

Bibliometric and Text Mining Analysis of SDGs Governance Research

Dr. Hassan Rabie*

Abstract

This paper presents a comprehensive analysis of the research on governance for the Sustainable Development Goals (SDGs) using bibliometric and text mining techniques. The study collected 6,341 articles from the Web of Science Core Collection database from January 2016 to December 2022. Bibliometric analysis revealed an increasing trend in publications related to SDGs governance, with a peak of 1,410 articles in 2022. The analysis also identified the top contributing countries, institutions, subject categories, and sources in the field. Text mining analysis employed topic modeling, specifically latent Dirichlet allocation (LDA), to identify key themes and topics within the collected articles. The findings highlighted the importance of engaging diverse stakeholders, achieving the SDGs, addressing environmental challenges, identifying research gaps, utilizing qualitative analysis methods, and exploring the relationship between governance and sustainability. The analysis also emphasized the significance of gender equality, corporate social responsibility, decision support systems, education policies, water resource management, and green innovation for sustainable development. The results provide valuable insights into the research landscape and hotspots in SDGs governance, facilitating evidence-based decision-making and future research directions.

Keywords - SDGs, Text Mining, Bibliometric analysis

* Associate Professor- Planning Techniques Center – Institute of National Planning.

تحليل ببيومتری وتعدين النصوص لأبحاث حوكمة أهداف التنمية المستدامة ملخص

تقدم هذه الورقة تحليلاً للبحوث المنشورة حول حوكمة أهداف التنمية المستدامة باستخدام التحليل الببيومتری وتعدين النصوص. جمعت الدراسة 6,341 مقالة من قاعدة بيانات Web-of-Science من يناير 2016 إلى ديسمبر 2022. أظهر التحليل الببيومتری اتجاهاً متزايداً في المقالات المتعلقة بحوكمة أهداف التنمية المستدامة، حيث بلغ عدد المقالات ذروته 1,410 مقالة في عام 2022. كما تم تحليل هذه الأبحاث طبقاً للبلد والمؤسسات البحثية والموضوعات الرئيسية والمصادر الأكثر أهمية في هذا المجال. وباستخدام تعدين النصوص من خلال تطبيق (Topic Modeling) لتحديد (Themes) الرئيسية لجميع المقالات، وتحديد خوارزمية (LDA)، لتحديد الموضوعات الرئيسية داخل المقالات المنشورة. أبرزت النتائج أهمية إشراك أصحاب المصلحة، وتحقيق أهداف التنمية المستدامة، ومعالجة التحديات البيئية، وتحديد الفجوات البحثية، واستخدام الأساليب الكمية، واستكشاف العلاقة بين الحوكمة والاستدامة. أظهرت النتائج أيضاً أهمية المساواة بين الجنسين، والمسؤولية الاجتماعية للشركات، وأنظمة دعم القرار، وسياسات التعليم، وإدارة موارد المياه، والابتكار الأخضر من أجل التنمية المستدامة. عرضت الورقة أهم النقاط البحثية والنقاط الأكثر تداولاً في حوكمة أهداف التنمية المستدامة، مما يساعد متخذي القرار على تحديد الاتجاهات المستقبلية للأبحاث.

الكلمات المفتاحية: التنمية المستدامة - تعدين النصوص - التحليل الببيومتری.

1. Introduction

In September 2015, member states of the United Nations (UN) agreed on the 2030 Agenda for Sustainable Development. With its 17 Sustainable Development Goals (SDGs, also called "Global Goals") and 169 targets, the agenda demonstrates the international commitment to achieve worldwide sustainable development in its social, economic, and environmental dimensions. What makes the SDGs special is the broad acceptance and commitment of the international community, the comprehensive definition of Sustainable Development (SD) in its different dimensions made measurable through 232 indicators, and the understanding that these sustainability goals are universal, integrated and indivisible (United-Nations 2015). The emergence of the goals can be understood in the context of and as a response to global problems emerging in the wake of globalization processes and increasing global interconnectedness.

The SDGs are integrated, and action in one of them will affect outcomes in others, making research on the SDGs broad, complex, and fragmented. To achieve the SDGs, all states are called upon to integrate SDGs into their national sustainability and development plans (United-Nations 2015). Therefore, governance for the SDGs needs to foster an enabling environment for collective action, ensure that the decision makers involved are held accountable and deal with emerging complex trade-offs between the

goals (Kanie and Biermann 2017, Glass and Newig 2019). Governance has been referred to as the “fourth pillar of sustainable development” (Glass and Newig 2019).

Governance refers to the institutions and processes involved in managing a society's resources and affairs (Elder et al. 2016). It includes how a community allocates and utilizes resources from the environment, develops and implements economic activity, directs the products of the economy to achieve social goals, and oversees domains like education, health, and welfare (Elder et al. 2016). Governance for the SDGs refers to the processes and institutions that are necessary to plan, implement, and monitor progress towards achieving the SDGs. It also refers to the coordination and management of efforts to achieve sustainable development goals. It matters about the participation of all stakeholders, including governments, civil society, the private sector, and international organizations. It involves creating an enabling environment for collective action, ensuring accountability of actors involved, and dealing with emerging complex trade-offs between the goals.

Good governance has been called ‘the Pandora’s box of the SDGs’¹. Good governance focuses on processes of decision-making and their institutional foundations. Governance ensures on values such as enhanced participation and inclusion, transparency, accountability and access to information, also, it focused on combatting corruption, securing basic human rights and the rule of law (Kanie and Biermann 2017).

Bibliometric and Text Mining

Bibliometric analysis is a quantitative method to analyze scientific publications and citations, in a specific subject area or field. This can include analyzing publication trends, identifying influential publications, authors, keyword analysis, and mapping the network of collaborations between authors, countries and institutions (Roy et al. 2022). Bibliometrics allows for a quantitative, data-driven analysis of large volumes of research publications to gain useful insights. There are also several bibliometric tools and software available to facilitate such analyses. Bibliometric analysis helps researchers and policymakers gain valuable insights into scientific publications and contribute to evidence-based decision-making.

Text mining concerns the use of text analysis methods to extract insights from large volumes of unstructured textual data. It is also known as text analysis or text analytics, which refers to the process of extracting meaningful information and insights from unstructured textual data. This can include identifying patterns and trends in language use, sentiment analysis, and topic modeling (Hwang et al. 2021, Roy et al. 2022).

