Pathos

Open Science Impact Pathways

Deliverable 1.2

Scoping Review of Open Science Impact

Deliverable Number and Name	D1.2 Scoping Review of Open Science Impact
Due Date of Deliverable	30.04.2023
Delivery Date	02.02.2024
Work Package	WP1
Type	R - Report
Authors	Thomas Klebel (KNOW), Nicki Lisa Cole (KNOW), Lena Tsipouri (OPIX), Eva Kormann (KNOW), Istvan Karasz (TGB), Sofia Liarti (OPIX), Lennart Stoy (TGB), Vincent Traag (ULEI), Silvia Vignetti (CSIL), Tony Ross-Hellauer (KNOW)
Reviewers	Ron Dekker (TGB), Silvia Vignetti (CSIL), Jessica Catalano (CSIL)
Approved by	Ioanna Grypari (ARC)
Dissemination Level	PU
Version	2.0
Number of Pages	88

The information in this document reflects only the author's views and the European Commission is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.

This project has received funding from the European Union's Horizon Europe framework programme under grant agreement No. 101058728. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.

Revision history

VERSION	DATE	REASON	REVISED BY
0.0	01.02.23	First Draft	ALL
0.1	14.04.23	Intermediate version	TK, NLC, LT, EK, TRH
0.2	21.04.23	Peer review	Ron Dekker, Silvia Vignetti, Jessica Catalano
0.3	27.04.23	Peer review comments addressed	TK, NLC, LT, EK, TRH
1.0	28.04.23	Final version after proofreading	TRH
1.1	21.01.24	Revisions responding to EC reviewer feedback	TRH, TK, NLC
1.2	31.01.24	Peer review	Silvia Vignetti, VT, LS
2.0	01.02.24	Revisions and final formatting	TK, TRH

Contributors

The following PathOS team members contributed via record screening and/or data-charting but did not contribute as authors. We gratefully acknowledge their efforts. Alkis Pitelis (OPIX), Ioanna Grypari (ARC), and Simon Apartis (CNRS)

Table of Contents

List of Tables

List of Figures

This work is licensed under a Creative Commons Attribution 4.0 International License.

The authors of this document have taken any available measure in order to ensure that its content is accurate, consistent and lawful. However, neither the project consortium as a whole nor the individual partners that implicitly or explicitly participated in the creation and publication of this document hold any sort of responsibility that might occur as a result of using its content. In case you believe that this document infringes in any way upon IPR held by you as a person or as a representative of an entity, please do notify us immediately.

This publication has been produced with the assistance of the European Union. The content of this publication is the sole responsibility of the PathOS consortium and can in no way be taken as a reflection of the views of the European Union.

PathOS is a project funded by the European Union (Grant Agreement No 101058728).

Abbreviations

APC	Article Processing Charge
CS	Citizen Science
D	Deliverable
EC	European Commission
EU	European Union
FAIR	Findable, Accessible, Interoperable, Reusable
OA	Open Access
OACA	Open Access citation advantage
OGD	Open Government Data
OS	Open Science
PathOS	Open Science Impact Pathways (Horizon Europe Project)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR	PRISMA Extension for Scoping Reviews
R	Report
R&D	Research & Development
R&I	Research & Innovation
SDG	Sustainable Development Goal

Executive Summary

This report details work to systematically scope evidence of the academic, societal and economic impacts of Open Science. It is guided by the PRISMA Extension for Scoping Reviews (PRISMA-ScR) methodological framework,and was limited to works in English since 2000 found in academic databases (Web of Science, Scopus) of peer-reviewed literature. This deliverable reports findings from the first stage of this work. Future work will extend this via snowball citation searching and web search for grey literature and will be published as three pre-prints.

Through systematic screening and assessment of over 30,000 initial records, we identified 479 relevant studies (311, 155 and 13 related to academic, societal and economic impact, respectively). Our findings show that evidence of impact is concentrated around Open Access (primarily academic impact) and Citizen Science (primarily societal impact), with little evidence of impact for other Open Science aspects, and very limited evidence of economic impact. Across types of impact, we found:

Academic impact: Open Access, especially impact as measured via citations, is most heavily studied. Evidence suggests an Open Access citation advantage; exclusion of authors from less-resourced regions and institutions due to APCs; and that “predatory publishing” threatens the quality of the research literature. Open/FAIR Data are associated with data reuse and a citation advantage for associated papers, but their role in fostering (computational) reproducibility seems less significant than expected. Open Code and Software produce efficiency gains in software development and may also increase citations of associated papers. Citizen Science increases efficiency and scope of data collection, but data quality is sometimes of issue. Open peer review shows neutral to positive effects on review quality.

Societal impact: The majority of studies relevant to societal impact concern Citizen Science, across a variety of types including educational, engagement and empowerment benefits for participants and their communities, the creation of data for use in governmental monitoring and administering of environments and natural resources, and impact in policy development, among others. Beyond CS, evidence is more limited. Some literature demonstrates societal impacts of OA, including public engagement with scientific literature, use in policy-making, and health-related outcomes. Beyond this, our search revealed little evidence. Especially relevant is the limited evidence (at this stage in our study) regarding the policy impact of OS (a recurrent claim in OS advocacy) and the societal impact of Open/FAIR Data.

Economic impact: Evidence here was scarce, with only 13 papers identified as relevant. Most papers referring to economic evidence use a theoretical rationale explaining why the academic and societal impacts eventually can turn into economic benefits but do not provide quantitative corroboration of this rationale. Evidence was most prevalent from the biomedical and health

domains. Some evidence gives positive indications of the potential of OA and Open/FAIR data to power economic activity, but this is still largely without rigorous quantification.

The report closes by reflecting on evidence gaps, including potential causes and solutions.

1. Introduction

PathOS (https://pathos-project.eu/), a 3-year 2m EUR project funded by the European Commission's Horizon Europe programme, aims to identify and quantify the Key Impact Pathways of Open Science (OS) relating to the research system and its interrelations with economic and societal actors. PathOS will enable a new understanding of OS impacts and their causal mechanisms. Its work plan encompasses actions to synthesise and structure current evidence, development of new methods and tools for measuring impact, iterative pilot-testing via in-depth case studies, innovative dissemination and networking, and co-creation synthesis activities culminating in policy recommendations. This is pivotal in order to develop effective OS policy in the EU.

