedmann, H., and Perfecto, I. (2019). "Development pathways toward 'zero hunger'". *World Development* 118: 1-14. https://doi.org/10.1016/j.worlddev.2019.02.004

Coelho L., Peng, T., Murphy, R. (2010). "Quantifying the distribution of probes between subcellular locations using unsupervised pattern unmixing". *Bioinformatics*, 26(12): i7-i12. https://doi.org/10.1093/bioinformatics/btq220

Cuesta, J., and Pico, J. (2020). "The gendered poverty effects of the COVID-19 pandemic in Colombia". *The European Journal of Development Research 32*(5): 1558-91. https://doi.org/10.1057/s41287-020-00328-2

Departamento Nacional de Planeación. DNP. (2021). *Documento CONPES 4069. Política Nacional de Ciencia, Tecnología e Innovación 2022-2031*. https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%c3%b3micos/4069.pdf

Diderichsen, F., Hallqvist, J., and Whitehead, M. (2019). "Differential vulnerability and susceptibility: how to make use of recent development in our understanding of mediation and interaction to tackle health inequalities". *International Journal of Epidemiology* 48(1), 268-74. https://doi.org/10.1093/ije/dyy167

Duran, D., Artene, A., Gogan, L., and Duran, V. (2015). "The objectives of sustainable development-ways to achieve welfare". *Procedia Economics and Finance* 26: 812-817. https://doi.org/10.1016/S2212-5671(15)00852-7

Griffiths, T., and Mark, S. (2004). "Finding scientific topics". *The Proceedings of the National Academy of Sciences* 101(Suppl 1):5228-35. https://doi.org/10.1073/pnas.0307752101

Hernández Palma, H., Núñez, W., and Jiménez, M. (2019). "Sistema de salud colombiano: integración para la calidad". *Criterio Libre* 18(31): 149-163. https://doi.org/10.18041/1900-0642/criteriolibre.2019v18n31.6134

Hofmann, T. (1999). "Probabilistic latent semantic indexing". *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. 50-57. https://doi.org/10.1145/312624.312649

Hofmann, T. (2001). "Unsupervised learning by probabilistic latent semantic analysis". *Machine Learning* 42(1-2): 177-196. https://doi.org/10.1023/A:1007617005950

Hong, L., Frias-Martinez, E., and Frias-Martinez, V. (2016). "Topic models to infer socio-economic maps". Thirtieth AAAI Conference on Artificial Intelligence 30(1): 3835-41. https://ojs.aaai.org/index.php/AAAI/article/view/9892.

Instituto Complutense de Estudios Internacionales. (2020). "Ciencia, tecnología e innovación para el cumplimiento de los objetivos de desarrollo sostenible en Iberoamérica". https://www.ucm.es/data/cont/media/www/27289//Relatori%CC%81a-CTI_12enero_2020.pdf

Jacobi, C., Atteveldt, W., and Welbers, K. (2016). "Quantitative analysis of large amounts of journalistic texts using topic modelling". *Digital Journalism*, 4(1): 89-106. https://doi.org/10.1080/21670811.2015.1093271

Khanal, U., Wilson, C., Sotavento, B., and Hoang, V. (2021). "Smallholder farmers' adaptation to climate change and its potential contribution to UN's sustainable development goals of zero hunger and no poverty". *Journal of Cleaner Production* 281, 124999. https://doi.org/10.1016/j.jclepro.2020.124999

Manzano-Nunez, R., Sarmiento, C., Villegas-Vargas, S., Angel-Barrios, J., Puyana, J., Peck, G., Castro, F., Gaviria, A., García, A. (2022). "Emergency surgery workforce and its inverse relationship with multidimensional poverty in Colombia". *European Journal of Trauma and Emergency Surgery* 48(2): 1159-1165. https://doi.org/10.1007/s00068-021-01690-4

Mei, Q., Xu Shen, X., and Zhai, C. (2007). "Automatic labeling of multinomial topic models. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, 490-499. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/1281192.1281246

Paul, M., and Dredze, M. (2013). "Drug extraction from the web: Summarizing drug experiences with multi-dimensional topic models". In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 168-178. Atlanta, Georgia. Association for Computational Linguistics.

