https://doi.org/10.51408/issi2025_157

Mapping National Research That Targets Sustainable Development Goals: The Responsible Visualization of Openalex Data for Societal Impact Measurements of Research

Robin Haunschild[1], Lutz Bornmann[2]

[1]*R.Haunschild@fkf.mpg.de*
Max Planck Institute for Solid State Research, Information Service,
Heisenbergstr. 1, 70569 Stuttgart (Germany)

[2]*L.Bornmann@fkf.mpg.de, bornmann@gv.mpg.de*
Science Policy and Strategy Department, Administrative Headquarters of the Max Planck Society,
Hofgartenstr. 8, 80539 Munich (Germany)

Abstract

As one approach to a framework for the responsible use of bibliometrics in research evaluation, we proposed to use global maps using OpenAlex to highlight concepts (i.e., fields of research) where countries are particularly active to achieve United Nations sustainable development goals (UN SDGs). As first examples in this research-in-progress paper, we used the USA and Japan. As to be expected, we found that the USA is very active in many concepts to achieve the SDGs (since the USA is very research active in general). We revealed for Japan that the country has two increased areas of activity to achieve the SDGs: One area is in Medicine and the other in Chemistry and Material Sciences. Our SDG mapping approach combines multiple aspects of the responsible use of bibliometrics in research evaluation: (1) By focusing on SDG relevant research, we provide an innovative approach for measuring target-oriented the societal impact of research. (2) Our approach goes beyond using simple counting of publications or citations by using maps to display complex results. (3) The usage of OpenAlex and free statistics software makes our procedure transparent and reproducible.

Introduction

In recent years, some initiatives have been started with the goal of reforming the way research is assessed (Rushforth & Hammarfelt, 2023). The initiatives include the Leiden Manifesto (Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015), the Metric Tide (Wilsdon et al., 2015), the Declaration on Research Assessment (DORA, https://sfdora.org/), and the Agreement on Reforming Research Assessment (CoARA, https://coara.eu). Whereas DORA focuses on reducing the use of journal-based citation impact indicators in research assessments, CoARA emphasizes the need for a more holistic approach to research evaluation (Thelwall, 2024). In this research-in-progress paper, we took up this call for a more holistic approach by introducing science maps visualizing national research that targets United Nations sustainable development goals (UN SDGs, https://sdgs.un.org/goals). The maps are intended to highlight the areas in which national research targets (worldwide) societal challenges. Most of the previous maps have focused on the visualization of traditional metrics, i.e., citation impact of publications.

In 2000, the UN established six Millennium Development Goals and in 2015, adopted the 2030 Agenda, which includes 17 interconnected SDGs. The Agenda

outlines an action plan for people, planet, and prosperity. At the Stockholm conference in 2022 (https://www.stockholm50.global), proposals were made to accelerate the achievement of the 2030 Agenda, focusing on SDGs for a healthy planet, social and economic progress, well-being, and resilience (Hernandez, Suazo López, & Domínguez Pacheco, 2022). It is one important goal of the science system to encompass societal products (outputs), societal use (societal references), and societal benefits (changes in society). It has been argued that society only reaps benefits from successful scientific studies when their results are converted into products (e.g., medications, diagnostic tools, machines, or devices) or services (e.g., government advising) (Bornmann, 2012, 2013). In recent years, some studies were published that have investigated whether scientific studies not only have societal impact but also specifically address SDGs (Ciarli, 2022; Purnell, 2022).

Using data from the OpenAlex database, we propose in this study overlay maps that visualize the national research that is especially active in worldwide SDG-relevant research. These overlay maps are visual tools used to represent the relationships and positions of national data within the worldwide scientific landscape. The maps overlay national data onto a base map that represents the entire science system. This helps to visualize how the national data fit into the larger context of scientific research. To demonstrate the overlay maps in this study, we present the maps for the USA and Japan.

Contribution of global overlay maps using OpenAlex to responsible bibliometric practices

In the development of the global overlay maps technique presented in this paper, we tried to follow the various guidelines for the responsible use of bibliometrics. The Leiden Manifesto (Hicks, et al., 2015) presents ten principles to guide research evaluation. The fourth principle suggests to use open data to foster transparency in research evaluation. CoARA also calls for the use of open datasets for and transparency in research evaluation. Since OpenAlex is openly available, we decided to use OpenAlex in order to follow both guidelines. The fourth principle of the Leiden Manifesto suggests that evaluation methodologies should be transparent. By laying our methodology out in this contribution, we also follow this principle. Adams, McVeigh, Pendlebury, and Szomszor (2019) argue for using profiles rather than metrics in research evaluation. Our proposal of using overlay maps to visualize contributions to reaching SDGs is one step in that direction. An earlier step into that direction was the introduction of beam plots for raw citations (Haunschild, Bornmann, & Adams, 2019). DORA suggests to consider a broad range of impact measures in research evaluation. CoARA also calls for "consideration of contributions to the research ecosystem, knowledge generation and scientific, technological, economic, cultural and societal impact" (CoARA, 2022). By providing a transparent methodology for the analysis of publications that targets SDGs, we extend the range of impact measures for research evaluation with the use of our maps.