It does so by collecting concrete evidence of the causal effects of OS by studying the pathways of OS practices, from input to output, outcome and impact, including the consideration of enabling factors and key barriers. Impacts and pathways will be developed in particular in the three areas of science, society and economy. By investigating, measuring and comparing its costs and benefits together with its pathways, PathOS will (i) bring a better understanding of the implications of OS for science, economy and society, (ii) provide recommendations to policy makers and other actors in the R&I ecosystem as to how and to what extent OS should be promoted in a balanced way, and (iii) develop innovative tools and methods using big data to augment traditional ones for studying the causal effects of OS. This will enable evidence-based OS policy prioritisation, maximum OS impact, and increased R&I capacity in EU research systems.

A key first step in this endeavour is to systematically scope, critically appraise, consolidate and valorise current knowledge from existing literature relating to quantifying and modelling OS impact, in order to provide a comprehensive, critical view of current evidence to underpin future project activities. This is especially necessary since, although a lot of work has been done by institutions, funders and even publishers to measure progress in the uptake or implementation of OS, much less exists to systematically assess its impact. In the absence of broader assessment for OS impact, the academic literature is a key source of information. But while a few studies have focused on the impact of individual aspects of OS or impact types, , to date (to our current knowledge) no comprehensive and systematic appraisal of the evidence on academic, societal and economic impacts of all aspects of OS has yet been done. In 2016, Tennant et al. (2016) conducted a narrative (i.e., non-systematic) review of academic, societal and economic impacts of Open Access (OA) which found there was "clearly much scope for additional research". In addition to it being somewhat outdated at this point, Tennant et al. (2016) only synthesised knowledge on one aspect of OS, namely Open Access. In 2019, Fell conducted a semi-systematic review of the economic impacts of OS (which closely followed the

PRISMA schema) (Fell, 2019). Regarding the societal impact of OS, there are reviews that demonstrate the societal impact of Citizen Science broadly (von Goenner et al., 2023) or in focused ways (Aristedou and Hreodotou, 2020; Walker et al., 2021; Bonney et al., 2016). What is hence missing is a broader synthesis of evidence of all types of impact (academic, societal, economic) across all aspects of OS (Open Access, Fair and Open Data, etc).

Impact is defined by the OECD-UNDP (2000) as: "Results of a programme or project that are assessed with reference to the development objectives or long-term goals of that programme or project; changes in a situation, whether planned or unplanned, positive or negative, that a programme or project helps to bring about."

As defined in PathOS Deliverable 1.1 "Open Science Intervention Logic" (Dekker, Karasz & Stoy, 2023), impacts may be direct or indirect, with primary and secondary effects, such as improved living standards, enhanced food security, higher export earnings, and savings from reduced imports. In the realm of Open Science, this includes, for example, fostering trust in the robustness of research results, boosting innovation for enterprises, and tackling Sustainable Development Goals challenges.

In the framework of this project, we use the PRISMA scoping review methodology to gather evidence of:

- **Academic impact:** Under academic impact, we understand the demonstrable contributions that OS has upon the workings of the academic research system itself (excluding its societal and economic/industry interactions). This includes changes, internal to the academic system, to the efficiency, productivity, quality, reproducibility and reuse of research and research processes, as well as changes in levels of education, collaboration and equity within the scientific system. Since performance in research is so often measured by citations, and in view of the fact that pilot-testing and prior knowledge reveals a vast literature regarding the effect of OS practices upon citations, we include this as a distinct theme.
- **Societal impact:** Includes impacts of research upon wider society, including the policy sphere, greater engagement, participation, education and trust among societal actors. It also includes contributions to societal impacts of mission-oriented research (research aiming explicitly to address pressing societal challenges), represented by the Sustainable Development Goals (SDGs; e.g., health, environment, gender equity). Given the immense societal impact of the COVID-19 pandemic in recent years, and the key role played by OS in the rapid response of society to this emergency, we also include this theme here.
- **Economic impact:** Refers to the impacts of OS on industry and the broader economy including changes in productivity, competitiveness, employment, income, and value. Broadly speaking, economic impacts may refer to efficiency and enablement gains. Efficiency gains (getting the same outputs with less inputs or more output with the same

input) include savings in access costs (e.g. reduced costs for firms to obtain or access research results as well as redundant costs avoided), in labour costs (or productivity improvements, access to research results becomes less time-consuming) and in transaction costs (savings in costs and time required to reach agreements necessary to access data or publications). Enablement gains may arise in the form of new outputs (e.g. product, service, collaboration) and/or increased productive processes that could not otherwise have been undertaken in a closed environment.

Mapping Open Science impact pathways involves outlining how inputs (resources) result in activities (more Open Access publishing or Citizen Science projects for example), and then how these connect to short-term outputs, medium-term outcomes and long-term impacts. These chains imply complex logics of causation. Identifying and estimating causal effects of OS is a more challenging activity than the description of pathways.

The key challenge is related to the attribution problem, i.e. the possibility to clearly attribute the observed change (e.g. a higher number of citations) to OS and not to other possible confounding factors. Providing evidence for the attribution (i.e. the causal claim) would require having a suitable counterfactual situation for comparison. In principle, studies on the impact of Open Science should compare two situations and investigate how impacts change: one where OS takes place, and one where OS does not take place. For example, such studies would ideally compare trust in science between a world where all scientific outputs are openly available (universal Open Access), and one where they are only available through paywalls or library subscriptions (universal closed access). Unfortunately, studies observing both situations (e.g. in carefully controlled experiments) are rare. Methods to infer causal effects from non-experimental data exist but are not yet applied widely. In many cases, we therefore must acknowledge that the literature does not provide sufficient evidence to identify causal effects, and restrain ourselves from drawing too strong conclusions for policy or advice. Beyond attribution, however, it is important to recognise that OS can contribute, together with other factors, to a number of scientific, societal and economic impacts. Understanding how these factors are intertwined with OS is also part of constructing impact pathways based on the evidence available.

