



Mapping scholarly publications related to the Sustainable Development Goals: Do independent bibliometric approaches get the same results?

Caroline S. Armitage^{iD}, Marta Lorenz^{iD}, and Susanne Mikki^{iD}

University of Bergen Library, University of Bergen, Bergen, Norway

an open access  journal



Keywords: bibliometrics, SDG, scientometrics, sustainability, Sustainable Development Goals

ABSTRACT

Many research and higher education institutions are interested in their contribution to achieving the United Nation's Sustainable Development Goals (SDG). Commercial services from Elsevier and Times Higher Education are addressing this by developing bibliometric queries for measuring SDG-related publications and SDG university rankings. However, such services should be evaluated carefully before use due to the challenging nature of interpreting the SDGs, delimiting relevance, and building queries. The aim of this bibliometric study was to build independent queries to find scholarly publications related to SDG 1, SDG 2, SDG 3, SDG 7, SDG 13, and SDG 14 using a consistent method based on SDG targets and indicators (the Bergen approach), and compare sets of publications retrieved by the Bergen and Elsevier approaches. Our results show that approach made a large difference, with little overlap in publications retrieved by the two approaches. We further demonstrate that different approaches can alter resulting country rankings. 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. Our results suggest that currently available SDG rankings and tools should be used with caution at their current stage of development.

1. INTRODUCTION

In 2015, the Sustainable Development Goals (SDGs) were adopted by the United Nations General Assembly as part of the 2030 Agenda for Sustainable Development. While their predecessors, the Millennium Development Goals, had success in tackling problems such as poverty, gender inequality, and disease, the SDGs were created with the reflection that more needed to be done “to integrate the economic, social and environmental aspects of sustainable development” (United Nations, 2015a, 2015b). The 17 SDGs therefore cover several broad themes and contain interrelated targets, aiming to stimulate action for people, planet, prosperity, peace, and partnership (United Nations, 2015b; Figure 1).

While the SDGs require political action, research and development of technology are also essential for achieving the goals. Research and technology are specifically mentioned in some of the targets (United Nations, 2015b) and are implicit in many others: For example, materials science is necessary for the development of clean, efficient energy systems (Chu, Cui, & Liu, 2017), relevant for SDG 7. Further selected examples include medical research for achieving targets under SDG 3, biological research for sustainable management and conservation of living resources (SDGs 14 and 15), and social sciences research to inform the governance aspects of all

Citation: Armitage, C. S., Lorenz, M., & Mikki, S. (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. https://doi.org/10.1162/qss_a_00071

DOI:
https://doi.org/10.1162/qss_a_00071

Supporting Information:
https://www.mitpressjournals.org/doi/suppl/10.1162/qss_a_00071

Received: 20 January 2020
Accepted: 29 April 2020

Corresponding Author:
Caroline S. Armitage
caroline.s.armitage@gmail.com

Handling Editor:
Ludo Waltman

Copyright: © 2020 Caroline S. Armitage, Marta Lorenz, and Susanne Mikki. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.





Figure 1. The UN's Sustainable Development Goals. Each goal consists of targets, giving 169 targets in total (the SDG icons and logo are property of the United Nations, reprinted for informational purposes in accordance with United Nations Department of Global Communications (2019); all rights reserved).

SDGs. Research collaboration and inter- and multidisciplinary research are also highly relevant and necessary for progress in sustainability (Leal Filho, Azeiteiro, et al., 2018; Reid, Bréchnac, & Tseh Lee, 2009). And although the process has its difficulties, scientific research provides objective input for political action via the science–policy interface (Gluckman, 2016). Research and higher education institutions can thus contribute to achieving the SDGs through research and other engagement (also see discussion in Körfgén, Förster, et al., 2018). The United Nations recognizes this via the United Nations Academic Impact initiative, stating that “The work of these institutions is vital to achieving the Sustainable Development Goals as they serve as incubators of new ideas, inventions and solutions to the many global challenges we face” (United Nations, 2019a).

Once research institutions commit to supporting the SDGs, they likely have an interest in examining their publishing output related to them. This may be exploratory; for example, which SDG-related research areas they focus on, how much they collaborate with specific partners, or to find relevant publications to make a policy brief. They may also wish to use the results to highlight their contribution to tackling societal and environmental problems, something that can help in building reputation and publicity. Some institutions are also interested in evaluation; the ability to present a high number in SDG-contribution rankings or otherwise evaluate how the institution compares to others (benchmarking). Interest in such questions is indicated by the development of commercial services for SDG rankings (for example the Times Higher Education [THE] University SDG Rankings) and services for finding and quantifying SDG-related publications, such as the SDG “research areas” in SciVal by Elsevier.

We believe that any services that measure or map SDG contributions should be evaluated carefully, because tools for ranking and quantification of research can have wide impacts. University rankings can influence strategic planning, institutional reorganization, and higher education priorities (Hazelkorn, 2009), and SciVal is advertised as allowing one to “visualize

your research performance, benchmark relative to peers, develop strategic partnerships, identify and analyse new, emerging research trends [...]” (Elsevier, 2019). These services thus have the potential to influence the evaluations, strategy, policy, and reputation of institutions, and thus have far-reaching impacts on research and society.

Both of the commercial services mentioned are based either entirely (SciVal) or partially (THE) on Boolean search queries containing SDG-related terms, which are then run against publication databases to find SDG-related scholarly publications. Similar approaches have been used in research papers by Jetten, Veldhuizen, et al. (2019), who identified academic publications concerning targets 2.1–2.4 of SDG 2, and Körfgen et al. (2018) who quantified research contributions to the SDGs from Austrian universities. However, there are four steps in the development of such queries that could vary greatly, and thus potentially have a large impact on any resulting insights: (a) interpretation of the themes and concepts of the SDG, (b) decisions around how publications must discuss these concepts to be considered as a “contribution” to the chosen interpretation of the SDG, (c) translation of concepts into a search query that will find contributing publications, and (d) data source.

The first step, interpretation, can be challenging, as the SDGs are discussed in different fora by different stakeholders and from different angles. They also have multiple titles (short and long versions) and have been translated into different languages, which can result in slightly different emphasis (for example, the English “Climate Action” vs. the Norwegian “Stopp Klimaendringer” [Stop climate changes]). What should be used as their basis when defining which themes and concepts are relevant to the SDG? In the second step, there is a further potentially subjective decision about how a publication must discuss these themes to be considered relevant to an SDG. Many of the targets concern an action (e.g., “end hunger”) as well as topics (e.g., “hunger”). Is a concrete, direct contribution to this action necessary (i.e., publications discussing ending hunger)? Or should indirect contributions be counted also (e.g., publications concerning crop technology)? Are some topics only relevant when discussed in conjunction with other topics? This also affects the third step: translation of these interpretations into search terms that will find relevant publications. Where should search terms originate from? The research community or policy documents? How should searches be structured to reflect the decisions made in the second step, and how should recall and precision be balanced? Finally, in the fourth step, there is the basis of the data used to measure research contribution. It is well known that publication databases vary in coverage of subject fields and languages (Aksnes & Sivertsen, 2019; Mongeon & Paul-Hus, 2016; Vera-Baceta, Thelwall, & Kousha, 2019), which will further affect any rankings and evaluations.

