bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Contrasting the open access dissemination of COVID-19 and SDG research

Vincent Larivière[1,2,3,4*], Isabel Basson[1], Jocalyn P. Clark,[5,6,7]

[1] École de bibliothéconomie et des sciences de l'information, Université de Montréal, Montréal, QC, Canada.
[2] School of Public Policy, Georgia Institute of Technology, Atlanta, Georgia, United States.
[3] Department of Science and Innovation-National Research Foundation Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University, Stellenbosch, Western Cape, South Africa.
[4] Observatoire des Sciences et des Technologies, Centre interuniversitaire de recherche sur la science et la technologie, Université du Québec à Montréal, Montréal, QC, Canada.
[5] International Editor, The BMJ, London, UK
[6] Adjunct Professor, Department of Medicine, University of Toronto, Toronto, Canada
[7] Honorary Associate Professor, Institute for Global Health, UCL, London, UK

*Correspondence to: vincent.lariviere@umontreal.ca

Abstract: This paper examines the extent to which research has been published open access in response to two global threats: COVID-19 and the Sustainable Development Goals (SDGs), including climate change. We compare the accessibility of COVID-19 content versus SDG literature using the Dimensions database between 2000 and 2021, classifying each publication as gold open access, green, bronze, hybrid, or closed. We found that 79.9% of COVID-19 research papers published between January 2020 and December 2021 was open access, with 39.0% published with gold open access licenses. In contrast, just 55.7% of SDG papers were open access in the same time period, with only 36.0% published with gold open access licenses. Papers related to the climate emergency overall had the second-lowest level of open access at just 55.5%. Papers published by the largest for-profit publishers that committed to both the SDG Publishers Compact and climate actions were not predominantly published open access. The paper highlights the need for continued efforts to promote open access publishing to facilitate scientific research and technological development to address global challenges.

One-Sentence Summary: In contrast to COVID-19 papers, research on UN Sustainable Development Goals including the climate emergency have not been made open access by leading global science publishers despite their corporate commitments to sustainability and climate action.

Collective efforts against global threats

Scientific research and technological development are key to solving global challenges. Such importance was clearly highlighted by the rapid production of knowledge for the COVID-19 pandemic response, which required both concerted basic scientific efforts across all fields and the development of new health technologies including vaccines, along with sizeable governmental investments. Within a month of the first reported coronavirus case, more than 160 funders, research institutions, and publishers committed to make such knowledge widely available. The pledge, initiated by the Wellcome Trust in January 2020, committed signatories to share rapidly and openly the research papers and data relevant to the outbreak, in order to inform the global health response and save lives. [1] Open access publishing increases visibility, discoverability, and impact by providing immediate and unlimited access to scientific research.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Such a collective commitment to COVID-19 data sharing and open access is remarkable, and begs the question of how such researchers and publishers have responded to other global threats. For example, in 2015, the 154 member states of the United Nations (UN) selected 17 Sustainable Development Goals (SDGs) as part of the 2030 agenda to guide and stimulate global action in areas of critical importance for humanity and the planet. Among these is a commitment to accelerated action on the climate emergency, considered by many as the greatest threat the world has ever faced [3] and in its final crucial decade to change course on devastating health and environmental risks. [4] In recognition, science publishers, including the Big 5 publishers that together represent over 50% market share of a global $28 billion annual publishing industry, [5] agreed to the SDG Publishers Compact. This commits the 270 signatories—which include publishers Elsevier, Sage, Springer Nature, Taylor & Francis, and Wiley—to develop sustainable practices and actively promote content on sustainability, justice, and safeguarding and strengthening the environment. [6]

The world's largest science publisher, Elsevier, furthermore has specific commitments to climate change explicitly acknowledging their role in supporting the "scholarly dialogue on climate change and its impacts though unique content, data and products" and ensuring that "research is communicated effectively, grounded in real world examples, and acted upon by the right people, [as] one of the greatest challenges we're facing."[7] Similarly, the other two largest publishers, Springer-Nature [8] and Wiley [9], have recognised their critical role in promoting climate action and research.