Topic modeling is a technique that automatically identifies topics or themes within a collection of documents. It helps in understanding the main themes and discussions present in the text data. Topic modeling is a powerful tool for extracting insights from

¹ <https://www.theguardian.com/public-leaders-network/2015/sep/25/good-governance-sustainable-development-goals-united-nations>

large collections of unstructured text data. It has applications in a wide range of fields and can help researchers and analysts to identify patterns and trends that might otherwise be difficult to detect. One such topic modelling method is latent Dirichlet allocation (LDA), which is a very popular unsupervised modelling technique used for topic modelling (Hwang et al. 2021, Roy et al. 2022).

In this paper, bibliometric and text mining will be employed to examine the literature on governance for the SDGs. These analyses can help identify key characteristics, themes, and trends in the research subject of governance for the SDGs. Furthermore, they can facilitate the mapping of collaboration networks and the assessment of the influence of authors and institutions in this field. Additionally, these analyses can offer insights into the language employed in the literature and the attitudes of researchers towards governance for the SDGs.

Literature Review

There are many types of literature reviews, including traditional reviews, scoping reviews, systematic reviews, meta-analyses, critical reviews, overviews, state-of-the-art reviews, rapid reviews, and bibliometric reviews. This study chose to conduct a bibliometric review for several reasons: (a) It provides a quantitative analysis, enabling greater reproducibility. (b) It incorporates a wide range of publications, allowing analysis of a large collection of papers. (c) Statistical tools can produce fairly precise estimates about research areas. (d) Various tools are available to conduct bibliometric analyses, such as VOSviewer, BibExcel, CiteSpace, and Pajek (Roy et al. 2022).

The study of Sustainable Development and its various aspects has attracted considerable interest from researchers over time. As they explore this important field, insights have emerged from a wide range of studies. These collective efforts contribute to forming a comprehensive understanding of Sustainable Development. By delving into different subjects, researchers are gradually revealing the interconnected nature of Sustainable Development. The combined findings from these studies gradually build a holistic view of what Sustainable Development truly entails. For example in the study of (Hassan et al. 2014), the authors embarked Sustainable Development research using Scopus database from 2000 to 2010. The study identified sub-areas of Sustainable Development (SD) and examined the research landscape at the country and institute levels. The study suggests that institutes strong in Sustainable Development overall may not be strong in all sub-areas, and vice versa. China has a strong publication output but not a strong citation count. The study's findings can be useful for government research agencies and institutes looking to find strategic partners for interdisciplinary research areas.

(Zhang et al. 2017) explored the realm of Water Footprint (WF) research from 2006 to 2015, employing the Web of Science database. Their exploration identified evolving trends, hotspots, and frontiers within WF research. A surge in research activity was evident, with the United States, China, and the Netherlands leading in productivity. Amidst this landscape, the University of Twente emerged as a hub of scholarly output.

The study underscores vital themes like accounting methodologies, water resource management, and the intricate water-food-energy nexus. By charting this course, the study sets a reference point for forthcoming investigations. Building upon this subject, (Adetoro et al. 2021) highlight the significance of Water Footprint, Water Sustainability, and Water Productivity (WFSP) in mitigating water scarcity in nations grappling with this challenge. The authors evaluated the global trends in research related to water footprint, water sustainability, and water productivity over time. They used bibliometric analysis to identify the most productive countries, institutions, and authors in this field and examine the research trends and themes. The study found that WFSP research has significantly increased over the past decades, with China, the United States, India, and the Netherlands having the most WFSP publications. However, publications in South Africa were relatively low. The study concludes that enhancing collaboration between multiple countries and authors in the field of WFSP research could provide important scientific solutions to global water scarcity. Rounding out this exploratory work, (Roy et al. 2022) delve into comprehending the landscape of SDG 6—water and sanitation. Using a combination of bibliometric and text mining methods, they analyzed 289 publications from the Web of Science database. The analysis included co-citation, co-occurrence, cooperation networks, theme networks, cluster analysis, word dynamics, and thematic evolution techniques. The findings revealed a lack of participation and collaboration with countries from the Global South in SDG 6 research. To address this, the study proposed the Sustainable Development of Water and Sanitation (SDWS) framework, which classifies the research domain by considering its connections to the environment, economy, and society. The study provides an overview of the water and sanitation research field and highlights areas for future investigation.

The work further unfolds, (Zhu and Hua 2017) analyzed the sustainable development research landscape. Through an analysis of a Web of Science dataset, they uncover the evolving knowledge structure of SD. They analyzed a dataset from the Web of Science database and used CiteSpace to examine the knowledge structure and evolution of SD. The analysis covered author, journal, and keyword networks, as well as categories and countries distribution and highly cited references. The United States and UK were prominent in SD studies, with China having the highest publication counts. The concept of nature capital played a significant role in interpreting SD, while the materials category and social sciences were identified as promising disciplinary frontiers. The study introduced a new indicator ($BC \times CB$) for keyword analysis and identified valuable keywords, citation maps, and visible hot research areas. Expanding on this topic, (Mesquita et al. 2017) consolidated the state of academic research on mining, sustainability, and sustainable development by organizing the results of previous studies within a systematic review on the SDGs. The paper uses the Web of Science Core Collection database as a database of record, as it is one of the most widespread databases of academic journals. The results shows that the focus of publications on mining and sustainability is mainly on the environmental dimensions of the SDGs, suggesting a need for more practical and academic work in the mining sector to address

the other goals that make up the SDGs framework. A comprehensive study by, (Körffgen et al. 2018) cast the spotlight on Austria's contribution to the United Nations' Sustainable Development Goals (SDGs) using a bibliometric lens. They emphasize the role of universities in integrating sustainability issues into their operations, research, education, and science-society interactions to support the implementation of the SDGs. The study finds that research related to SDG 3 (Good Health and Well-being) and SDG 4 (Quality Education) is well represented, while other SDGs are underrepresented. The results can assist universities in identifying thematic focuses for their research within the SDG framework and promote inter-university collaboration to address SDG implementation challenges. The findings can guide future research funding and policy decisions on sustainable development in Austria. (Salvia et al. 2019) journey into the realm of global SDG priority, revealing varying emphases across different regions. Through a bibliometric analysis, the authors examine scientific publications on the SDGs and identify the most frequently addressed goals and research areas. The findings reveal a connection between local challenges in specific regions and the main areas of interest among experts, with a notable focus on SDG 4 (Quality Education), SDG 11 (Sustainable Cities and Communities), and SDG 13 (Climate Action). The paper presented both positive and negative examples from different regions and assesses the global pursuit of various SDGs through research.