For these reasons, we believe it is essential to have an independent, transparent bibliometric method for finding publications related to the SDGs. This will allow institutions to compare different approaches, better understand where rankings come from, and evaluate how well a specific tool might work in their case. The aim of this study was therefore to build independent bibliometric search queries to find SDG-related scholarly publications, using a consistent and defined method that can be reused and built upon. Six SDGs were chosen for examination: SDG 1 (No poverty), SDG 2 (Zero hunger), SDG 3 (Good health and well-being), SDG 7 (Affordable and clean energy), SDG 13 (Climate action), and SDG 14 (Life below water). We then compared the results of our approach to the approach currently implemented in SciVal by Elsevier.

2. METHODS

2.1. Interpretation of the Themes of the SDGs and Which Publications “Contribute”

Each of the SDGs has targets (“Outcomes” and “Means of implementation”) and indicators. While the titles of some of the SDGs are relatively broad and open to subjective interpretation

(e.g., “Climate action”), the targets and indicators are much more specific about what should be achieved under the goal: They mention specific *actions* (e.g., “reduce”) and *topics* (e.g., “hunger,” “resilience,” and “tourism”; note that we use “topic” in a broad sense to also include states, characteristics, or activities). Using the SDG targets and indicators as the theoretical basis for interpretation thus allows some degree of objectivity, as one is working from a defined set of *topics* and *actions*. We therefore defined the themes of the SDGs based on targets and indicators. The list of SDG targets and indicators in Annex III of the IAEG-SDG 2017 Report to the UN Statistical Commission (Inter-agency Expert Group on SDG Indicators, 2017) was used, as this was the most recent complete list available at time of query development.

In terms of defining which research contributes to this interpretation of the SDG themes, we made a distinction between “direct contributions” and “indirect contributions.” We considered “direct contributions” to be publications that refer to *target concepts*—specific *topics* or *actions* in the targets or indicators. We considered “indirect contributions” to be publications that may be related to an SDG via *related concepts*—topics, actions, or research areas that are related to the *target concepts*. *Related concepts* are much more difficult to define and to limit than *target concepts*, as the latter can be defined from the targets and indicators themselves, while the former lacks an objective standard to base their inclusion upon—their inclusion would be much more based on a particular understanding of the concept and its relatedness. For this reason, we chose to build search queries aiming to find “direct contributions” only (Table 1, Example 1; Figure 2). This is not a judgment about the value of “indirect contributions” to achieving the SDGs but a practical approach, as defining them objectively and consistently across the SDGs is, in our opinion, extremely challenging. In addition, the “direct contribution” approach is well aligned with our method of interpreting the SDG themes. As the targets tend to be action oriented, this direct interpretation is unlikely to find basic/blue skies research where no application is discussed (Table 1, Example 2).

2.2. Translation of Interpretation into Search Queries

We took a Boolean search approach. Although this method of retrieval has been criticized, it maintains control and transparency (Hjørland, 2015), which was important both during query construction and for the current purpose of comparison. Some of the issues that may be criticized can also be avoided by careful construction of the search query (Hjørland, 2015), as outlined below. A Boolean approach has recently been used to map climate change, a broad interdisciplinary topic (Haunschild, Bornmann, & Marx, 2016).

The SDG targets mostly concern specific aspects of a topic or have a directional focus. In order to reflect this focus, the queries were often constructed with combinations of *topic terms* (e.g., “resilience” and “disaster”) together, or combinations of *topic terms* with *action terms* (e.g., “reduce,” “increase,” “establish,” “sharing”) (Table 1, Examples 3 and 4; Figure 2). Some terms were difficult to place into these types of terms (e.g., “policy,” which is both a way to bring about change and a *topic*), but they provided a rough framework for the construction of the queries. *Action terms* were not included on occasions when most works mentioning the topic(s) were considered likely to be directly relevant for the target, or when they would exclude much necessary research (Table 1, Example 5). A vocabulary reference list was used when adding *action terms* and commonly used *topic terms* to improve consistency across the SDGs.

The syntax of the queries was built with the aim of allowing maximum flexibility in phrasing and language use in publications while retaining specificity to the targets. Search terms were often truncated with wildcards, and combined with Boolean operators (AND, OR, very rarely NOT) in addition to the proximity operator NEAR (where NEAR/x allows x words between two

Table 1. Examples that illustrate our approaches to interpretation, definition of research contribution, or search query construction

Example	
Example 1	The FAO considers work on reducing food loss and waste to contribute to SDG2 (“Tackling food loss and waste is a defined target within the internationally agreed Sustainable Development Goals (SDG Target 12.3, which also contributes directly to SDG Target 12.5 and SDG Goal 2) and a key component of the Zero Hunger Challenge.” (Food and Agriculture Organization, 2019)). However, for our method of interpretation this link is indirect, as no “food waste”-related concepts are included in the targets or indicators of SDG 2. It was therefore not included in the search query; publications concerning food waste would only be found if they link this to topic’s concepts (e.g., ending hunger).
Example 2	Implementation of target 14.5 (“[...] conserve at least 10% of coastal and marine areas [...]”) will be supported and enhanced by basic research on marine biodiversity. However, according to our interpretation, this is an indirect contribution. Publications about marine conservation area establishment or management are more direct contributions.
Example 3	Target 14.3 (“minimize and address the impacts of ocean acidification [...]”) is not about ocean acidification generally; it is about the impacts. Our search query therefore requires that publications contain terms for “impacts” as well as “ocean acidification,” rather than just the latter.
Example 4	Target 2.5 concerns the maintenance of genetic diversity of agricultural resources and mentions “seed and plant banks.” The <i>topic terms</i> in the query were therefore expanded with “germplasm banks” and “gene banks,” as these have a similar function. These were combined with agricultural <i>topic terms</i> to limit the retrieved publications to those relevant to food production, rather than general conservation of genetic diversity. In the Bergen topic-approach, <i>action terms</i> such as “maintaining,” “preserving,” and “conserving” were removed, but the combination with agricultural terms was retained.
Example 5	Target 1.4 concerns access to economic resources and services, and “microfinance” is mentioned specifically. Since “microfinance” is a tool for providing access to financial services for low-income groups, it was not deemed necessary to combine this term with any <i>action terms</i> —the action relevant to the target is already inherent in the concept.
Example 6	Target 3.1 concerns reductions in child mortality (among other things). In the Bergen action-approach, the <i>topic terms</i> “mortality” and “child” were combined with <i>action terms</i> for “reduce”. In the Bergen topic-approach, the <i>action terms</i> were dropped; thus publications discussing child mortality in any context would be retrieved, regardless of whether they are about reducing it.

search terms). When combining related terms, use of the NEAR operator was preferred over combining terms in fixed phrases (e.g. “*sustainab**” NEAR/x “*aquaculture*” vs. “*sustainable aquaculture*”) to allow for language flexibility. Phrases were used for multiword concepts (e.g., “common fisheries policy”), or when a very close link between terms was required to exclude other subject areas.

One of the limitations of searching using terms in the text is that one must account for variations in the language use of authors. Although we attempted to be as flexible as possible in query construction, a danger with our interpretation and consequent use of *action terms* is that it may exclude publications that use unusual turns of phrase. Due to this, we developed two versions of the queries: one including all terms, hereafter referred to as the *Bergen action-approach* (BAA), and one where most of the *action terms* were removed and some of the NEAR terms combining *topic terms* loosened, hereafter the *Bergen topic-approach* (BTA). Combinations of *topic terms* were retained (Table 1, Examples 4 and 6). While the action-approach attempts to find literature that could directly contribute to achieving the SDG targets, the topic-approach finds literature related to the target concepts generally. Comparison of these two approaches allows examination of the effect of *action terms* and query structure.