In this context of commitments, we examine the extent to which research has been published open access, comparing accessibility of Coronavirus/COVID-19 content vs SDG, including climate change, literature. Using the Dimensions database of scientific research (www.dimensions.ai) from January 2000 to December 2021 we identified and compared Coronavirus and COVID research publications (including a subset of COVID-19 papers) with research publications classified as SDG related (including a subset of climate related papers), classifying each as to its open access status, defined here in declining levels of openness: gold (immediate and permanent access, author retains copyright and permissions), green (authors self-archive, publisher retains rights), bronze (article is made available but cannot be redistributed or reused), hybrid (published open access in a subscription journal), or not (see supplemental material for full Materials and Methods).

Not so open science

404,541 COVID-19 research papers were published between Jan 2020 and December 2021, of which 79.9% were open access. Overall, papers on any coronavirus amounted to 417,822 between 2000 and 2021, of which 79.8% are open access including 38.4% gold open access (*Figure 1*). Just 39.0% of the COVID-19 papers were published with gold open access licenses allowing full copyright and re-use allowances, as was required by the Wellcome data sharing pledge [1]. 26.2% of COVID-19 research publications were made available via so-called bronze OA, where a paywall is removed by the publisher but allows them to retain copyright and commercial rights and no obligation to maintain access for any period of time for non-paying readers. Nevertheless, 3 years after COVID-19 was declared a public health emergency of international concern by the World Health Organization, publishers have continued to make COVID-19 content freely available via various open access routes. This policy is at least partially responsible for the unprecedented volume science publishers have received of article

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

submissions, publications, and market growth, which is reported to be 8.1% for 2020, three times larger than the average annual absolute growth from 2013 to 2019. [5,10]

In contrast, 803,152 distinct papers on an SDG were published between Jan 2020 and the end of 2021, and just 55.7% of these were open access as of June 2022. 36% of the SDG papers were gold open access. Bronze, green, and hybrid open access licenses each accounted for less than 10% of papers. Levels of open access of relevant research varied by SDG, but none attained levels of COVID-19 (*Figure 2*). Most SDGs have openness rates between 55% and 67%, except for affordable and clean energy (SDG 7) for which a minority of papers are freely available. Among these, papers related to climate change ranks second lowest in in terms of open access, at just 55.5% overall. A third of research papers (33.8%) on the climate emergency are published with a gold open access license allowing reuse and author retention of copyright.

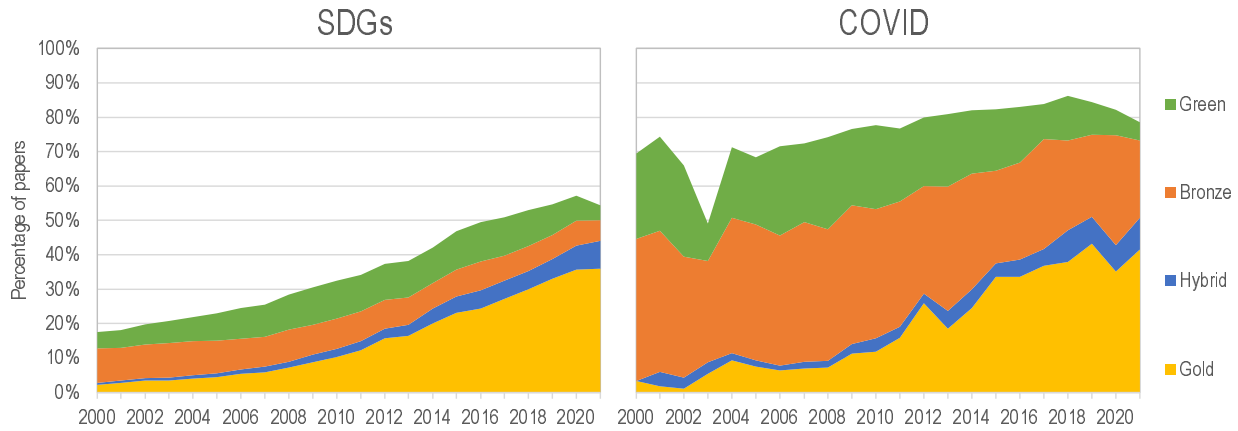

Figure 1 Percentage of open access for SDG-related and COVID-related research papers, by access type, 2000-2021.

