

UNRAVELLING THE BIOECONOMY - A SEMANTIC SIMILARITY ANALYSIS OF SCIENTIFIC PUBLICATIONS

Daniella F. dos Santos e Silva¹
daniella.fartes@hotmail.com
+55(21)993602871

Bárbara Bressan¹
bbressan@cgee.org.br
+55(61)981998982

Emilly Caroline C. Silva¹
esilva@cgee.org.br
+55(61)982087871

João Pedro F. Arbache¹
jarbache@cgee.org.br
+55(61)999813822

Marcelo Poppe¹
mpoppe@cgee.org.br
+55(61)999740111

¹Center for Strategic Studies and Management (CGEE) - SCS Quadra 9 - torre C – 4th floor, room 401 – 405. Ed. Parque da Cidade Corporate. Brasília. Brazil

ABSTRACT: In the last decade, the term bioeconomy has gained space both in scientific publications and in national strategies and plans as a path to a sustainable, low carbon economy. However, the search for the term in scientific databases still fails to capture its entire field, as it is a new and crosswise theme. Another challenge of studying the bioeconomy at a global level is the influence of each country's regional characteristics on the scope of the bioeconomy. Thus, the objective of this work is to identify which are the main fields of studies considered as part of the bioeconomy, and to analyse how regional characteristics influence the definition of these fields. To achieve this objective, scientific publications of the Web of Science database were studied through semantic similarity methodology using softwares developed by CGEE - InsightNet (iN) and InsightNet Browser (iNB).

Keywords: bioeconomy, biobased economy, semantic similarity analysis

1 CONTEXT

The term bioeconomy has appeared with increasing frequency both in scientific publications and in national strategies and plans. However, the search in scientific bases for the term still fails to capture its entire area of activity (Fig. 1) because it is a new and quite transversal theme. This article considers the following definition of bioeconomy:

“The bioeconomy comprises all economic activity derived from bioprocesses and bioproducts that contributes to efficient solutions in the use of biological resources - in the face of challenges in food, chemicals, materials, energy production, health, environmental services, and environmental protection - that promote the transition to a new model of sustainable development and social well-being” (CGEE, 2020).



Figure 1: Evolution of “bioeconomy” publications.

To create strategies and policies for the development of the bioeconomy, it is first necessary to define its scope. Several countries and international institutions have discussed the main areas of the bioeconomy activity, as well as the most relevant types of raw materials, processes and products [2][3][4][5][6]. However, there are still many discrepancies in the use of the term, both in research and policy debates [7]. In this context, the article aims to raise evidence about how the bioeconomy has been discussed in the academic field.

To access a representative sample of publications in bioeconomy, a methodology was built based on the semantic similarity between the publications extracted from the Web of Science (WoS) database, considering the abstracts, titles and the set of keywords provided by the authors and the journals. This methodology will be described below in section 2. In section 3, the results from the generated network will be presented and discussed in terms of thematic clusters, countries that publish the most, and the Brazilian context of publications. Finally, section 4 will present the conclusions of the work.

2 SNOWBALL METHODOLOGY

The methodology used to create the network of scientific publications on bioeconomy was based on the



Figure 5: Thematic clusters

Table I: Thematic Clusters

Thematic Clusters	
1	Forest biomass
2	Biogas
3	Grass
4	Valorization of waste and coproducts
5	Algae and lipids
6	Lignin
7	Biochar
8	Palm oil
9	Straw
10	Microalgae and wastewater treatment
11	Hydrogen
12	Torrefaction
13	Miscanthus x Giganteus
14	Circular economy
15	Corn biomass

3.2.1 Cluster 1 – Forest Biomass

Cluster focused on analyzing the production and use of forest biomass (whether residual or not) for power generation and product replacement (Fig. 6). The theme of forest biomass is treated in several ways: production methods, transformation processes, technical-economic feasibility analysis, life cycle analysis, land-use implications, public policies, ecosystem analysis, among others.