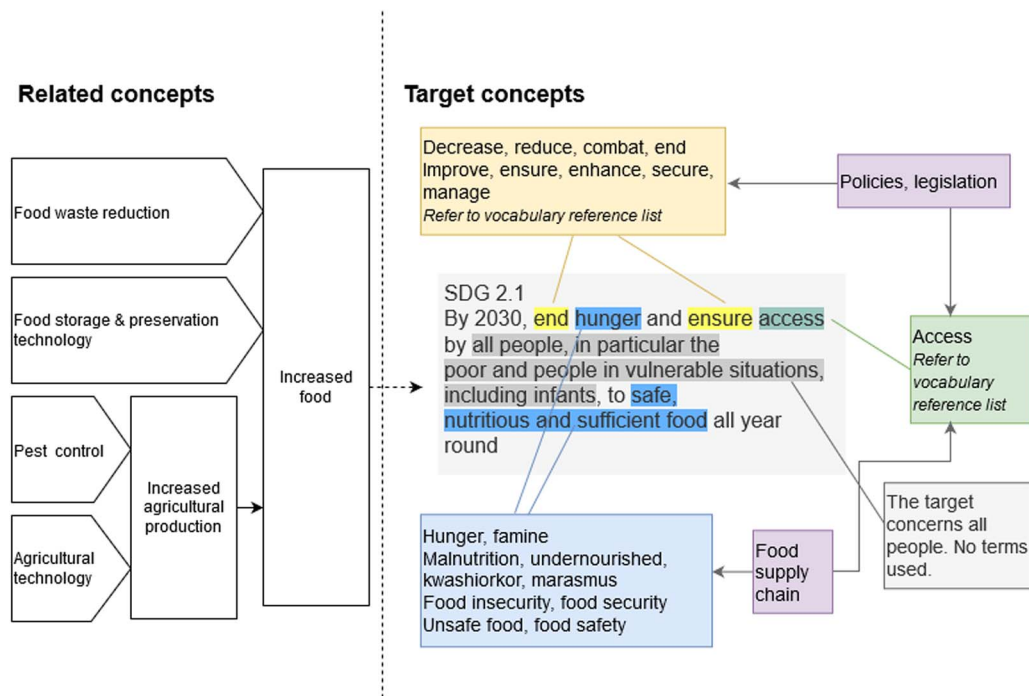


Figure 2. A diagram illustrating our interpretation of target 2.1 and the first steps in converting this into search terms. The *target concepts* section contains *topic* and *action terms* that are mentioned in the target. These are expanded with synonyms and subordinate topics (selected examples displayed). Some targets are aimed specifically at one particular group (e.g., the poor or small island developing states; gray highlighting); in these cases, these would be included in the query. Under *related concepts*, the elements contain research areas (e.g., pest control; arrows) that are not mentioned in the targets or indicators, but may contribute to a technology, knowledge, or state (e.g., increased food; boxes) that could potentially contribute to the target. These were not included explicitly in our queries.

Although the targets and indicators formed the basis of our interpretation of the SDGs, we did not limit our search queries to the terminology used in the targets. Terms were expanded considerably with synonyms and directly related/subordinate concepts (e.g. Table 1, Example 4; Figure 2). This allows retrieval of relevant publications even when authors do not explicitly relate their work to the SDGs. In addition to our own subject knowledge, we anchored our queries in terminology used by intergovernmental organizations and subject vocabularies, such as background notes from High-Level Political Forums on Sustainable Development (HLPF) (United Nations (ECESA plus), 2017a, 2017b, 2017c, 2017d; United Nations, 2018, 2019b) and resources from the United Nations, the World Health Organization (WHO), and the Food and Agriculture Organization (FAO). Controlled vocabulary thesauri were also examined to gather search terms, with Emtree® (Elsevier) used for SDG 2 and MeSH® (U.S. National Library of Medicine) for SDG 2 and SDG 3. Some of the targets needed particular attention because they concern categories (e.g., noncommunicable diseases, agriculture, LDCs), whereas specialist research publications are likely to refer to specific category members (e.g., melanoma, poultry, Angola). In these cases, the aforementioned resources were used as the basis for including category members as search terms. Details about the resources used in query development are included in the data files (Armitage, Lorenz, & Mikki, 2020).

The queries were developed to search in the title, abstract, and keywords of publications in a multidisciplinary database. Some of the *topic terms* in our queries are used across multiple academic fields; thus, they were used in combinations to limit the retrieved publications to the correct field. For example, for SDG 7 the term *energy* was used in specific combinations to

avoid results from biology or theoretical physics. For SDGs 2 and 3, terms also had to be used to limit some parts of the search to human-related studies. For SDG 14, where relevant terminology may apply to both marine and terrestrial environments, searching in the publication channel name (journal title) was used in addition to marine-related words to help limit the results to marine publications. The queries and notes on their development are included in the data files (Armitage et al., 2020).

2.3. Data Source

Web of Science Core Collection (Clarivate Analytics) (hereafter WoS) was used as the main database for testing and retrieving bibliographic information in the present study, as it allows advanced search functions and provides access to comprehensive article citation data for many different academic disciplines. Our searches were carried out using the search field *topic*, which searches in title, abstract, author keywords, and Keywords Plus® (plus publication channel name for part of SDG 14). WoS does not, however, cover all academic fields equally; natural sciences, technology, medicine, and health are well indexed, while the social sciences and particularly the arts and humanities are less well covered, in part due to its focus on articles (Aksnes & Sivertsen, 2019; Mongeon & Paul-Hus, 2016). We ran two analyses to estimate how well Nordic or Norwegian publications relevant to SDG 1 and SDG 3 would be covered, which suggested a rate of $\geq 93\%$ for SDG 3 and $\geq 79\%$ for SDG 1 (Supplementary Material 1). This should, however, be considered an upper estimate, given the location and high rate of English language publications in Norway and the Nordic countries.

To build the final corpus of literature, searching was limited to documents published for the years 2015 to 2018, as the SDGs were agreed upon in 2015 and implemented in January 2016 (United Nations, 2015b). The results were not restricted by language or document type; around 95–98% of publications were articles, while the rest were proceedings papers, corrections, and editorials.

2.4. Analysis

An evaluation of our BAA queries was done in two ways—for each SDG, we first visualized common keywords in 500 results, and second examined the titles and abstracts of 30 publications. The 500 publications for each SDG were chosen semirandomly by selecting 10 random search-result pages (each with 50 results) in WoS using a random number generator. The 30 publications were chosen from these 500 by assigning each publication a random number. The relevance of each publication was assessed by the three authors independently, and then discussed to get a consensus opinion on whether the publication was relevant, borderline relevant, or irrelevant. Keywords were visualized via network analysis in VOSViewer (Van Eck & Waltman, 2010) (full counting, keywords that occurred four or more times). Keywords were standardized by removing hyphens and using a thesaurus to combine forms of the same word.