We assess these impacts as they relate to various aspects of OS, namely:

- **Open Access:** We follow Suber (2012) and define Open Access (OA) literature as being research literature (articles, books, conference proceedings) that is “digital, online, free of charge, and free of most copyright and licensing restrictions.” OA can be achieved either via OA publishing (“gold” or “diamond” OA), where these criteria are fulfilled on publication, or author self-archiving of alternative versions in repositories (“green OA”).

- **Open and/or FAIR Data:** We define Open Data per the Open Definition[1]: Data is open if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness. We also include FAIR data, research data that is Findable, Accessible, Interoperable, and Re-Usable is also within scope (Wilkinson et al., 2016).
- **Open Methods:** This relates to enabling free-access to methods, protocols, materials, and other experimental elements, especially to enable the reuse and reproduction/replication of research.
- **Open Code/Software/Tools:** This refers to openly available *research* code, software, or tools. This entails code, software and tools which are specifically built and maintained for research purposes. Examples include software written to accompany specific analyses, statistical libraries/packages, or dedicated research software. General purpose open-source software is out of scope.
- **Citizen Science:** This is the practice of opening the research process itself to the broader public ("citizens"). Practices range from crowdsourcing data collection to "extreme citizen science", with public involvement into processes of problem definition, data analysis and interpretation, as well as dissemination (English et al., 2018). Citizen Science (CS) is increasingly part of common definitions of OS (e.g., the EC's approach to OS[2]). It is an important step in making research open to wider audiences, by fostering engagement beyond consumption, and is thus included within our scope.
- **Open Evaluation:** This includes alternative, open sources of metrics for quantitative evaluation of research and researchers, as well as Open Peer Review (Ross-Hellauer, 2017) for transparent, qualitative assessment of individual pieces of research.[3]

The review is guided by the methodological frameworks proposed by the PRISMA Extension for Scoping Reviews (PRISMA-ScR). As Tricco et al. (2018) state, scoping reviews are useful to "examine the extent (that is, size), range (variety), and nature (characteristics) of the evidence on a topic or question; determine the value of undertaking a systematic review; summarise findings from a body of knowledge that is heterogeneous in methods or discipline; or identify gaps in the literature to aid the planning and commissioning of future research". PRISMA specifies processes for searching formal databases of peer-reviewed material, as well as

[1] https://opendefinition.org/od/2.1/en/

[2] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en

[3] Note that the following elements are deemed out of scope for this review: Open Government Data and linked Open Data / semantic web are not in scope, as we here only include OS activities related to research. For this reason, we also exclude Open Educational Resources, which although sometimes grouped under OS, relates mainly to education rather than research. Elements of educational impact of OS resources will be brought up using the inclusion criteria in societal impact, however. Open Innovation (open processes within industry) is similarly out of scope. In addition, articles solely introducing data repositories without assessing their impact are also out of scope.

supplemental searching via hand-searching references of the included studies and references ("snowballing"), and structured web-search for "grey literature" resources like policy reports from major stakeholders.

The literature covered here encompasses a wide array of research methodologies, evidence, and data source types. This includes some studies where there is clear couterfactual evidence (e.g., randomized controlled trials), but also observational studies, surveys, case-studies and qualitative work. Hence, these types of evidence are not all equally suitable to support causal claims. Note that we perform a scoping review, not a systematic review, in order to identify which parts of OS and its impact is well studied and which is not. Hence, we even scope literature that might have methodological limitations. We do not conduct a systematic assessment of the strength of the causal claims made in these studies, nor do we evaluate the extent to which findings can be considered causal. Despite this, we do make an effort to point out particularly strong or weak examples of evidence throughout our review. This approach allows us to identify and highlight significant findings in the literature, while also acknowledging the limitations inherent in our methodological scope.

This deliverable presents the first stage of the literature synthesis - the database search. This deliverable is structured as follows: Section 2 "Methodology" details our process for identifying relevant literature, extracting relevant data and synthesising results (according to the PRISMA-SCR framework). Sections 3, 4, and 5, then detail the findings according to types of impact (academic, societal and economic respectively). Finally, Section 6 concludes by presenting an overall discussion and view on how work will continue towards finalised preprints and an online Zotero resource of relevant literature.

2. Methodology

Our protocol for the search was pre-registered on 31st October 2022 via the Open Science Framework.[4] The protocol covered search and data-charting strategy and ensures inclusion of all relevant and reliable literature concerning (1) the various elements of the initial PathOS impact pathways model (enabling factors, inputs, generating mechanisms, outcomes and impacts)[5], (2) all elements of OS, (3) varieties of impact including scientific, societal and economic, and including (4) all relevant data sources (including theoretical, scientometric, economic and qualitative work). In addition to the database search strategy (results of which are reported in this deliverable), the protocol also details methods for snowballing and search of stakeholder websites for additional relevant studies not captured in the database search.

The work was structured according to the following five steps:

1. Identifying the research question
2. Identifying relevant studies
3. Selection of eligible studies
4. Charting the data
5. Collating and summarising the results

2.1. Identifying the research question

The main research question is: *What **evidence** exists in the literature regarding the effect of OS on (1) academic, (2) societal, and (3) economic impact of research?*

Research sub-questions are:

- What types of positive or negative, direct or indirect academic, societal and/or economic impacts are observed?
- What kinds of mechanisms produce OS impact, i.e., what levels of correlation or causation can be observed between inputs, activities, outputs, outcomes and academic, societal and/or economic impacts?
- What specific enabling and/or inhibiting factors (drivers and barriers) are associated with these impacts?

[4] https://osf.io/m4rnc

[5] PathOS - D1.1 Open Science Intervention Logic https://zenodo.org/record/7801286#.ZEeCn87P1D8

- Which methods and/or indicators have been employed to assess academic, societal and/or economic impacts or estimate counterfactuals?
- What trade-offs between academic, societal and economic impacts are observed?
- What knowledge gaps emerge from this analysis?

This study used the PRISMA framework to align study selection with the research question and followed the relevant aspects of the PRISMA Extension for Scoping Reviews (PRISMA-ScR) to ensure thorough mapping, reporting and analysis of the literature.

2.2. Identifying relevant studies

Search was conducted for peer-reviewed published literature on the research area from January 2000.[6] Search was limited to articles published in English.