Methods and data

Assignments of papers to SDGs

Assignments of papers to SDGs is made in OpenAlex using the Aurora Universities SDG Classifier with a cut-off value of 0.4 as a compromise of achieving high recall and precision (OurResearch, 2025). Details about the classification algorithm were provided by Vanderfeesten, Jaworek, and Keßler (2022).

Data

We used an OpenAlex snapshot from August 2024 available to us via the German 'Kompetenznetzwerk Bibliometrie' (Schmidt et al., 2024). We extracted the SDG-relevant publications for (i) USA, (ii) Japan, and (iii) the world in the time period from 2014 to 2023. No restrictions on document types were imposed. Country information was extracted from the author's affiliations. Documents with multiple affiliations were fully counted as a paper for each of the collaborating co-authors. Thus, it is possible that some documents are counted for both countries included in this analysis. Table 1 shows the 17 SDGs with their number of papers in the time period investigated.

Table 1. SDGs with their number of papers in OpenAlex for the time period from 2014 to 2023 ordered decreasingly by the number of papers.

SDG		#Papers	%Papers
3	Good health and well-being	9,603,428	18.85
4	Quality education	5,244,104	10.29
7	Affordable and clean energy	4,326,404	8.49
2	Zero hunger	4,137,054	8.12
10	Reduced inequalities	3,733,772	7.33
16	Peace, justice, and strong institutions	3,634,677	7.13
8	Decent work and economic growth	2,971,404	5.83
11	Sustainable cities and communities	2,692,363	5.28
5	Gender equality	2,427,416	4.76
6	Clean water and sanitation	2,087,444	4.10
14	Life below water	2,030,626	3.99
9	Industry, innovation, and infrastructure	1,951,836	3.83
15	Life on land	1,909,980	3.75
13	Climate action	1,677,011	3.29
17	Partnership for the goals	1,160,840	2.28
1	No poverty	760,432	1.49
12	Responsible consumption and production	604,498	1.19

Overlay maps

Base maps have been used to create overlay maps. Base maps are intended to spatially position concepts from OpenAlex on a map based on citation relations between the concepts. In OpenAlex, concepts are abstract ideas that scholarly works are about. Concepts are assigned to works based on the title, abstract, and the title of the host venue using an automated classifier. Each work is tagged with multiple concepts although some works are not assigned to any concept. We indicated a concept where a country has reached or surpassed 10% of the world-wide SDG-relevant output in a concept with a red dot on the map. Concepts in which a country did not reach this 10% threshold are shown as gray dots. Thus, red dots indicate concepts with many publications of a country that are relevant for the worldwide research targeting SDGs. Data analysis and graphic production have been done using R (R Core Team, 2021) with the R packages 'tidyverse' (Wickham, 2017) and 'ggforce' (Pedersen, 2024).

We used the global base map for OpenAlex (2008-2022) as provided by Haunschild and Bornmann (2024a, 2024b). The base map provides coordinates for the concepts of level 0,1, and 2 of the science covered by OpenAlex. Concepts are one of the field classifications provided by OpenAlex. The maps also include a cluster assignment that is interpreted as a broad scientific classification: (i) Social Sciences and Humanities, (ii) Medicine, (iii) Physics and Engineering, (iv) Mathematics, Computer Sciences, and Theoretical Physics, (v) Biology, and (vi) Chemistry and Material Sciences.

Results

Figure 1 shows the overlay map for the USA. The six different scientific areas are roughly marked with circles and labels.