To compare our results with the Elsevier queries, we downloaded the most recent Elsevier query version available at the time of analysis (October 10, 2019; Jayabalasingham, Boverhof, et al., 2019). This was translated into WoS syntax and run in the WoS database with the same search settings as for the Bergen queries. The number of publications found for each set of queries was compared on a worldwide basis. Overlaps between the different sets of results were assessed, and benchmarking of contributions from selected countries was done to see the effect of approach on country rankings. A comparison of keywords of publications unique to each approach was undertaken via network analysis in VOSViewer (Van Eck &

Waltman, 2010) (full counting, keywords that occurred five or more times in the 500 top cited publications).

3. RESULTS

The keyword analysis of the publications retrieved by the Bergen action-approach queries indicated that they are SDG-related (Supplementary Figures 1–6), as did the examination of 30 random publications for each SDG. The number of irrelevant publications retrieved (out of 30) was zero for SDG 3, one for SDGs 13 and 14, two for SDG 7 and three for SDGs 1 and 2. The number of borderline relevant publications retrieved was zero for SDG 7, one for SDGs 1 and 3, two for SDG 14, and three for SDGs 2 and 13 (list available from Armitage et al., 2020). The percentages of relevant publications (not including borderline cases) were therefore 97% (SDG 3), 93% (SDG 7), 90% (SDG 14), 87% (SDGs 1 and 13), and 80% (SDG 2). When including borderline cases, the lowest relevance rate was 90%, indicating high precision of our queries.

The queries from the different approaches varied greatly in the number of publications retrieved. The Bergen action-approach found many more publications than the Elsevier approach for SDG 1, but the Elsevier approach found many more publications for SDGs 3, 7, 13, and 14 (Figure 3). When the BTA was used, this picture reversed somewhat; the topic-approach found more publications than the Elsevier approach for SDGs 1, 2, 3, and 7, but fewer for SDGs 13 and 14 (Figure 4).

The degree of overlap in the publication sets found by the Bergen action-approach and the Elsevier approach was very low; under 25% of the total publications found for each SDG were found by both approaches (Jaccard similarity index expressed as a percentage; Figure 3). This was not solely due to the larger of the publication sets encompassing the smaller publication set: A considerable proportion of publications found by the smaller sets was not present in the larger sets (Figure 3). The overlap in publication sets for SDG 2 was particularly striking because although the number of retrieved publications was of the same order of magnitude for both approaches, overlap was still low.

When the results from the BTA and Elsevier approach were compared, the degree of overlap was mostly larger than for the action-approach comparison. Nevertheless, it remained low (Figure 4). In particular, the SDG 1 overlap became even smaller due to the small number of Elsevier publications for this SDG. SDG 3 overlap rose to 56% (1,188,098 publications, the highest overlap of all the SDGs and comparisons); however, this still means that almost a million publications were only found by one of the two approaches (491,778 publications unique to the BTA and 425,090 unique to the Elsevier approach).

An analysis of keywords in the publication sets uniquely retrieved by either the BTA or the Elsevier approach was done for SDGs with the least overlap (SDGs 1, 2, and 14). While this cannot tell us whether the publications can be considered direct contributions to the SDGs (regarding our interpretation in Section 2.1), it can indicate whether they include *target concepts* or not. Because the focus here was on keywords and thus relatively wide, the BTA was used for the comparison. For SDG 1, both approaches had relatively clear clusters and mostly relevant concepts (although detailed examination of how keywords are combined would be needed to check relevance for some, e.g., “health”; Supplementary Figure 7). For SDGs 2 and 14, the Elsevier approach produced a less clear clustering pattern and had some themes that may not be relevant, according to our interpretation (although again, usage context would need to be examined). For SDG 2, the main themes in the BTA-unique set included food safety, obesity & weight, and agriculture & climate change, and biodiversity; the main themes

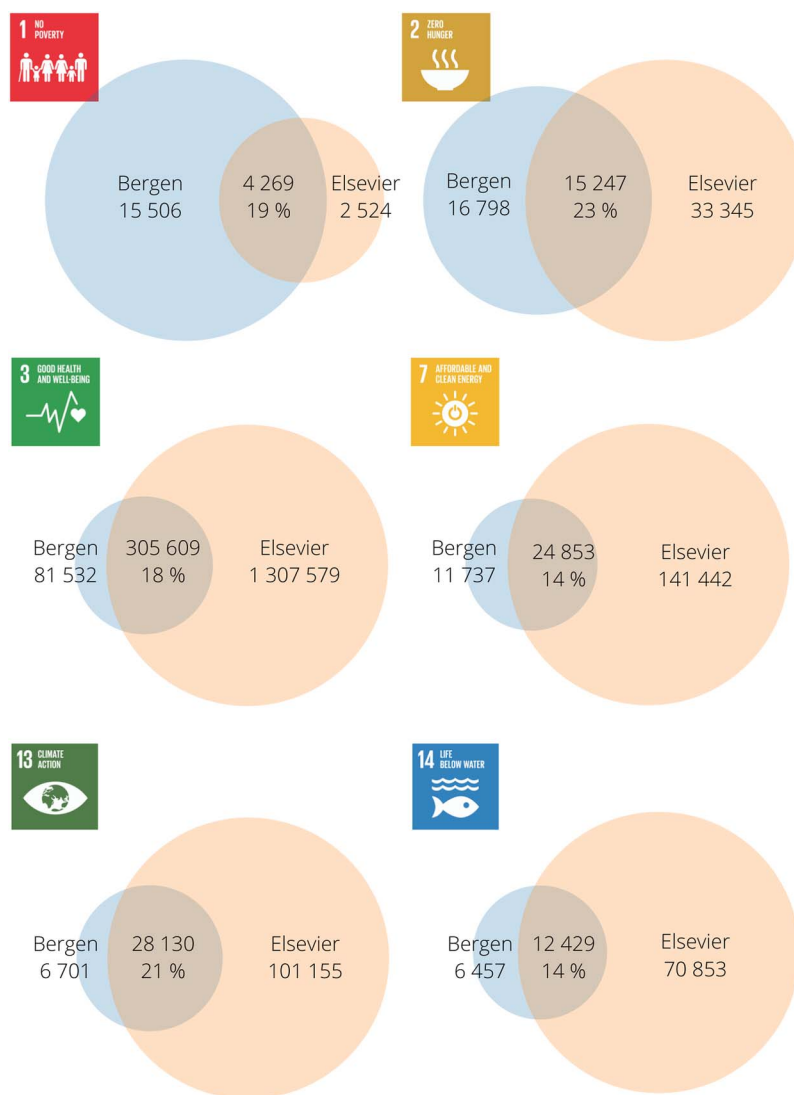


Figure 3. Venn diagrams showing the overlap in publications found by the Bergen action-approach queries and the Elsevier queries (number of publications found only by the Bergen or Elsevier queries is shown below the respective label, together with percentage (Jaccard similarity index) and number of shared publications in the middle). Publications retrieved from Web of Science Core Collection, publication years 2015–2018, all document types and languages.

in the Elsevier-approach-unique set included agriculture & climate change, heavy metals, and fertility & nutrients (Supplementary Figure 8). For SDG 14, the main themes in the BTA-unique set included marine pollution, climate change & ecosystem services, and fisheries & management & impacts; the main themes of the Elsevier-approach-unique set included climate change, temperature, evolution & diversity, and marine sediments & communities (Supplementary Figure 9).