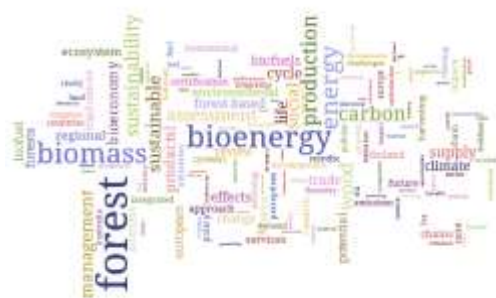


Figure 6: Word cloud from cluster 1

Table II: Top 5 countries in cluster 1

Cluster 1 - Forest	
Countries	Total
Finland	33
USA	31
Sweden	25
Germany	14
Canada	13

3.2.2 Cluster 2 – Biogas

Cluster 2 gathers articles aimed at the production, evaluation, and use of biogas. The central and more condensed part of the cluster presents review articles and articles about biogas assessments, such as advantages and disadvantages for different countries and regions, production potential, and energy generation potential. In the periphery of the cluster, it is possible to identify some specific characteristics, such as the production of biogas from composting, the anaerobic digestion process, and technical viability studies, and life cycle analyses (Fig. 7).



Figure 7: Word cloud from cluster 2

Table III: Top 5 countries in cluster 2

Cluster 2 - Biogas	
Countries	Total
Poland	21
USA	20
Italy	16
Germany	13
China	7

3.2.3 Cluster 3 – Grass

Cluster 3 focuses on energy crops, more specifically grasses, with emphasis on the *Miscanthus* and *Arundo* genera. It is divided into 2 main nuclei, with nodes scattered throughout its neighborhood. The lower nucleus has a large representation of Italian research on grasses, especially of the two genera mentioned above. The upper nucleus also encompasses other cultures, but the object of research in both is quite similar. The main focus is on bioenergy, but other aspects are also addressed, such as gas emissions, productivity, and soil dynamics (Fig. 8).



Figure 8: Word cloud from cluster 3

Table IV: Top 5 countries in cluster 3

Cluster 3 - Grass	
Countries	Total
USA	42
Italy	22
Germany	17
Canada	7
Netherlands	6

3.2.4 Cluster 4 – Valorization of waste and coproducts

Cluster 4 deals mainly with the recovery of waste and coproducts. In its central part, there is a concentration of articles focused on the use of sewage sludge for energy production. However, on the periphery of the cluster, there are groups of nodes that deal with the recovery of other residues and co-products, especially glycerol from the production of biodiesel and lignocellulosic residues (Fig. 9).



Figure 9: Word cloud from cluster 4

Table V: Top 5 countries in cluster 4

Cluster 4 - Valorization of waste and coproducts	
Countries	Total
Poland	12
Italy	8
Canada	7
Sweden	6
Brazil	5

3.2.5 Cluster 5 – Algae and lipids

Cluster 5 is large but more dispersed than those previously mentioned. Most publications are focused on topics related to the use of algae for the production of bioenergy and biofuels and also to lipid extraction

processes from algae and microalgae (Fig. 10).



Figure 10: Word cloud from cluster 5

Table VI: Top 5 countries in cluster 5

Cluster 5 - Algae and lipids	
Countries	Total
USA	15
India	10
China	8
Italy	4
Taiwan	4

3.2.6 Cluster 6 – Lignin

The main core of cluster 6 focuses on research on alternative uses of lignin. There is a rich range of topics, ranging from the production of paper, renewable chemicals, and fuels, use in biorefineries, to the physical, chemical, and biological properties of the material. On the periphery of the cluster, we observe related themes such as cellulose processing (Fig. 11).



Figure 11: Word cloud from cluster 6

Table VII: Top 5 countries in cluster 6

Cluster 6 - Lignin	
Countries	Total
USA	32
Canada	11
China	7
France	5
Brazil	4

3.2.7 Cluster 7 – Biochar

Cluster 7 stands out for being very dense and having all its nodes concentrated in the same space within the network, which represents the approach of the same theme: biochar. Biochar is a carbon-rich material used as a soil corrector to improve its quality. The cluster highlights topics such as the different biochar production

methods: pyrolysis, gasification, and hydrothermal conversion; as well as other topics: environmental and economic impacts, regulatory aspects, and risk analysis (Fig. 12).



Figure 12: Word cloud from cluster 7

Table VIII: Top 5 countries in cluster 7

Cluster 7 - Biochar	
Countries	Total
USA	19
China	8
Canada	7
[China, USA]	5
South Korea	4

3.2.8 Cluster 8 – Palm Oil

Cluster 8 is quite dense and has palm oil as its theme. Studies are mainly divided into reviews, technical and economic analyses, and life cycle analyses. In this cluster, Malaysia stands out as the country with the largest number of publications (Fig. 13).