The exploration continues as (Armitage et al. 2020) embark on a comparative expedition across independent bibliometric methodologies. The study uses three different bibliometric tools to analyze scientific publications related to the SDGs and assesses the similarities and differences in their outputs. The study provided insight into the use of bibliometric tools for mapping research related to the SDGs and emphasizes the need for critical reflection on the limitations and biases of these tools. The results show that different approaches can greatly alter the publications retrieved and resulting country rankings. The choice of search terms, how they are combined, and query structure play a role, related to differing interpretations of the SDGs and viewpoints on relevance. Therefore, currently available SDG rankings and tools should be used with caution at their current stage of development. Expanding on this subject, (Diaz-Lopez et al. 2021) analyzed the evolution of research on the SDGs through a systematic literature review and bibliometric analysis. The author used SciMAT software to identify hidden themes and their development in this field from 1990 to 2020. The results show a constantly evolving scientific field, from its initial focus on the millennium goals to the gradual inclusion of the current SDGs. The paper provided a comprehensive overview of the status quo and predicted the dynamic directions of future research, serving as a basis for the development of new strategies for the implementation of the SDGs.

The paper is relevant for researchers, policymakers, and practitioners interested in sustainable development. Building upon this subject, (Hwang et al. 2021) navigate the area of climate change awareness across different social groups in Korea, specifically within the realm of SDG 13.3. Their results showed that the Korean public had a relatively high awareness of early warning, while the government and academia had high

awareness of both mitigation and adaptation. Corporations in Korea were focused on climate change mitigation. The study proposed that this methodology could be used as an SDG indicator to measure differences in awareness and inform tailored policy formulation. (Payumo et al. 2021) used bibliometric methods and network analysis to examine research output and collaboration supporting the SDGs.

The study aims to explore means to detect and analyze research collaboration beyond the traditional definition of multiple, one-time co-authorship. The authors employed two additional lenses of collaboration: repeat collaboration and collaboration time point to quantify and visualize co-authorship data sourced from Microsoft Academic Graph. The results show an increased collaboration rate over time at the author and institutional levels. However, most collaborations in SDG-related research only happened once. The study also found that repeat collaboration happens more frequently, but after a longer duration, at the institutional level than at the author level. The authors further analyzed institutions and identified core institutions that could help influence more consistent collaboration and sustain, or grow, the research network for SDG-related research. The results have implications for understanding sustainable partnerships in research related to SDGs and other global challenges. The study recognizes the limitations and caveats of bibliometric data from abstract and citation databases of peer-reviewed research literature brought on by incomplete coverage and the difficulties of author disambiguation. The methodology builds on best practices on research evaluation developed throughout the years on quantitative science and technology studies.

Rounding out this exploratory work, (Shen et al. 2021) embark on a systematic exploration of food safety governance, weaving together a comprehensive narrative. Their bibliometric analysis evolving trends and illuminates future research pathways. They identified current trends and predicted future research directions. The findings highlighted food safety governance as a multidisciplinary field, with the United States leading in relevant articles and Wageningen University being the most influential institution. The development of food safety governance research was categorized into three processes: formulation of standards, implementation of standards, and co-governance across multiple sectors. Popular research areas included policy integration and public-private partnerships in food safety governance. Lower- and middle-income countries focused on food supply and system design, while higher-income countries emphasized food safety and nutrition. Future research directions include co-governance, online food governance, consumer willingness to purchase safe food, and food safety governance during pandemics.

(Bautista-Puig et al. 2021) focused on analyzing the research output on SDGs, with a specific focus on Higher Education Institutions (HEIs) and Research Centers (RCs). The study aims to map the global research of sustainability goals, describe thematic specialization, and assess SDGs interconnections. The authors conducted a bibliometric analysis to measure "core" research output and identified 25,299 bibliographic records, from which 21,653 (85.59%) are from HEIs and RCs. The

findings reveal the increasing participation of these organizations in this research, with 660 institutions involved in 2000-2005 and 1744 institutions involved in 2012-2017. The study also highlights the specialization of some institutions in the topic of sustainability goals, such as the London School of Hygiene & Tropical Medicine and the World Health Organization.

Bibliometric and text mining have proven to be valuable tools for understanding and getting insights about SDGs. Several studies have utilized bibliometric and text mining analysis to examine research trends, identify key contributors and institutions, and highlight thematic emphases within specific SDG areas. These analyses have provided insights into the research landscape, highlighting areas of strength and weakness, and facilitating interdisciplinary collaborations. By combining quantitative and qualitative approaches, these studies have contributed to a deeper understanding of SDGs, identified research gaps, and proposed frameworks for sustainable development in various domains. Overall, bibliometric and text analysis offer valuable tools for policymakers, researchers, and institutions to inform decision-making, and guide future research and policy directions in the context of SDGs.

Data Collection

In this study, the data were retrieved from Web-of-Science Core Collection database; it is considered the most extensive and comprehensive used database. The data covered all the articles related to governance and sustainable development goals from January 2016 to December 2022. The search words and terms included topic= [(“SDG”, OR “sustainable development goal?”, OR “sustainable development”) AND (“governance”)]. A total of 6,341 articles have been retrieved and will be analyzed using bibliometrics and text analysis.

We used the ‘bibliometrix’ package in R², Citespace software (V. 6.2.R3)³ and VOSviewer⁴ for the bibliometric analysis. We performed text mining analysis using Matlab⁵. The "bibliometrix" package in R is a tool used for bibliometric analysis. Bibliometric analysis involves quantitative analysis of publications, citations, and other bibliographic data to gain insights into the patterns of scholarly research and communication. This package is particularly useful for researchers and analysts who want to explore and understand the relationships among scientific publications, authors, journals, and keywords. The "bibliometrix" package provides functions to perform various bibliometric analyses, including: (1) Co-authorship Analysis: analyzing the collaboration patterns among authors. (2) Co-citation Analysis: to identify the most influential papers based on the citations they received. (3) Journal Analysis: analyzing publication trends, and journal co-citation networks.