The different approaches also made a difference to the country rankings. While the major contributors were often similarly ranked between the Bergen action-approach and Elsevier approach, this was not always the case, with large research nations such as China, the UK, and Australia swapping positions for SDGs 2, 13, and 14 (Figure 5). There was no SDG where

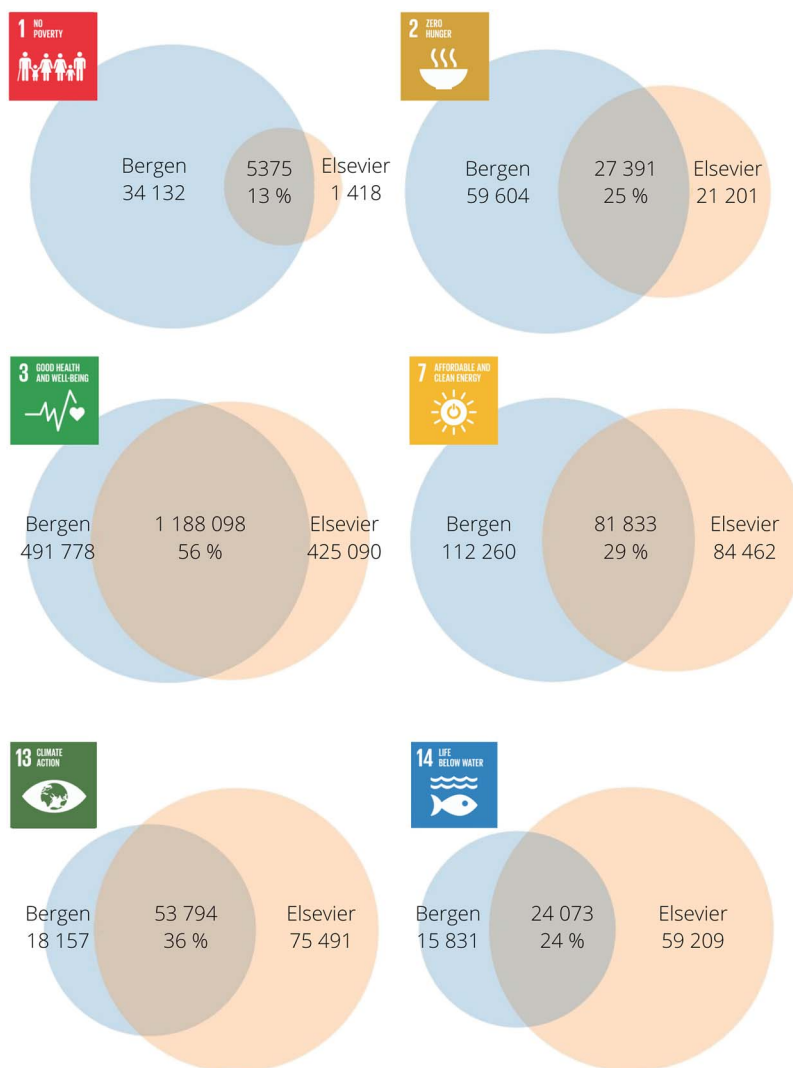


Figure 4. Venn diagrams showing the overlap in publications found by the Bergen topic-approach queries and the Elsevier queries (number of publications found only by the Bergen or Elsevier queries is shown below the respective label, together with percentage (Jaccard similarity index) and number of shared publications in the middle). Publications retrieved from Web of Science Core Collection, publication years 2015–2018, all document types and languages.

the countries remained in the same ranking order for both approaches; the smallest change was for SDG 3 (three countries changing position, less than 2% difference) and the largest for SDG 14 (nine countries changing position, four cases of over 2% difference; Figure 5). The top 10 countries for each SDG remained the same except for SDGs 7 and 14 (SDG 7: Japan 15th vs. 9th, Canada 10th vs. 11th; SDG 14: Japan 16th vs. 10th, Brazil 10th vs. 11th). The Bergen action-approach also found a larger contribution from LDCs to the total SDG-related publication set than the Elsevier approach. Across the six SDGs, Elsevier’s approach found a lower percentage contribution from the top-ranked LDCs than the Bergen action-approach for 53 of the 60 ranking places (Supplementary Figure 10). This difference was particularly large for SDGs 7, 13, and 14 (combined percentage contribution from the top 10 LDCs: SDG 7: BAA = 1%, Elsevier = 0.4%; SDG 13: BAA = 2.4%, Elsevier = 1.6%; SDG 14: BAA = 1.6%, Elsevier = 0.6%).

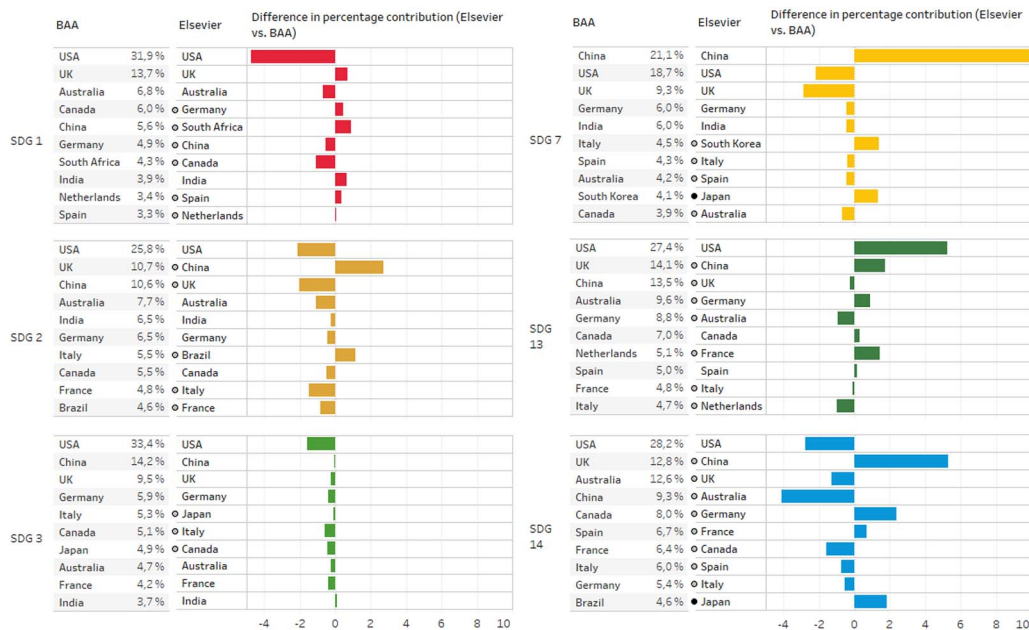


Figure 5. Comparison of percentage contribution of countries to the SDGs between the Bergen action-approach (BAA) and Elsevier queries. The top 10 countries in order of contribution (highest to lowest) are listed for each approach, with dots indicating a difference in rank (black = into the top 10). The percentages are the percentage contribution according to BAA (SDG-related publications from that country as a percentage of the total number of SDG-related publications), while the charts show the difference in each country’s (Elsevier top 10) percentage contribution when comparing the Elsevier percentage to the BAA percentage (e.g., for SDG 1, the USA’s percentage contribution was around 5% lower in the Elsevier results). Publications retrieved from Web of Science Core Collection, publication years 2015–2018, all document types and languages.