The authors developed a search strategy to conduct a search from 1 January 2000 of electronic databases (Scopus and Web of Science) for citations and literature using relevant keywords. Piloting by the authors identified the following search parameters (see Table 1).

Table 1 Concepts and respective search terms for identifying relevant studies

High-level concept	Element of OS	Impact	Academic Impact	Societal Impact	Economic Impact
Lower-level concepts	Open Science Open Access Open/FAIR Data Open Methods Open Code Citizen Science Open Evaluation	Effect Outcome	Efficiency Productivity Quality Education Reproducibility Reuse Citations Collaboration Equity, Diversity and Inclusion	Societal impact Trust Education/understanding Engagement Government policy Sustainable Development Goals Environment/climate Health COVID Participation	Economic impact Financial/monetary impact Cost/benefit analysis Input-output modelling Return on investment Productivity Innovation Patenting New products/services
Search terms	"open scien*" "science 2.0" "open data" "FAIR data" "open access" "open code" "citizen science" "open peer review" OR "open metric*"	impact* effect* outcome*	quality citation* integrity equi* collaborat* trust efficien* re-us* OR reus* productiv*	engag* educat* trust polic* sdg OR "sustainable development goal*" gender diversit* health environment* OR climat* covid* OR coronavirus* participat*	econom* financ* cost* mone* cba bca "input-output" "return on investment" "patent*" "innovation*" "efficiency gain*" "saving" "product*"

[6] OS only emerged as a concept, reform movement and policy priority since that time; hence limiting search post-2000 minimises false positives.

From this, the following primary search strings for Web of Science (WoS) and Scopus database searches were specified (see Table 2).

Table 2 Primary search strings for Web of Science (WoS) and Scopus

	ACADEMIC IMPACT	SOCIETAL IMPACT	ECONOMIC IMPACT
Web of Science	[To be run in all databases, for time 2000-01-01 to 2022-12-31] (TI= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR "open access" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR AB= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*" OR "open access publ*" OR "open access paper*" OR "open access journal*" OR "open access book*")) AND TS = ((impact* OR effect* OR outcom e*) AND (quality OR citation* OR integrity OR equi* OR collaborat* OR t rust OR efficien* OR re-us* OR reus* OR productiv*))	[To be run in all databases, for time 2000-01-01 to 2022-12-31] (TI= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR "open access" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR AB= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*" OR "open access publ*" OR "open access paper*" OR "open access journal*" OR "open access book*")) AND TS =((impact* OR effect* OR outcome*) AND (engag* OR educat* OR trust OR polic* OR (sdg OR "sustainable development goal*") OR (gender* OR diversit*) OR participat* OR health* OR (environment* OR climat*) OR (covid* OR coronavirus*)))	[To be run in all databases, for time 2000-01-01 to 2022-12-31] (TI= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR "open access" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR AB= ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*" OR "open access publ*" OR "open access paper*" OR "open access journal*" OR "open access book*")) AND TS = ((impact* OR effect* OR outcome*) AND (econom* OR financ* OR cost* OR mone* OR cba OR bca OR "input-output" OR "return on investment" OR "patent*" OR "innovation*" OR "product*" OR "efficiency gain*" OR "saving*"))
Scopus	TITLE-ABS ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open access" W/1 publ* OR paper* OR journal* OR book*) OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR TITLE ("open access") AND TITLE-ABS-KEY ((impact* OR effect* OR outcom e*) AND (quality OR citation* OR integrity OR equi* OR colla borat* OR trust OR efficien* O R re-us* OR reus* OR productiv*)) AND (PUBYEAR > 1999) AND (LIMIT-TO (LANGUAGE , "English"))	TITLE-ABS ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open access" W/1 publ* OR paper* OR journal* OR book*) OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR TITLE ("open access") AND TITLE-ABS-KEY ((impact* OR effect* OR outcome*) AND (engag* OR educat* OR trust OR polic* OR (sdg OR "sustainable development goal*") OR (gender* OR diversit*) OR participat* OR health* OR (environment* OR climat*) OR (covid* OR coronavirus*))) AND (PUBYEAR > 1999) AND (LIMIT-TO (LANGUAGE,"English"))	TITLE-ABS ("open scien*" OR "science 2.0" OR "open data" OR "FAIR data" OR ("open access" W/1 publ* OR paper* OR journal* OR book*) OR ("open code" OR "open software" OR "open tool*") OR "open method*" OR "citizen science" OR "open peer review" OR "open metric*") OR TITLE ("open access") AND TITLE-ABS-KEY ((impact* OR effect* OR outcome*) AND (econom* OR financ* OR cost* OR mone* OR cba OR bca OR "input-output" OR "return on investment" OR patent* OR innovation* OR product* OR "efficiency gain*" OR saving*)) AND (PUBYEAR > 1999) AND (LIMIT-TO (LANGUAGE , "English"))

Database results were retrieved on 2nd Nov 2022 (academic impact) and 8th Nov 2022 (societal and economic impact). Titles of results were first screened by one researcher to remove obvious false positives, by applying a broad inclusion approach (if the study may be at all relevant, it was included at this stage). Data regarding all studies from both database searches judged relevant via title search was then compiled and duplicates removed. In the next stage, title and abstract screening was undertaken by two researchers per aspect of impact who coded whether an article was (0) definitely out of scope, (1) unsure, potentially in scope, (2) definitely in scope. For economic impact, two reviewers undertook this step on all articles independently. For academic and societal impact, given observed high levels of agreement on articles out of scope in the process for economic impact, one reviewer first screened all articles and then the second reviewer independently assessed those articles judged 1 or 2. If at least one assessed an item to be of relevance (2), it was included (50% necessary percentage agreement). If not, the study was excluded and reasons detailed. In addition to screening articles for relevance, the first reviewer for each study also recorded to which research sub-questions the article was relevant.

Results from each database search were exported to dedicated libraries in the Zotero open source software. In Zotero, full-texts of all studies were gathered for enhanced screening. All reasonable attempts were made to obtain full-text copies of selected articles (e.g., via inter-library loans or contacting the authors). Where this was not possible, the study was excluded. Enhanced checking of full-text then determined whether full-text revealed the article to be eligible.