² <https://www.bibliometrix.org/home/>

³ <http://cluster.cis.drexel.edu/~cchen/citespace/>

⁴ <https://www.vosviewer.com/>

⁵ <https://www.mathworks.com/products/matlab.html>

VOSviewer is a software tool for constructing and visualizing bibliometric networks and scientometric data. These networks may for instance include journals, researchers, or individual publications, and they can be constructed based on citation, bibliographic coupling, co-citation, or co-authorship relations. VOSviewer also offers text mining functionality that can be used to construct and visualize co-occurrence networks of important terms extracted from a body of scientific literature⁶. Key features of VOSviewer include: (1) Network Visualization: VOSviewer allows you to create visualizations of networks, where nodes represent entities (such as authors, keywords, or documents) and edges represent relationships between them (such as co-authorship or co-occurrence). (2) Clustering and Mapping: The tool can automatically cluster similar entities and map them to visualize clusters and their relationships. Citespace is a software tool used for visualizing and analyzing the structure and evolution of scientific literature, particularly in the context of citation networks and bibliometric analysis. It is widely used in research and academia to uncover patterns, trends, and relationships in scholarly publications. Citespace was developed by Dr. Chaomei Chen⁷ at Drexel University and is designed to help researchers and analysts explore the dynamics of scientific research through visual representations. It primarily focuses on the analysis of citations, co-citations, and co-authorship networks to provide insights into the development of research fields, the influence of publications, and the evolution of scientific collaboration. Matlab is a high-level programming language and interactive environment primarily used for numerical computing, data analysis, and visualization. MATLAB provides various tools and functions that can be used for text mining and topic modeling tasks.

Bibliometric Analysis - Results and Discussion

Characteristics of Publication Outputs

A total of 6,341 publications were collected from the web-of-science database, which consisted of 4,905 articles (77.3%); 463 review (7.3%), 334 Proceeding Paper (5.3%). The other 642 publications are (book chapter, books, book review, etc.). Figure 1 shows the number of relevant articles published on SDGs governance. The number of governances for SDGs publications increased from 443 in 2016, to 1410 in 2022 -the peak year-. Figure 1 shows an overall growth trend from 2016 to 2022.

⁶ <https://www.vosviewer.com/>

⁷ <https://drexel.edu/cci/about/directory/C/Chen-Chaomei/>

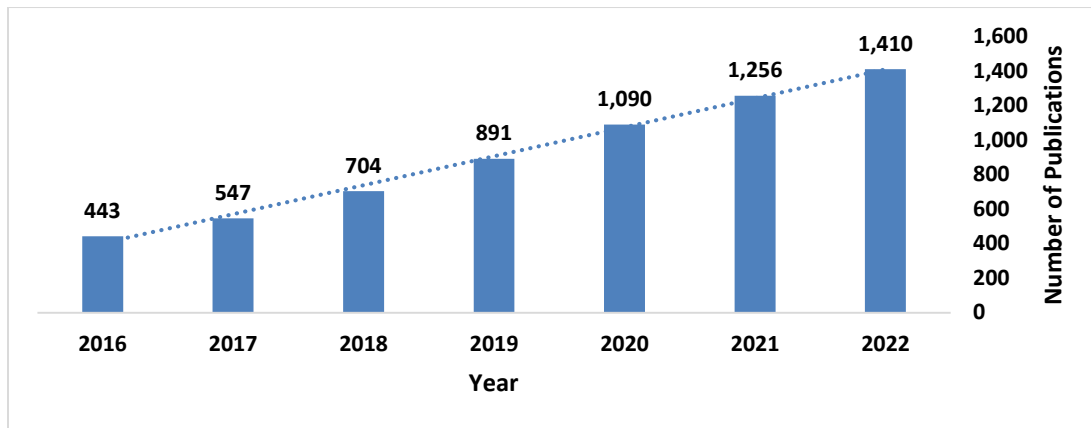


Figure 1: Publication of SDGs governance from 2016 to 2022

Subject Categories

The examination of subject categories can reveal common patterns in current research in SDGs governance, therefore, we used categories from the Web-of-Science database. Every article covered by the Web of Science core collection was assigned **at least one** of the subject categories. Analyzed results section in the Web of Science shows top 10 categories. The governance for SDGs related publications belongs more than 100 categories, of which the top three broad categories are environmental sciences (32.4%) which focuses on the study of the interactions between living and non-living systems, Environmental Studies (30.1%) which focus on the social, cultural, and political dimensions of environmental issues, and Green Sustainable Science Technology (24.9%) which studies the development and implementation of sustainable technologies, (Table 1).

Table 1: The Top Ten SDGs Governance Categories

#	Category	Count	%
1	Environmental Sciences	2,052	32.4%
2	Environmental Studies	1,911	30.1%
3	Green Sustainable Science Technology	1,580	24.9%
4	Management	484	7.6%
5	Business	450	7.1%
6	Economics	399	6.3%
7	Regional Urban Planning	345	5.4%
8	Development Studies	339	5.3%
9	International Relations	276	4.4%
10	Engineering Environmental	255	4.0%

Analysis of Sources

Source is a journal, book, etc. which published one or more articles included in our collection. According to our analysis, about 24% of the publications were published in the top ten journals, where, the top one is specialized in sustainability (Table 2). Among the sources, as per cumulative production dynamics, the top five are

Sustainability; J. Cleaner Production; Business Strategy and the Environment; Sustainable Development; and Marine Policy.

Table 2: The Top Ten SDGs Governance Sources

#	Sources	Articles	%
1	Sustainability	766	12.1%
2	Journal of Cleaner Production	180	2.8%
3	Business Strategy and the Environment	83	1.3%
4	Sustainable Development	80	1.3%
5	Marine Policy	75	1.2%
6	Environmental Science & Policy	72	1.1%
7	Corporate Social Responsibility and Environmental Management	70	1.1%
8	International Journal of Environmental Research and Public Health	64	1.0%
9	Frontiers in Environmental Science	59	0.9%
10	Water	56	0.9%

From the reference list, the top five most-cited local sources were Science of the Total J Clean Production (8,432), Sustainability (6,516), J. Business Ethics (4,754), J. Ecological Economics (2,732), and Business Strategy Environment (2,662).