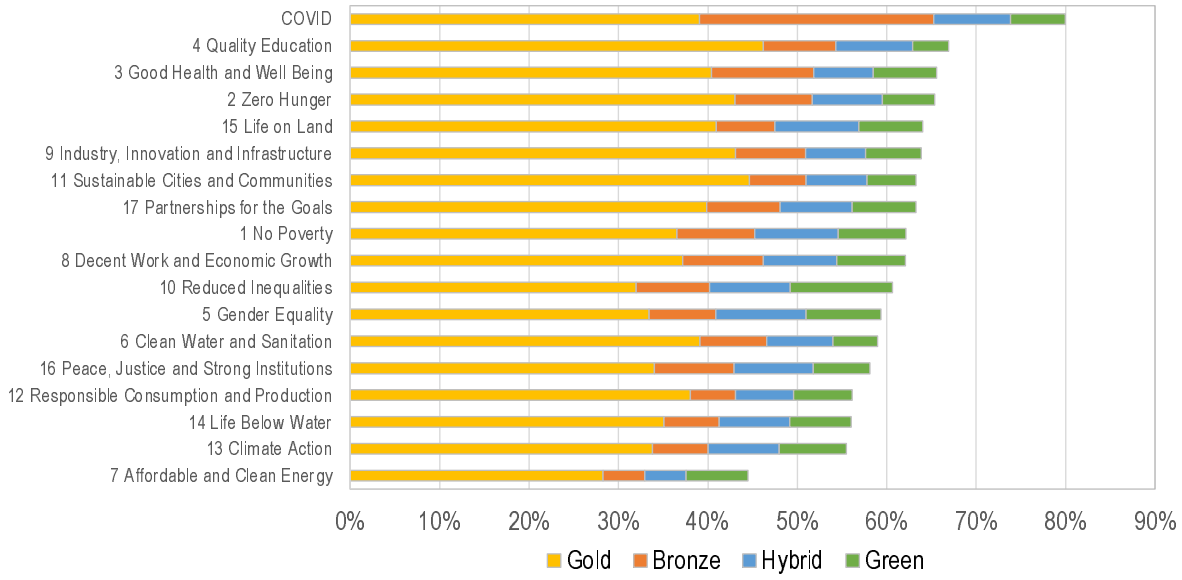

Figure 2 Number of papers by SDG and for COVID-related research, by access type, 2017-2021. Dimensions.ai database.

For work published in the largest publishers that committed to both the SDG Publishers Compact and climate actions, related papers are not predominantly published open access. For example, among all SDG research papers published by Elsevier, no SDG has more than 50% of its

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

research papers disseminated via open access publishing between 2017 and 2021. These levels range from 22.2% for papers on affordable and clean energy (SDG7) to 49.9% of those related to health and well-being (SDG3). Percentages are slightly higher for papers published by Springer-Nature or Wiley, with respectively ten and three SDGs having more than 50% of the related papers open access. For the climate emergency (SDG13), none of the three top publishers made more than 50% of the papers on this topic that they published open access. In contrast, the top 3 largest publishers have made almost all COVID-19 papers by June 2022 openly available to the research community and the general public: 92.1% of those published by Elsevier, 95.0% by Springer Nature, and 83.4% by Wiley.

Publishing paradoxes

What does this say about science publishing in the midst of global threats? COVID-19 seemed to spur collective action on the part of the research and publishing communities, which had no precedent but nonetheless demonstrated the ability to rapidly and openly disseminate crucial knowledge. Publishing executives and editors may have felt a moral obligation to contribute to collaborative efforts when global solidarity was required to overcome a novel infectious disease threat. An assessment of the Wellcome pledge similarly found that about half of COVID-19 papers published among its signatories were gold open access. [11] In contrast, despite the enormity and importance of the challenge at stake with the climate emergency – the very survival of the planet and the human race – the vast majority of papers published on the topic are not open access. This inconsistency was particularly acute for the top 3 largest publishers, which are home to the world's most prestigious and highly visible journals. The contrast between openness of research related to the COVID-19 pandemic but not to the priority global development goals or the climate emergency raises a critical tension in the scholarly communication field. It is a worrying contradiction in how and what knowledge is published and disseminated that health advocates, funders, and institutional leaders must demand more attention to.