2.3. Selection of eligible studies

Title and abstract screening was guided by the PRISMA framework, with specific eligibility criteria applied to ensure relevance for the study and its research questions. The selection process followed the recommendations in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) checklist. The following inclusion criteria were used:

- Articles on the academic (study 1), societal (study 2) and economic (study 3) impacts of OS and its various aspects (OA, Open/FAIR Data, Open Methods, Open Code, CS, Open Evaluation)
- Conducted internationally or nationally
- Published from 1 January 2000 until current
- Available in English
- Full-text available
- Study is a research article, review article, conference paper, or other peer-reviewed article containing primary evidence of OS impact

- All types of methodology (quantitative, qualitative, mixed, etc.) are eligible

These criteria were applied in both title/abstract and full-text screening phases.

2.4. Charting the data

A data charting form was used to electronically capture relevant information from each included study. A provisional version of the data-charting form was included in the pre-registration and this was then updated following Title/Abstract screening and in communication with the broader PathOS consortium to identify the fields below, which were used to extract the data (see Table 3).

Table 3 Categories extracted from included studies in the data charting process

DATA CHART HEADING	DESCRIPTION
Author	Name of author/s
Date	Date article sourced
Title of study	Title of the article or study
Publication year	Year that the article was published
Publication type	Journal, website, conference, etc.
DOI/URL	Unique identifier
Exclusion	Out of scope, non-English, duplicate
Justification	If a study was deemed to be out of scope, a justification had to be provided.
Study details and design (if applicable)	Type of study, empirical or review, etc. Notes on methods used in study (whether qualitative or quantitative, which population demographics studied, etc.)
Types of data sources included	Detail the data sources
Study aims	Overview of the main objectives of the study
Relevance to which aspect of Open Science	Open Access, Open/FAIR Data, Open Methods, Citizen Science, Open Evaluation, Open Science General
Relevance to which aspect of impact	**Academic:** (Provisional list: Quality, Citations, Integrity, Equity, Collaboration, Trust, Efficiency, Productivity, Reuse) **Societal:** (Provisional list: Engagement, Participation, Education, Trust, Policy, Sustainable development goals, Gender, Diversity, Health, Climate/Environment, COVID-19, Equity, Empowerment) **Economic:** (Provisional list: Economic impact, Financial/monetary impact, Costs, Cost-Benefit Analysis, Input-output, Return on investment, Patenting, Innovation, Productivity, Saving)

Key findings	Noteworthy results of the study that contribute to the scoping review question(s)
Coverage	Optional field to note any relevant information about the level of coverage of the study, e.g., only specific countries, disciplines, demographics covered
Confidence assessment	Optional field to note any concerns about reliability/generalisability of findings (e.g., conflict of interest, potential biases, small sample sizes, or other methodological issues) within the study

2.5. Collating, summarising, and reporting the results

Following the data-charting process, the study team discussed main emergent themes and delegated responsibilities for drafting individual sections into the narrative report below, summarising the extracted data. These results are described in relation to the research question and in the context of the overall study purpose.

As per our protocol, each section (academic, societal, economic impact) was led by distinct teams of PathOS researchers. The distributed, collaborative nature of this work led to some (agreed) differences in overall processes across the studies. Most notably, while the sections on academic and societal impact each present results according to elements of OS (OA, Open/FAIR Data, CS, etc.), the study team for economic impact decided that given the limited number of papers identified, that section would benefit from a thematic structure according to identified key themes (types of economic impact, business models, sectoral evidence, etc).

Following the presentations of results, Section 6 presents a final synthesis across all studies including gap identification of areas where further research is required.

The data on all studies from the data-charting process (including studies that were deemed out of scope) is available at https://doi.org/10.5281/zenodo.7870402.

3. Academic impact of Open Science

Authors: Thomas Klebel, Vincent Traag, Lennart Stoy, Tony Ross-Hellauer

OS is often claimed to have benefits for academia more broadly (Nielsen, 2013), but also researchers themselves (McKiernan et al., 2016). In this study, we systematically scoped the evidence to date regarding the *academic impact* of OS. Our analysis deemed 311 studies to provide evidence towards impacts of OS on academia. The main type of impact reported concerns citations, but other types of impact, such as towards quality, efficiency, and equity, were also found (Figure 1B). Related to the various forms of impacts, the largest share of the literature obtained was concerned with the impact of Open Access (188 studies), followed by Citizen Science (63), Open/FAIR data (29). Fewer studies reported on the impacts of general forms of OS (15), Open Evaluation (8), Open Code (6), and Open Methods (2) (see Figure 1A).