Analysis of Authors

The total number of authors is about 16,696 authors. 1,089 are authors as single authored -authored by only one author-articles. Fractional authorship is a way of quantifying the contributions made by individual authors to a published work, while if all co-authors made an equal contribution to each document. The most relevant authors, as per fractionalized publications, are Wang Y (11.45), Li Y (11.06), Liu Y (9.68), Li X (8.48) and Wang J (8.0). However, as per published articles, the top five are Wang Y (43), Li Y (41), Liu Y (30), Zhang Y (30), and Li X (28).

Table 3: The top ten relevant Authors for SDGs Governance

#	Authors	Articles	Articles Fractionalized
1	Wang Y	43	11.45
2	Li Y	41	11.06
3	Liu Y	30	9.68
4	Zhang Y	30	7.98
5	Li X	28	8.48
6	Liu X	27	7.5
7	Zhang J	25	6.5
8	Biermann F	24	7.83
9	Wang J	24	8
10	Wang X	24	6.17

Analysis of Affiliations

The ten most relevant global affiliations related to publications on governance of SDGs are University. University Utrecht (175 articles), Stockholm (109 articles), and Queensland Univ (106 articles) were the top three main productive Affiliations.

Table 4: The top ten SDGs Governance Affiliations

#	Affiliation	Articles	%
1	University Utrecht	175	2.7%
2	Stockholm University	109	1.7%
3	University Queensland	106	1.6%
4	University Sao Paulo	97	1.5%
5	Shanghai Jiao Tong University	96	1.5%
6	Tsinghua University	96	1.5%
7	University Toronto	95	1.5%
8	University British Columbia	88	1.4%
9	University Oxford	87	1.4%
10	Lund University	84	1.3%

Analysis of Countries

Among all the author's country that published governance of SDGs-related articles, the top five are China (4,167), USA (2,245), UK (2,087), Germany (1,329), and Australia (1,298), (Figure 2).

Country Scientific Production



Figure 2: Spatial distribution of total citations

Figure 3 reflects the cooperation network between different countries. The size of the circle represented the importance of each country in this research field, and the thickness of the connecting lines indicated the cooperation intensity between countries.



Figure 3: The cooperation network of the productive countries in SDGs governance from 2016 to 2022

Analysis of Publication Articles

Table 5 lists the most globally cited documents. Documents which included in the bibliographic collection.

Table 5: The top ten SDGs Governance Publication Articles

Paper	DOI	Total Citations
KEESSTRA S, 2018, LAND	10.3390/land7040133	356
KOLK A, 2016, J WORLD BUS	10.1016/j.jwb.2015.08.010	326
BIERMANN F, 2017, CURR OPIN ENVIRON SUSTAIN	10.1016/j.cosust.2017.01.010	324
CLARK WC, 2016, PROC NATL ACAD SCI U S A	10.1073/pnas.1601266113	324
ALI SH, 2017, NATURE	10.1038/nature21359	322
SEDDON N, 2020, PHILOS TRANS R SOC B-BIOL SCI	10.1098/rstb.2019.0120	313
BEBBINGTON J, 2018, ACCOUNT AUDIT ACCOUNT	10.1108/AAAJ-05-2017-2929	305
LIU J, 2018, NAT SUSTAIN	10.1038/s41893-018-0135-8	302
CLARK WC, 2016, PROC NATL ACAD SCI U S A-a	10.1073/pnas.0900231108	282
RASUL G, 2016, CLIM POLICY	10.1080/14693062.2015.1029865	282

Keyword Analysis and Hotspots

Through the analysis of title text, Table 6 shows the top ten most frequently used words are 'development' (1934), 'sustainable' (1907), 'governance' (1266), etc... From the analysis of the content of the abstracts, the top ten are 'development' (12756), 'sustainable' (9940), 'governance' (8221), etc... Through the analysis of authors' keywords, the top ten are 'sustainable development (1341)', governance (702),

sustainability (592) etc... Last, Through the analysis of Keywords Plus, the top ten are, governance (1777), management (776), sustainable development (629) etc...

If we want to explore temporal patterns more deeply, we can use cumulative word dynamics. By analyzing the cumulative word dynamics of Keywords Plus, we can gain insights into the evolving trends over time, for example ‘integrated assessment’ was the highest term used from 2017 to 2019, but it was exceeded by ‘state’, ‘community’ and ‘networks’ from 2017 to 2021. The keywords that are currently increasing are ‘corporate-social-responsibility’, ‘environmental performance’ and ‘firm performance’.

Table 6: Comparative table of top keywords on governance for SDGs

#	Titles	Abstract	Author Keywords	Keywords Plus
1	development (1934)	development (12756)	sustainable development (1341)	governance (1777)
2	sustainable (1907)	sustainable (9940)	governance (702)	management (776)
3	governance (1266)	governance (8221)	sustainability (592)	sustainable development (629)
4	sustainability (698)	environmental (4857)	sustainable development goals (421)	policy (454)
5	environmental (599)	study (4560)	sustainable (349)	framework (403)
6	goals (500)	social (4473)	development (330)	performance (402)
7	social (480)	research (4266)	climate change (213)	impact (400)
8	management (432)	sustainability (4206)	SDGs (177)	Climate change (367)
9	urban (412)	paper (3475)	corporate social responsibility (150)	challenges (310)
10	global (393)	policy (3461)	China (146)	politics (241)

A co-occurrence network is formed by connecting terms that occur together within a given unit of text. This network is constructed using a set of co-occurrence criteria, which establish connections between pairs of phrases. Co-occurrence networks provide a popular method for visualizing potential associations among individuals, organizations, concepts, and other entities within a text. The creation and visualization of co-occurrence networks have become feasible with the advent of electronically stored text that complies with text mining standards. From the analysis of the co-occurrence word network of Keywords Plus (Figure 4), we can see there are more than eight clusters. The highest occurring three co-words of major clusters, are ‘governance’, ‘politics, and ‘state’ in cluster 1; ‘management’, ‘sustainable development’, and ‘climate-change’ in cluster 2. ‘impact’, ‘performance’ and ‘corporate social-responsibility’ in cluster 3. ‘China’, ‘energy’, ‘economic-growth’

text corpus? How do these topics relate to each other? Which documents are most closely associated with specific topics?