A main element to consider is the role of bronze OA, which is a way to make content freely available without giving up commercial rights. It has the appearance of being and fulfilling "open access" commitments, whilst allowing publishers to retain copyright and block redistribution or reuse. We observe that for COVID-19 research, the bronze OA route was utilised by the largest 3 publishers. But for the other publishers, gold open access was predominantly provided. Some of this may be due to the choice of authors rather than publishers, but is unlikely to be a sizeable share of papers given the promise of many publishers to ensure COVID-19 content is made available. Among the top 3 largest publishers, therefore, access to COVID papers assigned bronze OA can be withdrawn at any time. Retention of copyright and reuse may also serve a commercial strategy to preserve dominant market position and historical huge profit margins. [5, 12] The extent to which publishers can market and sell gold open access publishing, which is often facilitated via article publication charges paid by the author, institution or funder, is evident in the SDG literature. In the Dimensions database, about 18.2% overall of research is published this way, compared with 12.3% for bronze OA [13]. A large proportion of COVID-19 publications, produced at volumes never seen before in the scientific literature, are thus owned by publishers rather than by their authors. The research asset created by the extraordinary collective response of scientists and funders to COVID-19 worldwide is not a wholly public one.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

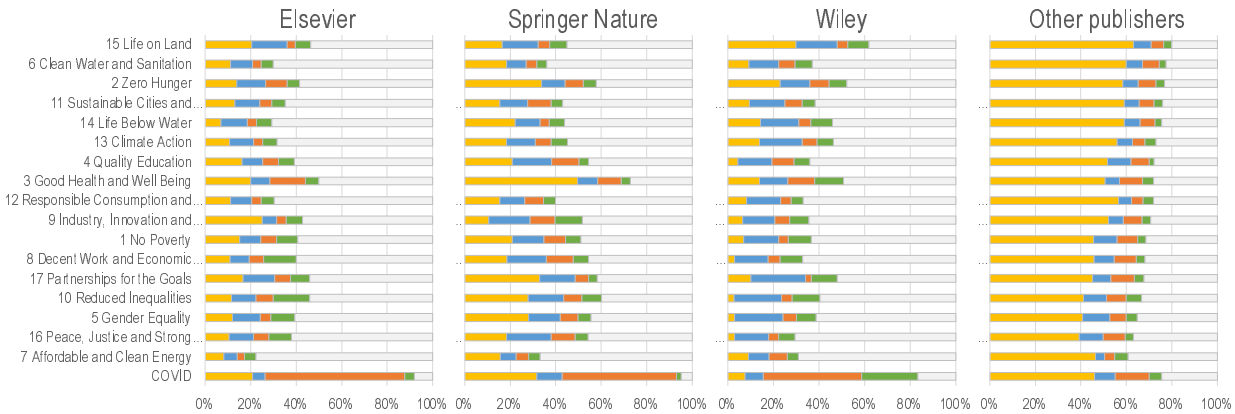

Figure 3 Percentage of open access (gold, hybrid, bronze and green) papers, for sustainable development goal research and COVID-related research, by publisher, 2015-2021.

The sustainability challenge

The importance of scientific knowledge and the global impacts of health threats are at their highest levels of public consciousness. Therefore, the general public and readers alike will increasingly expect scientific institutions to uphold higher-than-average moral standards and to be driven by common good principles rather than personal advancement. The SDG agenda is multi-faceted but no less important than COVID-19 as a global health challenge, and like climate change is not a mere scientific concern, but a social one; requiring concerted, collective global efforts to tackle and implement the blueprint for current and future peace and prosperity for people and the planet that was envisioned when member states agreed to the sustainable development priorities. Equally, publishers who agreed to sustainable development and climate action must be held to account for their commitments and plans. Facilitating SDG literature, including research relevant to the climate emergency, to be open to all would inform researchers, policymakers, and the public. Knowledge is critical to understand what is at stake, advance scientific outcomes, and help progress towards sustainability and climate goals, most of which have been severely derailed by the COVID-19 pandemic. [14]