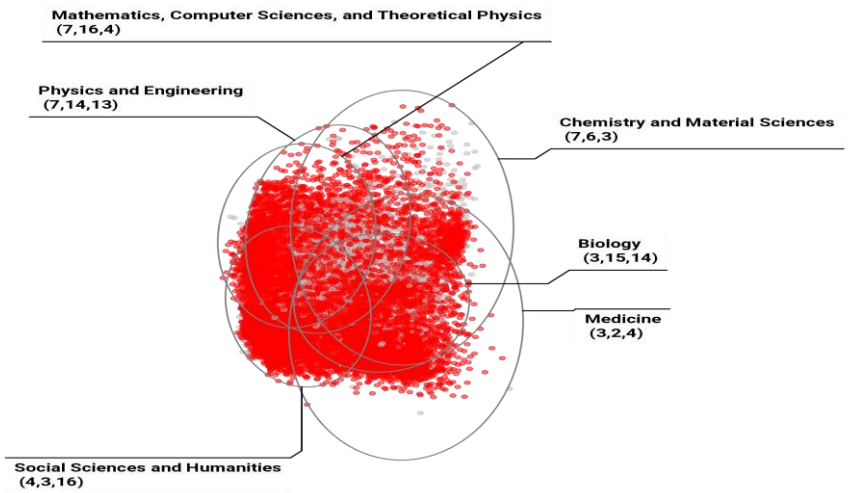

Figure 1. Overlay map of the USA where red dots show concepts with many SDG-relevant publications. The labels of the broad areas are extended by the top 3 SDG numbers in parentheses occurring in these areas (see Table 1).

The overlay map indicates by the many red dots that the USA surpasses the 10% threshold in many concepts (i.e., indicating high SDG-relevance in these fields) within all six broad scientific areas. This is not unexpected due to the very high publication output of the USA in general. The labels of the broad areas were extended by the top 3 SDG numbers in parentheses occurring in these areas (see Table 1). For example, in the case of Physics and Engineering, the top 3 SDGs are 7 ('Affordable and clean energy'), 14 ('Life below water'), and 13 ('Climate Action'). Overall, nine different SDGs occur as top 3 SDGs across all six different broad areas of science.
Figure 2 shows the overlay map for Japan. Due to the lower overall publication output of Japan compared to the USA, fewer red dots are visible. However, several red dots are visible in all broad scientific areas. Overall, eleven different SDGs occur as top 3 SDGs across all six different broad areas of science.
As Figure 2 reveals the map is able to point to Japanese research areas where the country significantly contributed to worldwide SDG relevant research. Two areas of aggregation of red dots indicating high SDG-relevance of Japanese research can be found in the lower-right part (Medicine) and middle-right part (Chemistry and Material Sciences) of the map. In the following, we will have a closer look at these two aggregations of concepts with high SDG-relevance.

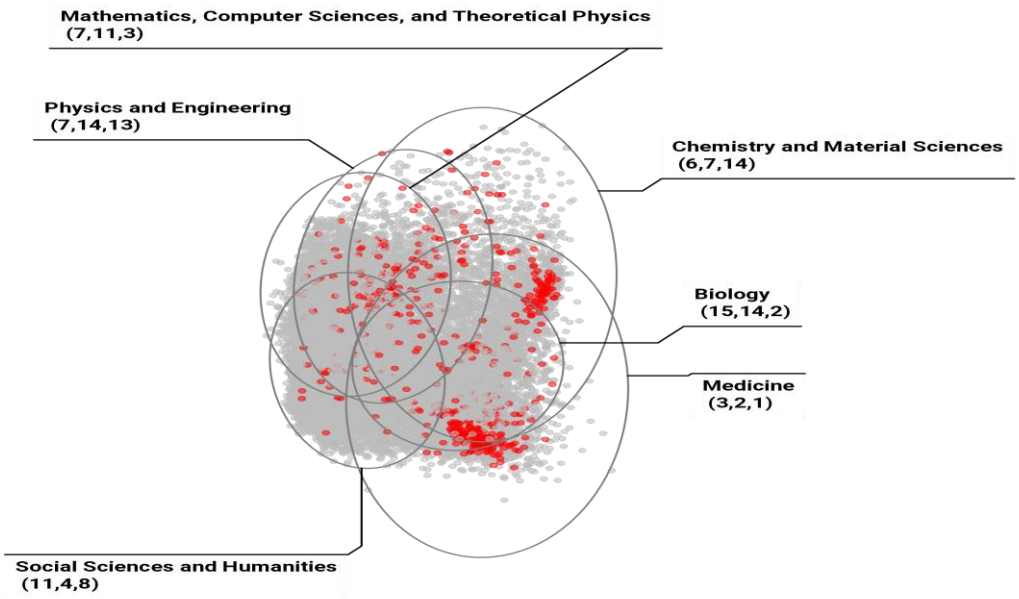

Figure 2. Overlay map of Japan where red dots show concepts with many SDG-relevant publications. The labels of the broad areas are extended by the top 3 SDG numbers in parentheses occurring in these areas (see Table 1).