Latent Dirichlet Allocation (LDA) is considered one of the most popular algorithms for topic modeling. LDA is a method for fitting a topic model that treats each document as a combination of topics, and each topic as a combination of words (Silge and Robinson 2017, Hwang et al. 2021). Therefore, each document can share words between them, rather than being divided into distinct groups or topics, which is similar to how natural language is used. LDA learnt the topic distribution for each document (article) and the word distribution for each topic, given the observed words in the documents and articles (Laureate et al. 2023). In this study, TM was employed to extract latent (suggested) topics and from the unstructured text extracted from web-of-science database (title and abstract text data) of the collected articles using Matlab⁸ code. Choosing the number of topics is one of the most important challenges in TM. If there are too few topics, the model may not be able to capture all the important themes in the data. If there are too many topics, the model may find spurious or redundant topics.

The goal is to choose a number of topics that minimize the perplexity compared to other numbers of topics. Perplexity is a common metric used to evaluate the quality of a topic model. It measures how well the model predicts a held-out (unseen) set of

documents (Laureate et al. 2023). Lower perplexity indicates better model performance. Therefore, we calculate perplexity for different numbers of topics and choose the one with the lowest perplexity.

Steps of text mining and topic modeling in MATLAB:

1. **Data Preparation:**
 - Export data from Web-of-Science.
 - Import your text data into MATLAB.
2. **Text Preprocessing:** Clean and preprocess the text data to prepare it for analysis:
 - Tokenization: Splitting text into words or phrases (tokens).
 - Removing stop words: like "and," "the," "in" that don't carry much meaning.
 - Removing punctuation and special characters.
 - Lowercasing all text for consistency.
3. **Topic Modeling:** Perform topic modeling on the preprocessed text data.
 - Fitting LDA (Latent Dirichlet Allocation) model

⁸ <https://www.mathworks.com/products/matlab.html>

Results

A Matlab code⁹ has been used to fit LDA models for a range of values for the number of topics, the best number of topics with low perplexity value was 35 topics. Table 7 presents the keywords associated with each topic model in the first column, while the second column provides suggestions for general topics/subjects that could be linked to the SDGs governance topic. We can highlight on some important topics, such as:

- Engaging diverse stakeholders in the governance process for achieving the SDGs is broad and encompasses several themes. These topics discuss the role of stakeholders in building partnerships to drive sustainable development. It also covers the importance of governance in promoting SD in the economy, industry, and environment, and the increasingly important role that government and society play in achieving the SDGs.
- Topics that cover the achieving SDGs: Strategies, Targets, and Achievements. These topics emphasize the importance of the SDGs and their targets in promoting SD, and the strategies and achievements of different countries and organizations in meeting these goals.
- Topics that focus on the effectiveness of governance in addressing environmental issues in marine and coastal areas, and the challenges of managing such systems. These topics would concentrate on global governance and environmental management of marine and coastal systems.
- Other topics are identifying gaps in literature and research on sustainable development. These topics could explore the gaps in literature and research on sustainable development, and the need for a comprehensive review and analysis of relevant papers, articles, and studies.
- Some topics could explore the use of qualitative analysis methods in governance research, and the application of different models and data analysis techniques to analyze and interpret research results.
- The relationship between governance and sustainability, and the key factors that influence this nexus, could also be explored. Additionally, the topic of creating value and solving problems through new solutions that promote sustainable development and the need for change and the potential benefits of innovative solutions in achieving the SDGs.
- Sustainable development and governance are crucial in promoting economic growth, environmental protection, and social welfare. It could discuss the role of government, private sector entities, and research in promoting sustainable development at the national/local level. It could also suggest methods for creating

⁹ <https://www.mathworks.com/help/textanalytics/ug/choose-number-of-topics-for-LDA-model.html>

value and solving problems through new solutions that promote sustainable development.

- Some topics could investigate how European countries develop different regions and member states to create a more cohesive Europe.
- Effective governance strategies for water resource management and conservation in natural and human environments are also essential in achieving SDGs. This topic could explore the challenges and opportunities for effective governance of water resources.
- Some topics could discuss the factors that have significant influence on sustainable development outcomes in China.
- Another topic could explore the importance of building capacity for sustainable development by using indicators to measure progress towards the SDGs. It could also investigate the need for effective evaluation systems to assess the impact of sustainable development initiatives.
- Gender equality and economic empowerment are critical in reducing poverty. This topic could explore the challenges and opportunities for promoting gender equality and economic empowerment, and the need to increase opportunities for women and many people around the world.
- Lastly, the impact of corporate social responsibility (CSR) on sustainability performance could be discussed, along with the challenges of global economic growth, urbanization, and health in countries and regions.

Table 7: Extracted topics and their suggested related topics

#	Topic Keywords	Potential/Suggested Topics
1	stakeholder, process, governance, local, community, actor, partnership, engagement	These keywords highlight the significance of stakeholder engagement, participatory processes, actor partnerships, and local community involvement in the governance and implementation of the SDGs.
2	role, policy, government, public, local, governance, sector, national, private	These keywords are centered around the roles of different actors, including governments, public, and private sectors, in policy formulation, implementation, and governance.
3	research, policy, provide, future, practice, practical	These keywords underscore the critical role of research in informing policies, providing practical insights, and identifying future directions.
4	governance, economy, industry, environment, important, government, promote, society	These keywords illustrate the importance of governance, the economy, industry, and the environment in SDGs governance, with the focus on promoting sustainability, balancing economic growth with environmental considerations, and addressing societal needs.
5	development, sustainable, goal, achieve, SDGs, target,	These keywords revolve around the development and implementation of strategies and agendas to achieve the SDGs. It focuses on setting targets, monitoring progress, and ultimately