Figure 13: Word cloud from cluster 8

Table IX: Top 5 countries in cluster 8

Cluster 8 - Palm Oil	
Countries	Total
Malaysia	23
England	23
USA	16
Australia	8
Canada	4

3.2.9 Cluster 9 – Straw

This cluster mainly deals with the use of straw from various agricultural products such as sugar cane, corn, and wheat in the generation of energy and the

transformation into bioproducts. In addition to the work on the use of straw, there is also a strong presence of themes aimed at the availability of this biomass and at assessing soil quality after the collection of straw. Also noteworthy is the presence of Brazil as the country that publishes the most on the topic. Brazilian publications are mainly focused on the use of sugarcane straw (Fig. 14).



Figure 14: Word cloud from cluster 9

Table X: Top 5 countries in cluster 9

Cluster 9 - Straw	
Countries	Total
Brazil	9
USA	8
China	7
England	6
Denmark	4

3.2.10 Cluster 10 – Microalgae and wastewater treatment

Cluster 10 is very cohesive, focusing on the use of microalgae in the treatment of effluents and subsequent transformation into energy and bioproducts. This cluster is very close to cluster 5 (algae and lipids), but the specificity of the topic in question was enough to generate a new modularity class. India stands out as the country that most publishes on the subject, while Brazil holds the fourth position (Fig. 15).



Figure 15: Word cloud from cluster 10

Table XI: Top 5 countries in cluster 10

Cluster 10 - Microalgae and wastewater treatment	
Countries	Total
India	9
USA	8
Malaysia	7

Brazil	6
China	6

3.2.11 Cluster 11 – Hydrogen

This cluster focuses on the production of hydrogen, mainly by biomass gasification, but also by other processes such as anaerobic digestion and fermentation. In addition to the technical publications focused on the hydrogen generation process, many review articles and life cycle analyses were also verified (Fig. 16).



Figure 16: Word cloud from cluster 11

Table XII: Top 5 countries in cluster 11

Cluster 11 - Hydrogen	
Countries	Total
USA	11
Italy	10
Germany	6
Canada	5
China	5

3.2.12 Cluster 12 – Torrefaction

Cluster 12 has as its central theme the biomass torrefaction process. Torrefaction is a thermochemical process that maintains about 75-95% of the energy content. Thus, this cluster deals with this topic, relating it mainly to energy use, the preparation of biomass for gasification processes, and the production of pellets (Fig. 12). The strong participation of the USA and Brazil among the 5 countries that most publish on the subject stands out (Table XIII).

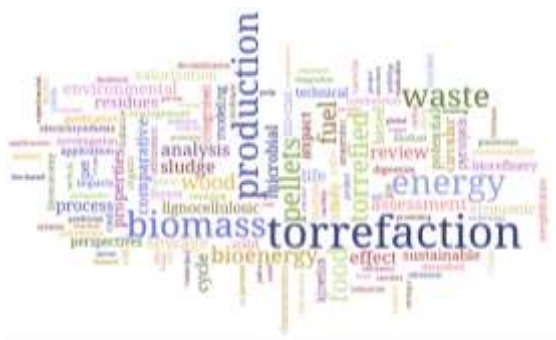


Figure 17: Word cloud from cluster 12

Table XIII: Top 5 countries in cluster 12

Cluster 12 - Torrefaction	
---------------------------	--

Countries	Total
USA	13
India	5
England	4
Italy	4
Brazil	3

3.2.13 Cluster 13 – Miscanthus x Giganteus

The main theme addressed in this cluster is the African and Eurasian plants *miscanthus x giganteus*, which have a large carbon capture capacity. The publications are mainly related to the production of bioenergy and biofuels and the evaluation of the productivity of *miscanthus x giganteus* (Fig. 18).



Figure 18: Word cloud from cluster 13

Table XIV: Top 5 countries in cluster 13

Cluster 13 - Miscanthus x Giganteus	
Countries	Total
USA	19
China	6
Germany	4
Croatia	3
England	3

3.2.14 Cluster 14 – Circular economy

Cluster 14 is strongly characterized by the theme of circular economy and also by “circular bioeconomy”. Within this theme, the publications deal mainly with new business models for bioproducts. It is also worth mentioning a subgroup within the cluster focused on circular economy in the Baltic countries (Fig. 19).