Concepts of high SDG-relevance in Medicine

Very prominent concepts by the number of SDG-relevant publications within the medical area of concept aggregation with many SDG-relevant publications are the concepts 'Resection', 'Dissection (medical)', and 'Aneurysm'. SDGs 3 ('Good health and well-being'), 2 ('Zero hunger'), and 1 ('No poverty') are the three most relevant SDGs for these concepts for Japanese publications. With the three SDGs, the medical area of concept aggregation reflects the medical research area as a whole. The concept 'Resection' refers to the surgical removal of all or part of an organ, tissue, or biological structure. The concept 'Dissection (medical)' refers to a tear within the wall of a blood vessel. The concept 'Aneurysm' refers to a bulge in the wall of a blood vessel.

Concepts of high SDG-relevance in Chemistry and Material Sciences

Within Chemistry and Material Sciences, concepts such as 'Total Synthesis', 'Diastereomer', and 'Trimethylsilyl' occur with a very high number of SDG-relevant publications in the area of concept aggregation with many SDG-relevant publications. The three concepts exhibit high numbers of publications in SDG 6 ('Clean water and sanitation'). The concept 'Total Synthesis' also contains many publications in SDG 14 ('Life below water'). The concept 'Total Synthesis' refers to a specialized area in organic chemistry that is concerned with synthesizing complex chemical compounds from substances found in nature. The concept 'Diastereomer' describes a specific type of stereoisomer within a compound. This concept also is closely related to organic chemistry. The concept 'Trimethylsilyl' refers to a functional group that is often used as a protective group in certain steps of chemical reactions.

Discussion

Following CoARA that emphases the need for a more holistic approach to research evaluation, we introduce here an approach that is based on overlay maps. The approach reveals national research areas contributing significantly (i.e., more than 10%) to the worldwide SDG-relevant research. We demonstrate our approach using the US and Japanese research as examples. Since the USA is one of the most research active countries in most disciplines, the US map also reveals high research activity with SDG relevance in most disciplines. Our approach is especially interesting for smaller countries with less publications than the USA to reveal their specific contributions to worldwide SDG-relevant research. In this study, we could identify two areas of Japanese research with high relevance for targeting SDGs: Medicine as well as Chemistry and Material Sciences. Since the development of our SDG overlay approach is research in progress, we plan to produce overlay maps also for other countries to reveal their specific SDG-relevant research.

It is one problem of the movements for reforming research assessments that they have not found broad acceptance and application. The results of Rushforth and de Rijcke (2024) show that "there is not yet a deep level of familiarity with international reform movements for responsible metrics and assessment in the United States. The lack of familiarity with the responsible metrics movements' 'responsibility

language' was manifest in: the lack of referencing specific points in responsible metrics statements; lack of awareness of the actors involved in enacting performative powers of metrics (e.g. nobody mentioned publishers); the propensity to present their own 'bottom up' responsibilities which were different from the reform movements' language, or were similar only by coincidence because all actors inhabit the same professional world". The study of Morgan-Thomas, Tsoukas, Dudau, and Gąska (2024) points out that "the limited incidence of non-journal outputs in institutional submissions, the high correspondence between expert score and an aggregate metrics (journal rank), and the non-significance of DORA affiliation, all point to declarations being potentially decoupled from practices". Since our approach is based on freely available OpenAlex data and targets a very relevant question in the area of societal impact measurements, i.e., national contributions to worldwide SDG-relevant research, we assume that there will be a 'market' for its application. We provide an overlay approach that goes beyond using simple counting of publications or citations by displaying national fields with high proportions of SDG-relevant publications on a map and discussing the results for different fields and SDGs.

Our approach is affected by an important limitation that has been addressed, e.g., by Mutz, Bornmann, and Haunschild (2025): the low agreement of different approaches for assigning SDGs to papers. In this study, we used the Aurora Universities SDG Classifier; other classifiers will probably lead to different assignments.

Acknowledgments

Access to OpenAlex bibliometric data has been supported via the German Competence Network for Bibliometrics, funded by the Federal Ministry of Education and Research (grant number: 16WIK2101A

References

Adams, J., McVeigh, M., Pendlebury, D., & Szomszor, M. (2019). *Profiles, not metrics*. Philadelphia, PA, USA: Clarivate Analytics.

Bornmann, L. (2012). Measuring the societal impact of research. *EMBO Reports, 13*(8), 673-676.