#	Topic Keywords	Potential/Suggested Topics
	agenda, strategy, achievement	achieving the SDGs through coordinated efforts and sustainable practices.
6	governance, ocean, address, marine, environmental, issue, coastal, management	These keywords are centered around addressing governance issues in the context of ocean, coastal and marine environmental management.
7	value, solution, problem, good, new, potential, future, understanding, need, create	These keywords focus on the value and potential of creating new solutions to address problems and create a better future, and understanding the issues of development innovative approaches to achieve SDGs.
8	management, framework, purpose, planning, fishery, examine, governance	These keywords revolve around the management and governance of fisheries. It involves the purposeful planning and implementation of frameworks and strategies to examine and improve fisheries management practices.
9	European, take, country, develop, state, different, region, policy, member	These keywords center around the European region and the diverse approaches taken by different countries and member states to develop and implement policies aligned with the SDGs.
10	result, show, effect, finding, indicate, factor, spatial, suggest, significant, China	These keywords highlight the results and findings of studies, particularly with regard to the effects and factors influencing sustainable development. The focus is on China research.
11	sustainability, governance, framework, global, consider, political, new, institutional	These keywords revolve around the emergence of new frameworks and models for sustainability governance on a global scale. It considers the political and institutional dimensions of sustainable development.
12	study, nexus, explore, factor, identify, relationship	These keywords are focused on studies that explore the nexus and relationships between different factors within the context of SD.
13	system, make, decision, need, good, support, governance	These keywords revolve around decision-making and the need for robust systems and good governance to support sustainable development.
14	education, policy, forest, national, sustainability, level, implementation, regional	These keywords encompass the integration and implementation of education policies at national and regional levels to promote forest sustainability.
15	analysis, data, method, study, model, apply, qualitative	These keywords focus on research methodology and data analysis. It involves the use of various methods and models to study and analyze qualitative data.
16	water, resource, management, natural, human, area, environment, ecological, quality, conservation	These keywords revolved around the management of water resources, both natural and human-made. It includes considerations of the ecological aspects of water management, as well as the preservation of water quality and the protection of the environment in water-related areas.

#	Topic Keywords	Potential/Suggested Topics
17	country, develop, level, low, market, high, cost, among, find, income	These keywords revolved around country-level development and the disparities in income, market conditions, and costs between low-income and high-income countries.
18	indicator, SDG, measure, assess, progress, evaluation,	These keywords center around the development and use of indicators to assess and measure sustainability and progress towards the SDGs.
19	literature, research, review, study, gap, analysis, present	These keywords revolved around conducting a literature review or research study to identify gaps in existing knowledge and present the findings.
20	people, world, poverty, woman, economic, gender	These keywords emphasize the need to address poverty and gender-related issues to improve the lives of people worldwide.
21	social, corporate, governance, responsibility, environmental, economic, dimension,	These keywords involved the integration of social, environmental, and economic dimensions in corporate governance and responsibility.
22	challenge, face, global, economic, growth, urban, health, country, increase	These keywords revolve around the challenges faced in achieving global economic growth, particularly in urban areas, while addressing issues related to health and increasing disparities among countries and regions.
23	green, environmental, innovation, impact, effect, pollution, emission, reduce	These keywords center around promoting green and environmental innovation to reduce the impact of pollution, emissions, and other environmental issues.
24	business, governance, sustainability, network	These keywords revolve around the intersection of business, governance, and sustainability.
25	society, key, international, group, focus, agreement, cooperation, discussion	These keywords involve the role of society and key international groups in promoting sustainable development. It includes three main focuses: agreement, cooperation, and discussion.
26	company, performance, firm, report, financial, ESG, corporate, disclosure, information, CSR	These keywords revolved around the company's performance, financial reporting, and corporate disclosure related to environmental, social, and governance (ESG) factors, with the emphasizes on the importance of corporate social responsibility (CSR).
27	tourism, area, rural, time, development, China,	These keywords center around tourism development, particularly in rural areas in China.
28	study, case, approach, design, methodology, recent, river	These keywords involve the study of river cases using a methodological approach and research design.
29	risk, Africa, framework, south, disaster, challenge, global, opportunity, policy	These keywords involve the risk and disaster management framework in Africa, particularly in South Africa, with the focus on understanding the challenges and opportunities.

#	Topic Keywords	Potential/Suggested Topics
30	cultural, framework, theory, perspective, pandemic, covid19, heritage, theoretical	These keywords involve the development of a cultural framework and theoretical perspective to understand the impact of the COVID-19 pandemic on cultural heritage.
31	change, climate, action, adaptation, global, policy, mitigation	These keywords involve climate change action and policy, with the focus on addressing the challenges posed by climate change through both adaptation and mitigation measures, and the need for transformative action.
32	city, urban, smart, datum, technology, planning, digital, infrastructure, governance	These keywords involve smart city planning and governance in the context of urban areas and utilizing new technologies and digital infrastructure to enhance urban development and improve the quality of life for residents.
33	energy, food, production, supply, economy, security, chain, consumption, sector	These keywords revolve around the energy and food sectors. It encompasses the production, supply, and consumption of energy and food, considering their impact on the economy, security, and sustainability.
34	development, agenda, SDGs, nations, global, achieve	These keywords revolve around the development and achievement of sustainable goals on a global scale.
35	health, service, ecosystem, human, public, reserve, wellbeing	These keywords involve the intersection of health, ecosystem, and human wellbeing. It emphasizes the importance of providing equitable and accessible health services as a fundamental right.

Conclusion

The objective of this study was to consolidate and analyze a compilation of research papers related to the governance of SDGs. The research papers were gathered from the Web of Science, and their findings and insights were analyzed using bibliometric and text mining methods. Research on the governance of SDGs has been analyzed using bibliometric and text mining techniques. The analysis of scientific production on SDGs in universities from different countries revealed the formation of clusters of authors, organizations, and countries, with keywords related to the focus of the study being prominent. A comprehensive review of contributions to SDGs research has been conducted, identifying hidden themes and their development over time. This review provides a comprehensive overview of the status and future directions of SDGs research, serving as a basis for the development of new strategies for implementation.