References

1. Wellcome Trust. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak. (2020). Available at : https://wellcome.ac.uk/press-release/sharing-research-data-and-findings-relevant-novel-coronavirus-COVID-19-outbreak

2. United Nations General Assembly (2015). Resolution adopted by the General Assembly on 25 September 2015. Available at : https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf

3. I Fry. Report of the Special Rapporteur on the promotion and protection of human rights in the context of climate change. (October 2022.) Available at: https://daccess-ods.un.org/tmp/167568.810284138.html

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4. IPCC. Climate Change 2023: Synthesis Report. Available at: https://www.ipcc.ch/report/ar6/syr/

5. Association of Scientific, Technical and Medical Publishers. STM Global Brief 2021 – Economics and Market Size, Oct 19, 2021. https://www.stm-assoc.org/wp-content/uploads/2022_08_24_STM_White_Report_a4_v15.pdf

6. SDG Publishers Compact. (2021). https://www.un.org/sustainabledevelopment/sdg-publishers-compact/

7. Elsevier. Climate Action. Available at: https://www.elsevier.com/about/corporate-responsibility/climate-action (visited April 5, 2023)

8. Springer Nature. https://group.springernature.com/gp/group/taking-responsibility

9. https://www.wiley.com/en-us/network/publishing/research-publishing/trending-stories/becoming-greener-where-are-we-now

10. C Petrou. Scientific Output in the Year of COVID, An Update. The Scholarly Kitchen. https://scholarlykitchen.sspnet.org/2021/02/23/guest-post-scientific-output-in-the-year-of-COVID-an-update/

11. M Hagve. The money behind academic publishing. *Tidsskr Nor Legeforen* 2020. doi: 10.4045/tidsskr.20.0118

12. A Chiarelli, L Loffreda, E Cox, R Johnson, C Ferguson, E Gagné, D Campbell, A Emecz. From intent to impact: Investigating the effects of open sharing commitments. (2022). 10.5281/zenodo.6620914.

13. Basson I, Simard M-A, Ouangré ZA, Sugimoto CR, Larivière V (2022) The effect of data sources on the measurement of open access: A comparison of Dimensions and the Web of Science. PLOS ONE 17(3): e0265545. https://doi.org/10.1371/journal.pone.0265545

14. United Nations The Sustainable Development Goals Report 2020. (2020). Available at: https://unstats.un.org/sdgs/report/2020/The-Sustainable-Development-Goals-Report-2020.pdf

Funding: Funding for this analysis was provided by the Canada Research Chairs program, grant number 950-231768, and by Open Society Foundations, Grant number OR2021-83000.

Author contributions:

Conceptualization: VL, JC

Data Curation: VL

Funding acquisition: VL

Investigation: VL, IB

Methodology: VL

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Project administration: VL

Resources: VL

Supervision: VL

Visualization: VL

Writing - Original Draft: VL, IB, JC

Writing - Review & Editing: VL, IB, JC

Competing interests:

Supplementary Materials

Figs. S1 to S2

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Materials

Using the Dimensions database from January 2000 to June 2022, we created two datasets: one of COVID-related research (including research on coronaviruses in general) and one related to SDGs. The COVID dataset was created using the COVID*, SARS-COV-2 and Coronavirus* search strings, which has been the standard query to retrieve papers on the topic.[1] This query retrieved 417,822 research papers between 2000 and 2021; most of which have been published since 2020 (96.8%). The definition of research related to SDGs remains more challenging, as some of the SDGs remain poorly defined.[2] However, recent advances in natural language processing and machine learning led to the creation of a validated dataset of SDG related research, which we have leveraged for this analysis.[3] The SDG-related dataset consists of 4,636,992 paper/SDG combinations (as papers can be associated with more than one SDG). The smallest SDG from a number of paper point of view is SDG 17 (Partnerships for the Goals) with 9,150 papers, followed SDG 9 (Industry, Innovation and Infrastructure), which contains 14,509. At the other end of the spectrum, SDG 3 (Good Health and Well Being) has the highest numbers of papers (1,178,622), followed by SDG 7 (Affordable and Clean Energy), with 1,047,233 papers (Fig. S1). Open access of papers was determined by Unpaywall, as embedded in the Dimensions.ai database. Following Piwowar et al.,[4] open access status was divided into four mutually exclusive category: gold open access (paper published in a fully open access journal—with or without APCs), hybrid open access (paper published with an open license in a subscription journal), bronze open access (paper freely available on the journal website, but without a clear license), and green open access (papers solely available in a repository). It is worth recalling that the open access status is determined as of now (or a couple months ago) and does not represent the availability at the time of publication.