4. DISCUSSION

This work indicates that methodological approach makes a large difference to the publication set retrieved, and by extension can affect subsequent rankings. This demonstrates the importance of independent evaluation and open methodology, and has implications for the use of commercial services currently offering SDG-related rankings or bibliometric data. Elsevier itself is relatively open about its queries being under development, but this is not clear in the SciVal platform itself (as of November 2019). We could also not find any caveats about the THE queries (Times Higher Education, 2019). The danger with this is that their services’ results could be used uncritically, as they provide easy access to rankings and data for leaders, administrators, and research managers. University rankings are used in important decisions despite their demonstrated shortcomings (Hazelkorn, 2009; and discussion in Schmoch, 2015), and they have additional issues when it comes to the SDGs (Torabian, 2019). Rankings and bibliometric information based on unreliable data foundations have even more potential for adverse consequences.

4.1. Comparison of Approaches

In the development of the current version of the Elsevier queries, Jayabalasingham et al. state that they compared their original search queries to the targets of the SDGs to assess relevance (Jayabalasingham et al., 2019; Mu & James, 2019). Given that the Bergen approach is also based on the targets, it is surprising that there was so little overlap in the results of these approaches. The differences seem to reflect different choices made in the first three steps of the

process—interpretation of the themes of the SDG, decisions about how a theme must be discussed in a publication to be considered “relevant,” and how these themes are translated into search terms. In other words, it is not just the practical differences in query construction, but also differing perspectives on which concepts, and which combinations of concepts, are “relevant” for a target. Deeper examination of the Bergen and Elsevier queries revealed that there are large differences between the approaches in query structure and the way that terms are combined, the use of *action terms*, and differences in which *topic terms* are included.

The Elsevier queries combine topics relatively rarely, making them much broader than even the BTA queries (an effect which is further magnified when *action terms* are combined with *topic terms* in the Bergen action-approach). A particularly striking example is SDG 14, where the topic “marine” is used in two parts of the Elsevier query (combined with AND), with the result being that a publication only has to use the word “marine” to count as SDG 14-related. For our interpretation of SDG 14 relevance, this is far too broad, as SDG 14 is not concerned with everything to do with the marine environment. The effect of its inclusion is indicated by the keyword analysis, where keyword clusters not obviously closely related to the targets of SDG 14 appear (e.g., marine viruses/genomics/evolution, marine sediments/molecular phylogeny; Supplementary Figure 9). Similarly, using the phrase *climate change* is sufficient to be counted as related to SDG 13 in the Elsevier approach, whereas the Bergen approaches required this term to be used in combination with terms with which it is combined in the targets (e.g., climate change adaptation, mitigation). Another example from SDG 2 is the topic “agricultural production.” For the Elsevier approach, this phrase alone is enough to result in inclusion. For the Bergen approaches, terms for this topic had to be used with terms for small-scale food production or sustainability, because increased agricultural production is referred to in these specific contexts in the targets (targets 2.3 and 2.4).

Combinations of *topic terms* were not the only difference between the approaches; in many cases there were differences in which *topic terms* were included at all. The Elsevier approach includes some *topic terms* that were not considered part of the SDGs by the Bergen approaches; for example, “ocean circulation modelling” for SDG 14 and “fertiliser” for SDG 2. This likely explains the presence of keyword clusters in the Elsevier-approach-exclusive results related to temperature variability and modeling for SDG 14 and fertilizer/nutrients for SDG 2 (Supplementary Figures 8 and 9). Similarly, the Bergen approaches include many *topic terms* that are not included in the Elsevier queries. In particular, when developing the Bergen approaches we tried to include category members as terms when categories were used in the targets, as we are aware that researchers are often publishing for a specialised audience and thus use specific terminology. For example, the SDG targets might refer to “neglected tropical diseases” or “least developed countries,” but a publication might talk about “dengue” or “Ghana.” By only using the category as a search term, such publications are likely to be excluded unless broad keywords have been added.

The use of *action terms* clearly also makes a large difference to the number of publications retrieved. The Bergen action-approach retrieved a much smaller number of publications than the BTA, despite the same *topic terms* being used. These *action terms* make a very large difference in some cases; for example, searching for “poverty” without *action terms* (e.g., “reduc*” OR “decreas*”) may find many publications where poverty is a causative factor (e.g., poverty-associated diseases) but do not directly discuss combating it.

Finally, the use of operators is also likely to be responsible for some of the differences in retrieved publications between the Elsevier and Bergen approaches. To combine terms, Elsevier used the operator AND (and sometimes NOT), while the Bergen approaches used NEAR much

more often. Regardless of any differences in *topic terms*, this simultaneously broadens and limits the results of Elsevier approach: Combining terms with AND instead of NEAR makes the search wider, but using phrases instead of NEAR may result in missing relevant publications. For example, the use of “*extreme poverty*” OR “*poverty alleviation*” OR “*poverty eradication*” OR “*poverty reduction*” excludes any works discussing “ending poverty,” “eradicating poverty,” or “eliminating child poverty.” On the other hand, combinations used in the Bergen approaches which allow flexibility can permit some nonrelevant results (e.g., “*economic*” NEAR/10 “*control*” AND..., part of the Bergen queries for SDG 1, results in some publications from cancer studies due to “control” and “economic” used in two unrelated sentences). This shows that it is not just the *topic terms* that must be agreed upon for a bibliometric tool for measuring SDG-related output, but also how *topic terms* are combined, and whether *action terms* are also necessary for increasing precision and the inclusion of relevant publications.

4.2. Other Approaches

In this study we have demonstrated the differences in results produced by our approaches and Elsevier’s approach; however, there have also been other approaches to the same problem. The Sustainable Development Solutions Network (SDSN; Australia, New Zealand, and Pacific), the AURORA network (AURORA Universities Network, n.d.), and SIRIS Academic (Duran-Silva, Fuster, et al., 2019) have also been working on keyword-based bibliometric queries to find literature related to the SDGs. A comparison of the SDSN and Elsevier queries was shown during a webinar on Elsevier’s Research Intelligence channel (BrightTalk) and indicated that the SDSN queries take a much broader approach (e.g., “cities AND land” for SDG 11) and find many more publications than the Elsevier queries (except for SDG 3). In that comparison, the Elsevier queries were characterized as a very focused approach aiming to minimize false positives by using very precise keywords (Mu & James, 2019). However, this characterization is not reflected in the results of the present study.

We were not granted access to the detailed methodology for the Times Higher Education (THE) approach. While the THE approach is developed in partnership with Elsevier (and Vertigo Ventures; Times Higher Education, 2019), the queries being used by THE (at present) are not the same as those from Elsevier currently being used in SciVal, and appear instead to be based on an earlier version (Mu & James, 2019). We were not able to ascertain if the original Elsevier queries have been further modified by THE. We could therefore not do a comparison with their approach. However, the original version of the Elsevier queries, according to Elsevier’s own testing and documentation, finds many publications that are not relevant to the SDG targets (Jayabalasingham et al., 2019; Mu & James, 2019). This suggests that if THE is using these original queries, then their results might diverge from our approach even more so than Elsevier’s. In addition, the specifics of how THE integrate their data on SDG-related scholarly publications with their other criteria for assessment is relatively unclear from their public methodology (Times Higher Education, 2019). University rankings can have wide reaching impacts (Hazelkorn, 2009), and, in THE’s own words, “The data we compile to produce these rankings are trusted by governments and universities and are a vital resource for students when they are making decisions about where to study” (Times Higher Education, 2018). It is therefore unfortunate that we could not confirm their method of evaluation well enough to compare results.