Through bibliometric analysis of research status, research output and trends, and research hotspots, several conclusions were drawn. As to the annual publishing trend, the number of articles about SDGs governance had been increasing rapidly after 2016. As to published countries and academic institutes, China, United Kingdom, and USA had made significant contributions in this field. University Utrecht, Stockholm University and University Queensland were the three most influential institutions in the field. Research into SDGs governance field involved cross-disciplinary studies in

Environmental Sciences, Environmental Studies, Green Sustainable Science Technology, and other disciplines. Sustainability, Journal of Cleaner Production, and Business Strategy and the Environment were the top three journals with most publications. The research hotspots mainly covered Corporate-social-responsibility, Environmental Performance and Firm Performance of SDGs governance.

Through text mining of modeling topic, several conclusions were drawn. Achieving the SDGs requires the engagement of diverse stakeholders, including local communities, in the governance process. This involves building partnerships and implementing policies and programs that promote sustainable development at the national and local levels. Effective governance is also critical in promoting sustainable development in the economy, industry, and environment. Research plays an essential role in informing policy and practice in sustainable development, and effective communication of research findings to policymakers is necessary. The emerging governance frameworks need to promote global sustainability, taking into account political and institutional considerations. The key factors that influence the nexus between governance and sustainability need to be identified. Building effective decision support systems for good governance is necessary. There is a need to integrate forest sustainability into national education policies and strategies, explore the use of qualitative analysis methods in governance research, and identify effective governance strategies for water resource management and conservation. Gender equality and economic empowerment are necessary to reduce poverty. Corporate social responsibility plays a significant role in sustainability performance. The challenges of global economic growth, urbanization, and health in countries and regions need to be addressed, and green innovation and sustainable enterprise are essential in reducing environmental pollution and emissions.

Acknowledgment

I would like to acknowledge Prophet Muhammad (Peace Be Upon Him) who has taken all human being from the darkness to the lightness.

References:

- Adetoro, A. A., M. Ngidi, Y. S. Nyam and I. R. Orimoloye (2021). "Temporal evaluation of global trends in water footprint, water sustainability and water productivity research." Scientific African **12**: e00732.
- Armitage, C. S., M. Lorenz and S. Mikki (2020). "Mapping scholarly publications related to the Sustainable Development Goals: Do independent bibliometric approaches get the same results?" Quantitative Science Studies **1**(3): 1092-1108.
- Bautista-Puig, N., A. M. Aleixo, S. Leal, U. Azeiteiro and R. Costas (2021). "Unveiling the research landscape of sustainable development goals and their inclusion in higher education institutions and research centers: major trends in 2000–2017." Frontiers in Sustainability **2**: 12.
- Churchill, R. and L. Singh (2022). "The evolution of topic modeling." ACM Computing Surveys **54**(10s): 1-35.
- Diaz-Lopez, C., C. Martin-Blanco, J. J. De la Torre Bayo, B. Rubio-Rivera and M. Zamorano (2021). "Analyzing the scientific evolution of the sustainable development goals." Applied Sciences **11**(18): 8286.
- Elder, M., M. Bengtsson and L. Akenji (2016). "An optimistic analysis of the means of implementation for sustainable development goals: Thinking about goals as means." Sustainability **8**(9): 962.
- Glass, L.-M. and J. Newig (2019). "Governance for achieving the Sustainable Development Goals: How important are participation, policy coherence, reflexivity, adaptation and democratic institutions?" Earth System Governance **2**: 100031.
- Hassan, S.-U., P. Haddawy and J. Zhu (2014). "A bibliometric study of the world's research activity in sustainable development and its sub-areas using scientific literature." Scientometrics **99**: 549-579.
- Hassani, H., C. Beneki, S. Unger, M. T. Mazinani and M. R. Yeganegi (2020). "Text mining in big data analytics." Big Data and Cognitive Computing **4**(1): 1.
- Hwang, H., S. An, E. Lee, S. Han and C.-h. Lee (2021). "Cross-societal analysis of climate change awareness and its relation to SDG 13: A knowledge synthesis from text mining." Sustainability **13**(10): 5596.
- Kanie, N. and F. Biermann (2017). Governing through goals: Sustainable development goals as governance innovation, Mit Press.
- Körfgen, A., K. Förster, I. Glatz, S. Maier, B. Becsi, A. Meyer, H. Kromp-Kolb and J. Stötter (2018). "It's a hit! Mapping Austrian research contributions to the sustainable development goals." Sustainability **10**(9): 3295.
- Laureate, C. D. P., W. Buntine and H. Linger (2023). "A systematic review of the use of topic models for short text social media analysis." Artificial Intelligence Review: 1-33.
- Mesquita, R. F. d., B. Klein, A. Xavier and F. R. N. Matos (2017). "Mining and the sustainable development goals: a systematic literature review."
- Payumo, J., G. He, A. C. Manjunatha, D. Higgins and S. Calvert (2021). "Mapping collaborations and partnerships in SDG research." Frontiers in Research Metrics and Analytics **5**: 612442.
- Roy, A., A. Basu, Y. Su, Y. Li and X. Dong (2022). "Understanding Recent Trends in Global Sustainable Development Goal 6 Research: Scientometric, Text Mining and an Improved Framework for Future Research." Sustainability **14**(4): 2208.
- Salvia, A. L., W. Leal Filho, L. L. Brandli and J. S. Griebeler (2019). "Assessing research trends related to Sustainable Development Goals: Local and global issues." Journal of Cleaner Production **208**: 841-849.
- Shen, C., M. Wei and Y. Sheng (2021). "A bibliometric analysis of food safety governance research from 1999 to 2019." Food Science & Nutrition **9**(4): 2316-2334.

- Silge, J. and D. Robinson (2017). Text mining with R: A tidy approach, " O'Reilly Media, Inc."
- Takizawa, P. A. (2023). "Using a topic model to map and analyze a large curriculum." Plos one **18**(4): e0284513.
- United-Nations (2015). "Transforming our world: the 2030 Agenda for Sustainable Development." United Nations: New York, NY, USA.
- Vayansky, I. and S. A. Kumar (2020). "A review of topic modeling methods." Information Systems **94**: 101582.
- Zhang, Y., K. Huang, Y. Yu and B. Yang (2017). "Mapping of water footprint research: A bibliometric analysis during 2006–2015." Journal of Cleaner Production **149**: 70-79.
- Zhu, J. and W. Hua (2017). "Visualizing the knowledge domain of sustainable development research between 1987 and 2015: A bibliometric analysis." Scientometrics **110**(2): 893-914.