Paul, M., and Dredze, M. (2014). "Discovering health topics in social media using topic models". *PloS One 9*(8): e103408. https://doi.org/10.1371/journal.pone.0103408

Pinilla-Roncancio, M. (2018). "The reality of disability: Multidimensional poverty of people with disability and their families in Latin America". *Disability and Health Journal* 11(3): 398-404. https://doi.org/10.1016/j.dhjo.2017.12.007

Ramage, D., Rosen, E., Chuang, J., Manning, C., and McFarland, A. (2009). "Topic modeling for the social sciences". W*orkshop on applications for topic models NIPS* 5: 1-4.

Ramírez, J., Díaz, Y., and Bedoya, J. (2017). "Property tax revenues and multidimensional poverty reduction in Colombia: A spatial approach". *World Development* 94: 406-421. https://doi.org/10.1016/j.worlddev.2017.02.005

Rentería-Ramos, R., Hurtado-Heredia, R., and Urdinola, P. (2019). "Morbi-Mortality of the Victims of Internal Conflict and Poor Population in the Risaralda Province, Colombia". *International Journal of Environmental Research and Public Health* 16(9): 1644. https://doi.org/10.3390/ijerph16091644

Rogers S., Girolami, M., Campbell, C., Breitling R. (2005). "The latent process decomposition of cDNA microarray data sets". *IEEE/ACM transactions on computational biology and bioinformatics* 2(2):143-56.

Roncancio, D., and Nardocci, A. (2020). "Social vulnerability in Colombia". *International Journal of Disaster Risk Reduction 50*, 101872. https://doi.org/10.1016/j.ijdrr.2020.101872

Scimago. (2022). "Scimago Journal and Country Rank, 2019". Last modified April, 2022. https://www.scimagojr.com/countryrank.php?year=2020

Shivashankar, S., Srivathsan, S., Ravindran, B., Tendulkar, A. (2011). "Multi-view methods for protein structure comparison using latent Dirichlet allocation". *Bioinformatics* 27(13): i61-i68. https://doi.org/10.1093/bioinformatics/btr249

Stevens, K., Kegelmeyer, P., Andrzejewski, D., and Buttler, D. (2012). "Exploring topic coherence over many models and many topics". In *Proceedings of the 2012 joint conference on empirical methods in natural language processing and computational natural language learning*, 952-961. Association for Computational Linguistics, USA.

Sunderland, T., O'Connor, A., Muir, G., Nerfa, L., Rota, G., Widmark, C., Bahar, N., and Ickowitz, A. (2019). "Zero hunger: Challenging the hegmony of monoculture agriculture for forests and people". In *Sustainable Development Goals: Their Impacts on Forests and People,* Edited by Katila, P., Pierce, C., Jong, W., Galloway, G., Pacheco, P., and Winkel, G. 48-71. Cambridge: Cambridge University Press.

United Nations Educational, Scientific and Cultural Organization. (2017). *La UNESCO Avanza. La Agenda 2030 para el Desarrollo Sostenible*. Paris: UNESCO. https://es.unesco.org/creativity/sites/creativity/files/247785sp_1_1_1.compressed.pdf.

Uribe-Tirado, A., Vallejo, J., y Betancur, D. (2016). "Somos visibles y tenemos impacto. Análisis desde datos de acceso abierto, altmetrics y otros de la Revista Interamericana de Bibliotecología". *Revista Interamericana de Bibliotecología* 19(3): 243-75

Xin L., and Lei, L. (2021). "A bibliometric analysis of topic modelling studies (2000–2017)". *Journal of Information Science 47*(2): 161-175. https://doi.org/10.1177/0165551519877049

Zhao, W., Chen, J., Perkins, R., Liu, Z., Ge, W., Ding, Y., and Zou, W. (2015). "A heuristic approach to determine an appropriate number of topics in topic modeling". *BMC bioinformatics* 16(S8): 1-10. https://doi.org/10.1186/1471-2105-16-S13-S8

DOI: https://doi.org/10.15332/25005421.8848

ISSN: 1657-107X
e-ISSN: 2500-5421