Boolean searching is not the only approach taken to examining SDG-related publication output. Machine learning has been suggested as an approach (Mu & James, 2019), but to our knowledge has not been applied yet. Citation analysis has also been used, where a body of SDG-related literature is built from publications who cite publications that use the phrase *Sustainable*

Development Goal(s) (Nakamura, Pendlebury, et al., 2019). A potential danger with this is that one risks missing works from researchers who are not engaged with the SDGs despite working on relevant topics. However, it does avoid some of the issues with interpretation of SDG themes and relevance by shifting the responsibility for this to the authors of the publications referring to the SDGs.

4.3. Challenges

We carried out our comparisons within the same database using the same settings; these should therefore be valid. However, the number of publications and country rankings presented here will be affected by the coverage of the WoS database. Our analyses indicated that we may be missing around 7% of the relevant medical literature and 20% of the social sciences publications (Supplementary Material 1). These percentages are likely an underestimate when considering global publications, given the underrepresentation of non-English-language publications in large international databases. In addition, the WoS database focuses strongly on articles. Measuring other publication types (e.g., theses, reports) would require a different data source.

To make our searches both flexible and specific, we used proximity operators and a complex structure. This means that to run these searches as they are, one is limited to databases which support advanced search structure (e.g., WoS, Scopus). However, the queries do not rely on controlled vocabularies, thesauri, or subject indexing, and are designed to prevent fuzzy searching (lemmatization); therefore, the queries should function the same in any database (provided proximity operators are supported and queries are translated into the appropriate syntax). The queries can also function as a resource for gathering SDG-related keywords.

Interpreting the SDGs from the targets was not without its challenges. The SDG targets vary in formulation and degree of specificity, with a report by the International Council for Science assessing only 29% of the targets as “well-developed” (ICSU & ISSC, 2015). In addition, the targets vary in how directly they can be linked to research activities, and interpretation may be impeded by differences in the title of the goal, HLPF documents, and targets. An example of these last two points is SDG 13. Here there is a focus on combating climate change in both the long title of the goal (“Take urgent action to combat climate change and its impacts”) and the HLPF background note (which states that addressing climate change impacts “require a two-pronged approach—reduction in the greenhouse gas emissions, and adaption planning” [United Nations, 2019b]). This implies that research and technological development for reducing greenhouse gases (e.g., carbon capture and storage) would be relevant. However, the targets are mostly focused on the political perspective (funding and implementing of policies for climate change measures [13.A, 13.2] or educating and improving capacity for climate change adaption, mitigation, etc. [13.3]). Reduction of greenhouse gases is mentioned only in relation to national policies or plans (indicator 13.2.1) and in terms of improving human and institutional capacity for mitigation (13.3). Interpreting the relevance of research to target 13.3 therefore becomes a question of what contributes to “human and institutional capacity,” which can be relatively subjective (the UN definition of capacity is “[...] the ability of people, organizations and society as a whole to manage their affairs successfully” [United Nations Development Group, 2017]). In addition, the use of terms such as “impact reduction” makes it relatively unclear whether impacts on nonhuman entities (ecosystems, plants, etc.) are relevant; we interpreted it in the broadest sense.

Another challenge concerning relevance comes from the use of search terms. As publications are retrieved based on their keywords, abstract, and title, they may be retrieved or not retrieved based on author awareness of relating their work to wider issues. For example, it is

not uncommon for publications about renewable energy (SDG 7) to have a sentence in the abstract relating their work to climate mitigation, and thus may be retrieved by the SDG 13 search. This means that differences in writing style and strategic positioning of buzzwords may affect representation in the retrieved publication set. In the same way, the results of this particular study will be affected by the WoS process of applying Keywords Plus to publications, because if our search found a publication relevant based on Keywords Plus, this has been mediated by WoS's interpretation of that concept and whether it was appropriate for the publication.

Finally, we emphasize that the approach outlined here is limited to measuring scholarly publishing. Assessing institutions' wider engagement with the SDGs requires an integrated approach which takes into account a variety of institutional activities (e.g., courses offered in sustainability, measures to reduce inequality, climate-friendly policies), not least because the pursuit of certain indicators or rankings (related to publishing) may encourage practices that are in conflict with the SDGs themselves (see Torabian, 2019).

4.4. Summary and the Way Forward

Interpretation of the themes of the SDGs, making decisions about what counts as a "contribution," and translating this into functioning search queries are not simple tasks. This study has shown that two independent approaches can deliver two widely different sets of results. Differences in the terms included and how they are combined makes large differences to the final result. The results suggest that it would be premature to trust commercial SDG analyses for anything other than exploratory purposes at this stage in their development.

The open methodology of the Elsevier approach facilitated this comparison and demonstrates how open science can facilitate independent testing and potentially stimulate the advancement of methods and tools. According to Mu and James (2019), the next stage of the Elsevier approach is to develop a website where people can give feedback on whether a document addresses the SDG or not, and, if yes, which of the targets it refers to. This type of crowdsourcing may further development, but due to the wide variety of perspectives involved, crowdsourcing may not be ideal for developing a consensus on interpretation of SDG themes and making decisions about which research "contributes." We suggest that developing multiple approaches to reflect different perspectives and uses (e.g., our action- and topic-approaches), perhaps with an additional even-wider approach to cover indirect contributions, could be beneficial. This would allow for multiple interpretations of what it means to be "SDG-related research," allow researchers and managers to use the appropriate tool for their needs at that time, and allow bibliometric measurements to cover the diversity in SDG-related research.

ACKNOWLEDGMENTS

We thank the reviewers for their comments, and our University of Bergen colleagues Regina K ufner Lein for help developing our SDG 3 query and Katja Enberg for input on our SDG 14 query.

AUTHOR CONTRIBUTIONS

Caroline Armitage: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing—original draft, Writing—review & editing.
Marta Lorenz: Conceptualization, Formal analysis, Investigation, Methodology, Validation,

Writing—review & editing. Susanne Mikki: Conceptualization, Methodology, Project administration, Validation, Writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

No funding has been received for this research.

DATA AVAILABILITY

The queries developed in this study, and the data for the testing of 30 random publications, are available in DataverseNO (UiB Open Research Data) at <https://doi.org/10.18710/98CMDR>. Metadata directly from databases (used to create keyword analyses and compare coverage) are not available.