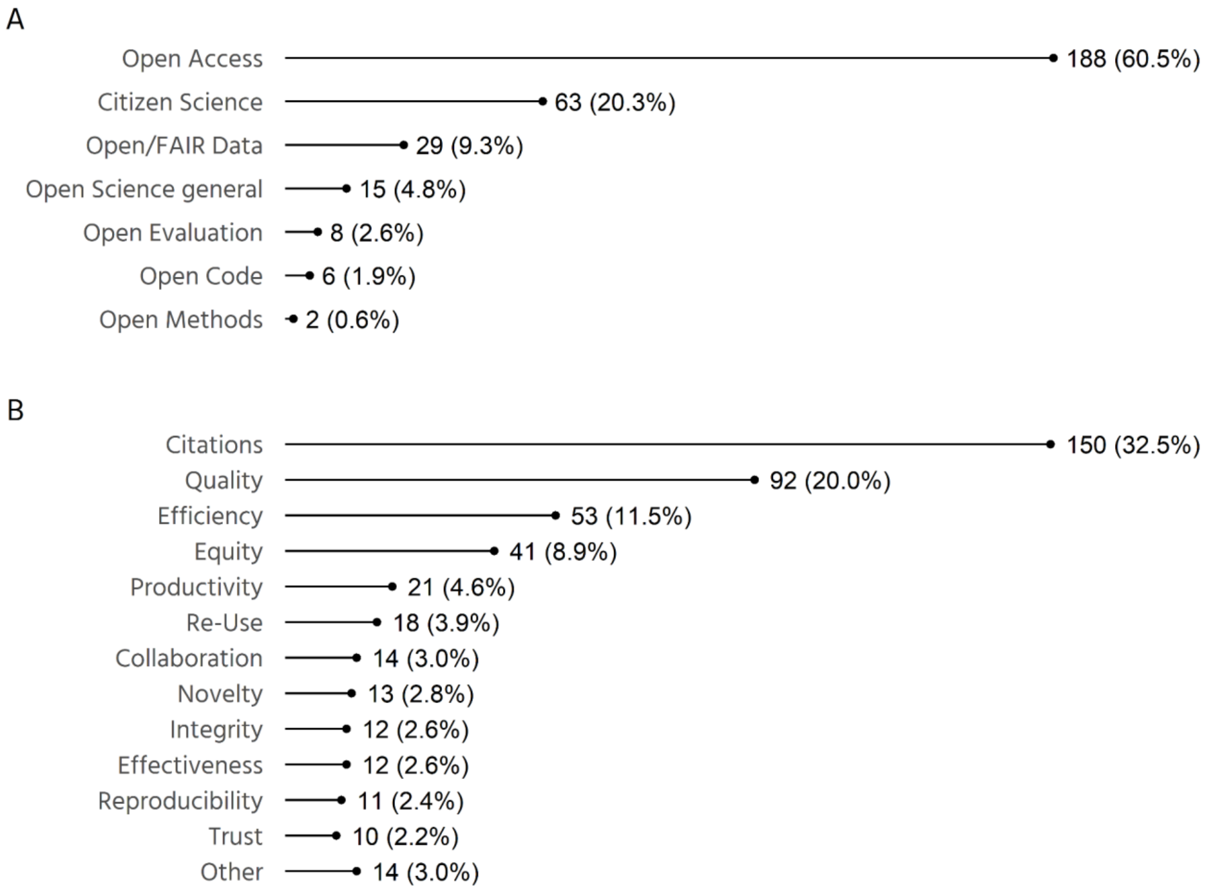

Figure 1 (A) Number of studies reporting academic impact of Open Science by relevance to aspects of Open Science (B) Number of studies reporting academic impact of Open Science by type of impact

Note that the numbers reported in (B) do not sum to 311 because some studies reported multiple types of impact.

3.1. Statistical summary

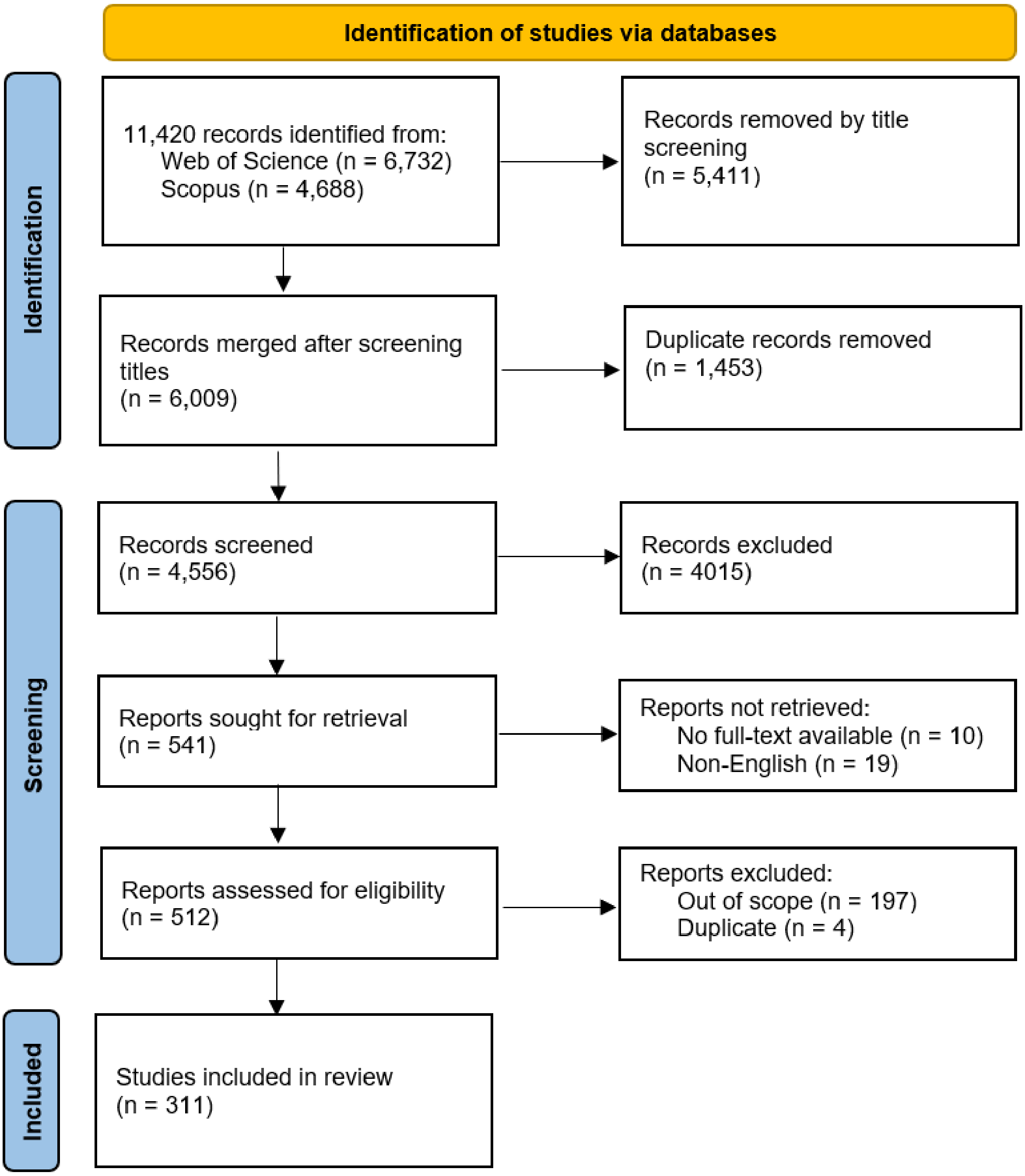

Figure 2 PRISMA diagram for scoping of academic impact

3.2. Open Access

Our review identified 188 relevant articles related to OA, making it the most studied element of OS for academic impact. Of these, most studies (136) investigate effects on citations and the so-called Open Access Citation Advantage (OACA). Other areas of inquiry include the effect of OA on quality (31 relevant studies), equity (28 relevant studies), productivity (15), efficiency (12), and further areas (19 codings towards various attributes). In the following, we analyse the retrieved literature along the following dimensions: Open Access Citation Advantage, equity in OA publishing, and the changing landscape of scholarly communication brought about by OA publishing.