Bornmann, L. (2013). What is societal impact of research and how can it be assessed? A literature survey. *Journal of the American Society of Information Science and Technology, 64*(2), 217-233.

Ciarli, T. (Ed.). (2022). *Changing directions: Steering science, technology and innovation towards the sustainable development goals*. SPRU, University of Sussex, UK: STRINGS.

CoARA. (2022). Agreement on Reforming Research Assessment. Retrieved February 7, 2025, from https://coara.eu/app/uploads/2022/09/2022_07_19_rra_agreement_final.pdf

Haunschild, R., & Bornmann, L. (2024a). Global base maps produced with OpenAlex. Retrieved February 20, 2025, from https://doi.org/10.17617/1.daf7-fq06

Haunschild, R., & Bornmann, L. (2024b). The use of OpenAlex to produce meaningful bibliometric global overlay maps of science on the individual, institutional, and national levels. *PLOS ONE, 19*(12), e0308041. doi: 10.1371/journal.pone.0308041.

Haunschild, R., Bornmann, L., & Adams, J. (2019). R package for producing beamplots as a preferred alternative to the h index when assessing single researchers (based on downloads from Web of Science). *Scientometrics, 120*(2), 925-927. doi: 10.1007/s11192-019-03147-3.

Hernandez, C., Suazo López, F., & Domínguez Pacheco, F. A. (2022). The Sustainable Development Goals index: An analysis (2000-2022). *Transdisciplinary Journal of Engineering & Science, 13*(0). doi: 10.22545/2022/00213.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature, 520*(7548), 429-431.

Morgan-Thomas, A., Tsoukas, S., Dudau, A., & Gąska, P. (2024). Beyond declarations: Metrics, rankings and responsible assessment. *Research Policy, 53*(10), 105093. doi: 10.1016/j.respol.2024.105093.

Mutz, R., Bornmann, L., & Haunschild, R. (2025). How to use assignments of United Nations sustainable development goals (SDGs) to scientific papers in research evaluation? The proposal of a gold standard combining assignments from different data providers. *Scientometrics*. doi: 10.1007/s11192-025-05254-w.

OurResearch. (2025). How do you classify works as contributing to the UN SDGs? Retrieved February 24, 2025, from https://help.openalex.org/hc/en-us/articles/27972124390679-How-do-you-classify-works-as-contributing-to-the-UN-SDGs

Pedersen, T. L. (2024). ggforce: Accelerating 'ggplot2'. Retrieved February 24, 2025, from https://CRAN.R-project.org/package=ggforce

Purnell, P. J. (2022). A comparison of different methods of identifying publications related to the United Nations sustainable development goals: Case study of SDG 13: Climate action. *Quantitative Science Studies, 3*(4), 976-1002. doi: 10.1162/qss_a_00215.

R Core Team. (2021). R: A language and environment for statistical computing. Retrieved February 20, 2025, from https://www.R-project.org/

Rushforth, A., & de Rijcke, S. (2024). Practicing responsible research assessment: Qualitative study of faculty hiring, promotion, and tenure assessments in the United States. *Research Evaluation, 33*, rvae007. doi: 10.1093/reseval/rvae007.

Rushforth, A., & Hammarfelt, B. (2023). The rise of responsible metrics as a professional reform movement: A collective action frames account. *Quantitative Science Studies, 4*(4), 879-897.

Schmidt, M., Rimmert, C., Stephen, D., Lenke, C., Donner, P., Gärtner, S., . . . Stahlschmidt, S. (2024). The data infrastructure of the German Kompetenznetzwerk Bibliometrie: An enabling intermediary between raw data and analysis. Retrieved October 28, 2024, from https://doi.org/10.5281/zenodo.13935407

Thelwall, M. (2024). Quantitative methods in research evaluation: Citation indicators, altmetrics, and artificial intelligence. Retrieved July 5, 2024, from https://arxiv.org/abs/2407.00135

Vanderfeesten, M., Jaworek, R., & Keßler, L. (2022). AI for mapping multi-lingual academic papers to the United Nations' Sustainable Development Goals (SDGs). Retrieved February 20, 2025, from https://dx.doi.org/10.5281/zenodo.5603019

Wickham, H. (2017). Tidyverse: Easily install and load the 'Tidyverse'. R package version 1.2.1. Retrieved June 22, 2020, from https://CRAN.R-project.org/package=tidyverse

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., . . . Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. Bristol, UK: Higher Education Funding Council for England (HEFCE).