Figure 19: Word cloud from cluster 14

Table XV: Top 5 countries in cluster 14

Cluster 14 - Circular economy	
Countries	Total
Finland	9
Germany	5
Netherlands	5
India	4
Spain	4

3.2.15 Cluster 15 – Corn biomass

In this cluster, there is an emphasis on the research of cellulosic ethanol with corn biomass, such as corn stover - a set of leaves, stems, and corncobs. It is possible to verify that the United States stands out in this field of research due to the great interest that the country has in the subject (Fig. 20). Several country-specific studies can be found, such as the life cycle analyzes of maize grain and maize straw in the United States. Although with a smaller number of publications, Brazil also appears within the 5 countries that publish the most in cluster 15 (Table XVI).



Figure 20: Word cloud from cluster 15

Table XVI: Top 5 countries in cluster 15

Cluster 15 - Corn biomass	
Countries	Total
USA	51
[Canada, USA]	6
[China, USA]	5
Brazil	2
Serbia	2

3.2.15 Considerations for clusters analysis

The analysis of the clusters was able to identify several segments of the bioeconomy, mainly those associated with the production and processing of renewable raw materials (RRM). Of the 15 clusters, 9 were directly associated with an RRM, while the other 6 were related to the topics of bioproducts, processing, and circular economy. This result reinforces the idea that the main dimension of the bioeconomy is related to the substitution of fossil raw materials for renewable ones. It is important to mention that the characterization of the network was not exhaustive, with the 15 clusters representing 42% of the network (a total of 1947 articles). Several themes that can be found in the network, such as biodiesel (239 articles), sugar cane (437 articles), and bioplastics (50 articles), did not form

semantic similarity clusters large enough to be selected in the characterization process. Thus, to deepen the discussion on the generated network, analyzes were made for the 5 countries that publish the most on bioeconomy - the USA, Germany, Italy, China, and England - and also for Brazil. The analysis by country is important in the context of the bioeconomy due to the strong influence that regional characteristics have on the theme. As seen in this section, some countries showed a great predominance in publications on specific topics, such as Finland in the forest biomass cluster (cluster 1), Malaysia in the palm oil cluster (cluster 8), Brazil in the straw cluster (cluster 9) and the USA in the corn biomass cluster (cluster 10).

3.3 Top 5 countries – Number of publications in the network

Fig. 26, in section 3.4, presents the evolution of the number of articles from the 11 countries that have published the most, according to the generated network, until the year 2019. It is observed that the increase of publications takes place in the mid-2000s and since then has been following a growth trend. It is worth highlighting in this graph the accelerated growth of publications from Germany and China as of 2014. The results for each of the five analyzed countries are presented below.

3.3.1 United States of America

The USA is the country with the largest number of publications with a total of 896 articles, which corresponds to 19% of the network generated. By looking at Fig. 21, it is possible to identify a keyword profile similar to the complete network word cloud (Fig. 3), with the biofuels and bioenergy sectors as the most present in the set of publications.



Figure 21: USA word cloud

3.3.2 Germany

Fig. 22 shows the keyword cloud for German publications. The country produced 494 articles - 11% of the network - of which 206 were in cooperation with other countries. This cooperation was widely distributed, with even the main partners having small participation: Italy (4%), Holland (3%), and China (3%)

Brazil has a total of 215 publications on the network (5%), of which 42% are in partnership with other countries. The map in Fig. 28 shows Brazil's main publication partners.



Figure 28: Brazil's main publication partners.

Fig. 29 shows the 15 organizations that publish the most in Brazil. It is possible to verify that 12 of the 15 are federal and state universities. It is worth mentioning Embrapa and CNPEM as research institutions that also have a strong role in the subject of bioeconomy.

Fig. 30 shows the map of Brazil with the number of publications by state. These figures make clear the supremacy of the Southeast region, especially the state of São Paulo, in scientific production on bioeconomy. Considering the enormous biodiversity that exists throughout the national territory, this result highlights the need for more STI institutions in bioeconomy beyond the southeastern region.