3.2.1. Open Access Citation Advantage

The effect of OA publishing on citations, especially with regard to differences with closed access, has been well-studied. A major discussion has been whether there exists an "Open Access Citation Advantage" (henceforth OACA), the assumption being that easier access to literature leads to higher readership and subsequently higher citation counts. Given the very large number of studies identified as relevant to this topic (136 studies relevant to citations), we here mainly report on previously conducted reviews, complemented by more recent studies.

There is substantial heterogeneity in methodological approaches among the studies identified as relevant to the OACA in our review. Citation impact is measured in a variety of ways, comparing raw or standardised citations to articles, Journal Impact Factors between OA and non-OA journals, or comparing a number of further available indicators relying on citations. In addition, most studies tend to focus on certain types of OA (gold, green, hybrid, bronze, see Piwowar et al., 2018)[7] or certain research fields. Another source of heterogeneity when assessing studies on the OACA are the various definitions of OA itself. While the introduction of the Unpaywall service[8] has led to some standardisation in recent years, earlier studies relied on very different ways of defining and selecting individual publications or journals as being OA.

Key amongst the OACA-relevant studies identified in our review is the recent systematic review by Langham-Putrow et al. (2021), which synthesises evidence for and against the OACA. The authors report finding substantial heterogeneity in how studies aim to measure the OACA, for which reason they did not conduct a quantitative meta-analysis. Moreover, they found remarkably high levels of risk of bias. In their assessment, only three of 134 studies were found to have a low risk of bias across the domains "population", "data collection", "study design", and "results". One of the three studies found to have a low risk of bias detected an OACA in general, one reported an OACA in subsets of data, and one found no OACA. They conclude that there is

[7] In addition, "diamond" OA more recently refers to OA journals with no author facing charges, which can be understood as a subset of gold OA. Only one study identified in our review analysed diamond OA journals.

[8] https://unpaywall.org/

a “need for reporting guidelines for bibliometric studies”. Disregarding risk of bias to report across all 134 relevant studies, the authors did identify an OACA in general, reporting that 47.8% of reviewed studies found an OACA, with the remaining studies finding no OACA (27.6%) or only in subsets (23.9%) with respect to journal, discipline, or time. The authors did not find differences in how often studies reported an OACA by type of OA, but did find that studies with broader disciplinary coverage tended to find an OACA more often.

However, the degree of risk of bias across the literature surveyed by Langham-Putrow et al. (2021) shows that although the literature seems to indicate that OA is generally associated with higher citations, causal effects are difficult to substantiate. The main sources for risk of bias reported by the authors were poorly described samples or samples insufficient to support conclusions, and missing justification for the choice of study period (e.g., length of citation window). In addition, a large number of confounding factors have been suggested to affect estimates of the OACA, which were also apparent in our review (e.g., “Journal Impact Factor, number of authors, length of article, type of study”). Besides the confounding factors reported by Langham-Putrow et al., the studies identified in our review indicate further factors which might bias causal estimates of the OACA (Craig et al., 2007):

- **Selection bias**: In the case of hybrid[9] OA publishing, there is the postulate that researchers might only choose to make their best works OA, which in turn would lead to increased citations based on publication quality, rather than OA status (Gaulé & Maystre, 2011; Moed, 2007). However, Gargouri et al. (2010) find the OACA to be independent of self-selection, thus questioning the existence of this bias.
- **Early view bias:** Another postulate discussed in the literature is the early-view effect on citations - publications made available as preprints (i.e., prior to publication in a peer-reviewed journal) might receive more citations, since citations usually take some time to accrue (Moed, 2007).
- **Research funding:** Another potential source of bias concerns research funding, which often mandates OA. However, research emanating from publicly funded projects might be of higher quality than the average research conducted in a field, due to selection effects at proposal stage. Not controlling for funding could therefore introduce spurious effects (Lansingh & Carter, 2009).
- **Other confounding factors:** A common approach in studies assessing the OACA is to compare journal impact factors between OA and non-OA journals. However, these studies often do not control for other factors influencing impact factors, such as the extent to which journals are established and central to sub-disciplines. Comparing new

[9] Closed access journals that offer the option to publish manuscripts OA under an open licence by paying an APC are commonly referred to as “hybrid”.

(OA) to established (non-OA) journals regarding their citation performance ignores these aspects and leads to confounded and thus biased estimates.

Given this complexity, a major evidence gap seems to be that our review did not find any *systematic account* of the causal pathways related to the OACA. A systematised model of known and potential causal pathways might bring clarity to the study of the OACA and inform design choices for future studies.

The studies included in our review indicate that potential positive effects of OA on citations might differ according to the type of OA. Most studies on Gold OA journals find them to have lower impact factors than closed-access journals (Dorta-González et al., 2017; Dorta-González & Santana-Jiménez, 2018; Eger et al., 2021; Piwowar et al., 2018), but this might be partly driven by other factors than OA, including those above as well as the fact that many Gold OA journals are relatively new and thus have had less time to build reputation than closed access journals. Studies comparing citations received between hybrid OA and closed articles from the same journals tend to find an OACA (Abbasi et al., 2019; Sotudeh et al., 2019; L. Zhang et al., 2021; but see Mueller-Langer & Watt, 2018), but this might be driven by selection bias and effects of research funding. Studies assessing citations towards green OA articles tend to report finding an OACA (X. Chen et al., 2021; Clayson et al., 2021; De Filippo & Mañana-Rodríguez, 2020; Eger et al., 2021; Piwowar et al., 2018; Young & Brandes, 2020).

While we thus observe differences between the types of OA, it must be noted that these differences might not be systematic, as indicated by the analysis of Langham-Putrow et al. (2021). A systematic examination will be conducted in future work. Assuming that the differences uncovered in our review represent actual differences, our results suggest that it might be futile to investigate and claim an *overall* OACA, but that efforts should likely focus on context-specific accounts of an OACA.