[1] Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., ... & Kohlmeier, S. (2020). Cord-19: The COVID-19 open research dataset. ArXiv. Dataset available at: https://www.kaggle.com/datasets/allen-institute-for-ai/CORD-19-research-challenge

[2] 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.

[3] Wastl, Jürgen, & Diwersy, Mario. (2019, December 19). Phase 1 and Phase 2 Summary of SDG Project by Springer Nature, VSNU/UKB, Digital Science. Zenodo. https://doi.org/10.5281/zenodo.3904447

[4] Piwowar, H., Priem, J., Larivière, V., Alperin, J.P., Matthias, L., Norlander, B., Farley, A., West, J., Haustein, S. (2018) The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. PeerJ 6:e4375 https://doi.org/10.7717/peerj.4375

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

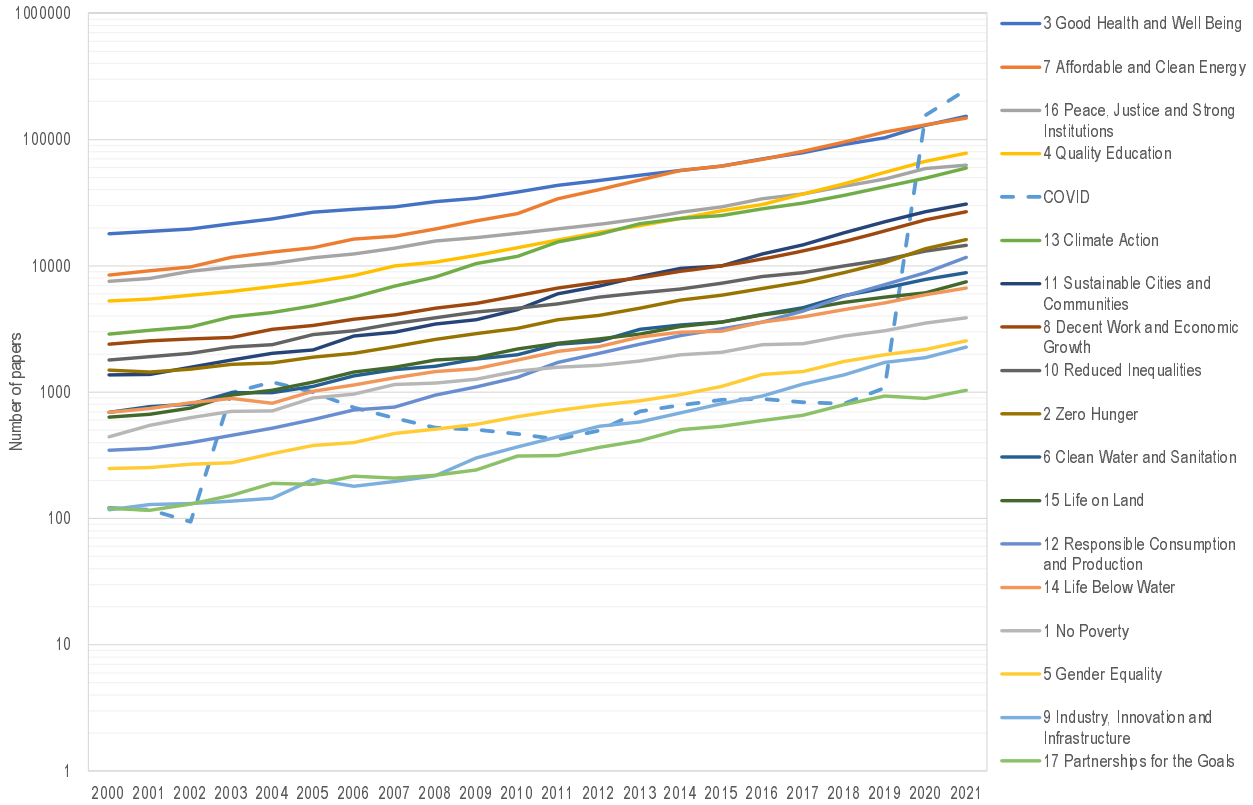

Figure S1 Number of papers by SDG and for COVID-related research, 2000-2021. Dimensions.ai database.

Country	SDG1	SDG2	SDG3	SDG4	SDG5	SDG6	SDG7	SDG8	SDG9	SDG10	SDG11	SDG12	SDG13	SDG14	SDG15	SDG16	SDG17	COVID
United States	47%	60%	61%	37%	35%	58%	48%	39%	43%	42%	45%	45%	57%	59%	55%	29%	48%	82%
China	44%	45%	52%	46%	27%	36%	29%	40%	46%	39%	47%	37%	35%	31%	50%	35%	41%	82%
United Kingdom	77%	81%	80%	73%	72%	83%	79%	73%	80%	79%	78%	79%	82%	78%	82%	65%	74%	87%
India	40%	53%	67%	54%	40%	37%	21%	32%	33%	40%	38%	24%	27%	21%	39%	36%	31%	82%
Italy	48%	64%	60%	54%	54%	58%	50%	49%	55%	53%	55%	57%	53%	50%	55%	43%	67%	85%