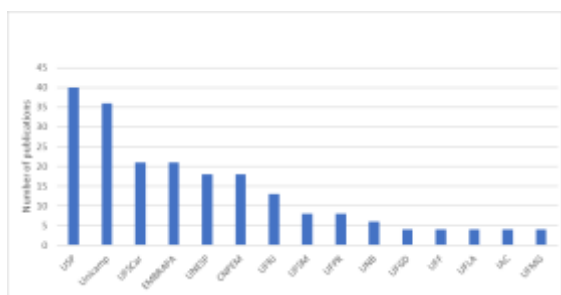


Figure 29: Main institutions that publish about bioeconomy in Brazil

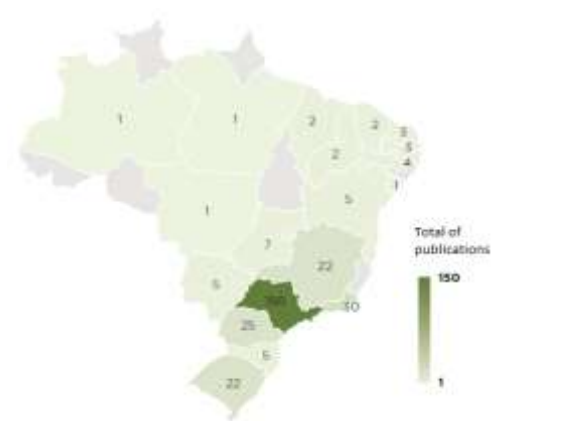


Figure 30: Distribution of publications by Brazilian states.

4 CONCLUSIONS

This article presented the results of a semantic similarity network built from a snowball methodology on the term bioeconomy. The results emphasized mainly: the influence of regional characteristics on the types of publications on bioeconomy, the predominance of the biofuels and bioenergy sectors in the subject, and the emphasis on treating bioeconomy through the dimension of biomass or renewable raw materials. The specific results for Brazil showed how the country has been standing out in recent years in terms of the number of publications on the topic. The analysis also showed that research in bioeconomy is still very concentrated in the Southeast region, mainly in the state of São Paulo.

The collected data presents some limitations, such as the use of a single database, the Web of Science database; a methodology based on a small initial sample set (1675 articles); and a partial characterization of the network (15 clusters). Future works will aim to reduce the limitations of the study on the bioeconomy by seeking new data sources, such as other scientific and patent databases, and broadening the scope of the search on the topic.

5 REFERENCES

- [1] CGEE, 2020. Espaço conceitual da Bioeconomia. Brasília: Centro de Gestão e Estudos Estratégicos, 2020.
- [2] European Union, European Commission. A sustainable Bioeconomy for Europe: strengthening the connection between economy, society and the environment. Updated Bioeconomy Strategy. Bélgica: 2018. Available in: <https://ec.europa.eu/research/bioeconomy/pdf/ec_bioeconomy_strategy_2018.pdf#view=fit&pagemode=none>. Access in: nov 2020.
- [3] Germany, The Federal Government. National Bioeconomy Strategy. Berlin: jul. de 2020. Available in: <https://bioeconomie.de/sites/default/files/bmbf_national-bioeconomy-strategy_en_0.pdf>. Access in: nov 2020.
- [4] Nordic co-operation. Development of the Nordic Bioeconomy. NCM reporting: Test centers for green energy solutions - Biorefineries and business needs. 04 fev. 2016. Available in: <<https://www.norden.org/en/publication/development-nordic-bioeconomy>>. Access in: nov 2020.
- [5] United States of America, The White House. National Bioeconomy Blueprint. Washington: abr. 2012. Available in: <https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/national_bioeconomy_blueprint_april_2012.pdf>. Access in: nov 2020.
- [6] Brazil, Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC). Plano de Ação em Ciência, Tecnologia e Inovação em Bioeconomia. 2018. Brasília, DF: Centro de Gestão e Estudos Estratégicos, 2018.
- [7] M. Bugge; T. Hansen; A. Klitkou. What is the bioeconomy? A review of the literature. Sustainability (Switzerland) (2016), v. 8, n. 7.
- [8] D. D'amato; N. Droste; B. Allen; M. Kettunen; K. Lähinen; J. Korhonen, P. Leskinen, B. D. Matthies; A. Toppinen. Green, circular, bio economy: A

comparative analysis of sustainability avenues.
Journal of Cleaner Production (2017) v. 168, p. 716–
734.

8 ACKNOWLEDGEMENTS

- The Authors are grateful to the technical teams of the CGEE and,
- CGBE/MCTI team of the ODBio project.



MINISTRY OF
SCIENCE, TECHNOLOGY
AND INNOVATION