REFERENCES

- Aksnes, D. W., & Sivertsen, G. (2019). A criteria-based assessment of the coverage of Scopus and Web of Science. *Journal of Data and Information Science*, 4(1), 1–21. <https://doi.org/10.2478/jdis-2019-0001>
- Armitage, C. S., Lorenz, M., & Mikki, S. (2020). *Replication data for: Mapping scholarly publications related to the Sustainable Development Goals: Do independent bibliometric approaches get the same results?* [Data set]. DataverseNO. <https://doi.org/10.18710/98CMDR>
- AURORA Universities Network. (n.d.). *SDG analysis: Bibliometrics of relevance*. Retrieved March 30, 2020, from <https://aurora-network.global/project/sdg-analysis-bibliometrics-relevance/>
- Chu, S., Cui, Y., & Liu, N. (2017). The path towards sustainable energy. *Nature Materials*, 16(1), 16–22. <https://doi.org/10.1038/nmat4834>
- Duran-Silva, N., Fuster, E., Massucci, F. A., & Quinquillà, A. (2019). A controlled vocabulary defining the semantic perimeter of Sustainable Development Goals (Version 1.2) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.3567769>
- Elsevier. (2019). *SciVal*. Retrieved December 9, 2019, <http://www.elsevier.com/solutions/scival>
- Food and Agriculture Organization. (2019). *Policy Support and Governance; Food Loss and Food Waste*. Retrieved October 20, 2019, from <http://www.fao.org/policy-support/policy-themes/food-loss-food-waste/en/>
- Gluckman, P. (2016). The science-policy interface. *Science*, 353(6303), 969. <https://doi.org/10.1126/science.aai8837>
- Haunschild, R., Bornmann, L., & Marx, W. (2016). Climate Change Research in View of Bibliometrics. *PLOS ONE*, 11(7), e0160393. <https://doi.org/10.1371/journal.pone.0160393>
- Hazelkorn, E. (2009). Rankings and the battle for world-class excellence. *Higher Education Management and Policy*, 21(1), 1–22. <https://doi.org/10.1787/hemp-v21-art4-en>
- Hjørland, B. (2015). Classical databases and knowledge organization: A case for boolean retrieval and human decision-making during searches. *Journal of the Association for Information Science and Technology*, 66(8), 1559–1575. <https://doi.org/10.1002/asi.23250>
- ICSU & ISSC. (2015). *Review of Targets for the Sustainable Development Goals: The Science Perspective*. Retrieved from <http://council.science/publications/review-of-targets-for-the-sustainable-development-goals-the-science-perspective-2015>
- Inter-agency Expert Group on SDG Indicators. (2017). *Report of the Inter-agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2017.2)*. United Nations. <http://unstats.un.org/unsd/statcom/48th-session/documents/2017-2-IAEG-SDGs-E.pdf>
- Jayabalasingham, B., Boverhof, R., Agnew, K., & Klein, L. (2019). *Identifying research supporting the United Nations Sustainable Development Goals*. Mendeley Data, v1 (2019-10-10). [Data set]. <https://doi.org/10.17632/87txkw7khs.1>
- Jetten, T. H., Veldhuizen, L. J. L., Siebert, M., & Darroch, P. I. (2019). *An explorative study on a university's outreach in the field of UN Sustainable Development Goal 2* [Research report]. Retrieved from Wageningen University & Research: <https://doi.org/10.18174/476199>
- Körffgen, A., Förster, K., Glatz, I., Maier, S., Becsi, B., Meyer, A., Kromp-Kolb, H., & Stötter, J. (2018). It's a Hit! Mapping Austrian Research Contributions to the Sustainable Development Goals. *Sustainability*, 10(9), 3295. <https://doi.org/10.3390/su10093295>
- Leal Filho, W., Azeiteiro, U., Alves, F., Pace, P., Mifsud, M., Brandli, L., Caeiro, S. S., & Disterheft, A. (2018). Reinvigorating the sustainable development research agenda: The role of the sustainable development goals (SDG). *International Journal of Sustainable Development & World Ecology*, 25(2), 131–142. <https://doi.org/10.1080/13504509.2017.1342103>
- Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213–228. <https://doi.org/10.1007/s11192-015-1765-5>
- Mu, J., & James, C. (2019). *Analyzing UN SDG research with SciVal—The case of University of Auckland* [Webinar]. Retrieved from BrightTALK: <https://www.brighttalk.com/webcast/13819/375997>
- Nakamura, M., Pendlebury, D., Schnell, J., & Szomszor, M. (2019). *Navigating the Structure of Research on Sustainable Development Goals*. Retrieved from the Institute for Scientific Information: <https://clarivate.com/webofsciencegroup/campaigns/sustainable-development-goals/>

- Reid, W. V., Bréchnignac, C., & Tseh Lee, Y. (2009). Earth system research priorities. *Science*, 325(5938), 245. <https://doi.org/10.1126/science.1178591>
- Schmoch, U. (2015). The informative value of international university rankings: Some methodological remarks. In *Incentives and Performance: Governance of Research Organizations* (pp. 141–154). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-09785-5_9
- Times Higher Education. (2018). *How to participate in Times Higher Education rankings*. Retrieved December 12, 2019, from <https://www.timeshighereducation.com/how-participate-times-higher-education-rankings>
- Times Higher Education. (2019). *THE University Impact Rankings 2019: Methodology*. Retrieved December 12, 2019, from <https://www.timeshighereducation.com/world-university-rankings/methodology-impact-rankings-2019>
- Torabian, J. (2019). Revisiting Global University Rankings and Their Indicators in the Age of Sustainable Development. *Sustainability: The Journal of Record*, 12(3), 167–172. <https://doi.org/10.1089/sus.2018.0037>
- United Nations. (2015a). *The Millennium Development Goals Report 2015 Summary*. [https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG 2015 Summary web_english.pdf](https://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%2015%20Summary%20web%20english.pdf)
- United Nations. (2015b). *Transforming our world: The 2030 Agenda for Sustainable Development*. Retrieved December 2, 2019, from <https://sustainabledevelopment.un.org/post2015/transformingourworld>
- United Nations. (2018). *2018 HLPF Review of SDG implementation: SDG 7- Ensure access to affordable, reliable, sustainable and modern energy for all*. https://sustainabledevelopment.un.org/content/documents/195532018_background_notes_SDG_7Final1.pdf
- United Nations. (2019a). *Academic Impact; About UNAI*. Retrieved October 14, 2019, from <http://academicimpact.un.org/content/about-unai>
- United Nations. (2019b). *Review of SDG implementation and inter-relations among goals. Discussion on SDG 13—Climate action*. https://sustainabledevelopment.un.org/content/documents/23627BN_SDG13.pdf
- United Nations (ECESA plus). (2017a). *2017 HLPF Thematic Review of SDG 1: End poverty in all its forms everywhere*. https://sustainabledevelopment.un.org/content/documents/14379SDG1format-final_OD.pdf
- United Nations (ECESA plus). (2017b). *2017 HLPF Thematic review of SDG2: End hunger, achieve food security and improved nutrition, and promote sustainable agriculture*. https://sustainabledevelopment.un.org/content/documents/14371SDG2_format_revised_FINAL_28_04.pdf
- United Nations (ECESA plus). (2017c). *2017 HLPF Thematic Review of SDG3: Ensure healthy lives and promote well-being for all at all ages*. https://sustainabledevelopment.un.org/content/documents/14367SDG3format-rev_MD_OD.pdf
- United Nations (ECESA plus). (2017d). *2017 HLPF Thematic Review of SDG 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development*. <http://sustainabledevelopment.un.org/content/documents/14375SDG14format-revOD.pdf>
- United Nations Department of Global Communications. (2019). *Sustainable Development Goals: Guidelines for the use of the SDG logo including the colour wheel, and 17 icons*. Retrieved from <https://www.un.org/sustainabledevelopment/news/communications-material/>
- United Nations Development Group. (2017). *Capacity Development, UNDAF companion guidance*. Retrieved from <https://unsdg.un.org/resources/capacity-development-undaf-companion-guidance>
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Vera-Baceta, M. A., Thelwall, M., & Kousha, K. (2019). Web of Science and Scopus language coverage. *Scientometrics*, 121(3), 1803–1813. <https://doi.org/10.1007/s11192-019-03264-z>