Indonesia	88%	90%	89%	94%	89%	92%	90%	90%	89%	91%	94%	91%	92%	90%	92%	90%	89%	87%
Brazil	74%	59%	70%	88%	80%	60%	36%	65%	60%	78%	61%	44%	41%	34%	52%	84%	67%	89%
Spain	60%	64%	58%	71%	65%	58%	60%	58%	62%	66%	65%	58%	59%	48%	57%	68%	67%	90%
Germany	51%	63%	62%	57%	55%	59%	55%	59%	67%	59%	61%	59%	65%	68%	62%	42%	52%	84%
Canada	43%	57%	60%	42%	38%	41%	31%	41%	28%	46%	40%	38%	45%	47%	53%	34%	60%	84%
France	49%	59%	63%	54%	55%	59%	64%	49%	61%	63%	63%	58%	69%	71%	70%	38%	47%	87%
Australia	35%	55%	61%	40%	37%	45%	41%	40%	45%	46%	45%	39%	49%	51%	56%	33%	50%	82%
Iran	45%	43%	55%	60%	50%	35%	20%	52%	28%	49%	38%	24%	23%	26%	42%	57%	75%	82%
Japan	44%	55%	58%	30%	35%	45%	37%	39%	38%	49%	53%	44%	46%	48%	47%	33%	49%	83%
Turkey	35%	42%	54%	47%	40%	27%	24%	40%	28%	46%	37%	34%	25%	22%	30%	33%	60%	82%
Russia	63%	74%	69%	72%	82%	72%	59%	72%	80%	63%	87%	74%	58%	50%	74%	74%	72%	71%
South Korea	55%	64%	69%	51%	42%	42%	29%	41%	50%	47%	52%	36%	38%	37%	66%	36%	41%	91%
Saudi Arabia	29%	70%	77%	67%	67%	52%	48%	51%	41%	63%	60%	46%	52%	68%	50%	50%	67%	89%
Poland	75%	74%	78%	73%	58%	77%	68%	74%	89%	71%	82%	78%	70%	69%	72%	72%	64%	87%
Netherlands	77%	80%	78%	79%	74%	80%	80%	81%	87%	81%	81%	72%	85%	81%	81%	75%	80%	90%
World	56%	63%	65%	62%	55%	54%	41%	56%	60%	57%	59%	50%	52%	51%	60%	52%	60%	84%

Figure S2 Percentage of Open Access papers, by SDG and country, 2017-2021. Dimensions.ai database.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.18.541286; this version posted May 21, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.