3.2.2. Equity in Open Access publishing

There are multiple pathways models of OA publishing, but particularly prominent are those involving author facing charges, so-called Article Processing Charges (APCs). This business model has been identified as being a threat to equity in publishing very early (Papin-Ramcharan & Dawe, 2006), but is still rampant today (Alordiah et al., 2021; Siler et al., 2018). Of particular importance is the reported link between APCs and journals perceived quality, as measured by citation-based metrics. Given that APCs have been found to be positively correlated with journal metrics (Asai, 2019; Björk & Solomon, 2015; Ezema, 2021; Gray, 2020; Maddi & Sapinho, 2021; Schönfelder, 2020), it is evident that access to publishing in the most recognized journals is only available to those with sufficient funding. In addition, waivers have been found to be ineffective at countering this issue, in particular for researchers from lower- and upper-middle income countries (Asai, 2021). Taken together, these dynamics create a barrier for researchers aiming to publish their work, stratifying global publishing even further and undermining initial goals of the OA movement in terms of democratisation. Thus, even though OA increases equity in terms of physical access, it has been found to decrease equity on the side of publishing research.

3.2.3. Changes in the scholarly publishing landscape

The advent of OA publishing has been accompanied by multiple changes in the landscape of scholarly communication and publishing (see e.g. Eysenbach, 2010). While the internet and modes of online publication have enabled OA, the move to OA publishing has in turn led to new forms of publishing, such as OA mega-journals, predatory publishing, and an uptake in preprinting, which in turn has been sparking new modes of organising peer-review (such as the model of post-publication review powered by platforms like F1000, which was also adapted by the EC in its Open Publication Platform, or introduced at eLife in 2023).

A common concern in assessed studies touching on this area is with potential detrimental effects on quality in scientific publishing. Predatory publishing is by definition expected to lead to publications of lower quality, due to lack of basic editorial practices including peer-review. While empirical investigations of the quality of articles published in potentially predatory journals indeed find them to be of lower quality (Bianchini et al., 2020; Clements et al., 2018), these articles have been found to accrue much fewer citations than comparable articles, if any at all, and therefore have a low impact on science (Björk et al., 2020). The use of blacklists (such as Beall's list) to identify predatory journals has been subject of substantial debate and can be considered a secondary impact of OA publishing on equity, given its purported effect on "divisiveness, discrimination and stigmatization" (Teixeira da Silva & Kimotho, 2022).

Beyond predatory publishing, multiple studies have analysed the impact of OA publishing on article quality. While acceptance rates have been found to be higher among OA journals (Sugimoto et al., 2013), there is conflicting evidence as to differences in actual article quality. Some authors have found no difference in quality between OA and non-OA articles (Hall & Hendricks, 2020; Meerpohl et al., 2011; Pastorino et al., 2016), while others report higher quality among non-OA articles (Jerčić Martinić-Cezar & Marušić, 2019). In addition, evidence on the effect of OA mega-journals on article quality is sparse (Spezi et al., 2017). Retraction rates have been found to be higher among OA mega-journals (Erfanmanesh & Teixeira da Silva, 2019), but this is arguably related to better editorial practices rather than lower-quality research, as indicated by the finding that OA journals provide more detailed information on the reason for retractions (Peterson, 2013).

Finally, the move to OA publishing has changed business models of publishers, within established publishers and beyond. Quantifying costs and benefits of OA business models and self-archiving, Houghton (2009) reports "substantial net benefits in the longer term" in terms of academic, societal and economic impact for "more open access", due to increases in speed and breadth of access, as well as lower costs of publishing (a) for publishers and (b) for the whole system of scholarly communication.

A substantial driver for lower costs under OA publishing is the move from print to electronic distribution (Houghton, 2009). A publisher's revenue under gold OA depends on the number of accepted publications and the level of APCs, which might incentivise publishers to lower their acceptance criteria (Asai, 2019; van Vlokhoven, 2019)[10]. This has given rise to the concerns about article quality discussed above, but also led to business models where OA mega-journals cross-subsidise more selective and thus less profitable journals from the same publisher (Spezi et al., 2017).

3.3. Open/FAIR Data

Our review identified 29 studies which assessed the academic impact of Open Data. Of these, 13 studies reported impacts on data re-use, seven studies reported impacts on citations, five on equity, four on ethics, quality and reproducibility each, with further impacts on efficiency, productivity, collaboration, and trust.

Overall, few publications that are included in the review provide concrete evidence of impact of Open Data. Various publications study effects of data sharing policies, instead of effects of data sharing itself. The most obvious direct effect of data sharing is whether the data is actually reused, and most literature focuses on this. In contrast to OA, if data is not available at all, it simply cannot be reused, whereas research that is not OA can still be read and cited. Therefore, most studies do not compare usage statistics with a possible alternative situation or counterfactual. In principle, it would be of interest to study, for example, the reuse of data that is openly available with the reuse of data that is available only upon request. However, there are substantial practical limitations to detecting reuse of data that is available only upon request. Finally, some studies analyse whether making data openly available has some effect related to the publication that makes the data openly available itself. For instance, similar to OA, the so-called "citation advantage" of Open Data is studied. Other studies focus instead on whether studies that make data openly available increase the robustness of the results, related also to questions of reproducibility.

[10] Note however that many OA journals listed in DOAJ do not charge an APC at all (see footnote above on the terminology of "diamond OA"). Thus, alternative funding models for OA publishing which do not necessarily incentivise low acceptance thresholds exist (Spezi et al., 2017).

3.3.1. Data reuse

In a wide variety of fields there are ongoing discussions of data sharing. A number of publications examine the use of data sets from existing repositories, such as in agriculture (Ali & Dahlhaus, 2022), biodiversity (Khan et al., 2021), ocean science (Tanhua et al., 2019) and genomics (Xia & Liu, 2013). Other publications introduce novel repositories, such as for topography (Crosby et al., 2020), COVID-19 (Harrison et al., 2021) or neurology (Markiewicz et al., 2021). The publications referenced above do not represent an exhaustive list of repositories that are introduced, and only include publications that also report on actual usage.

Most papers discuss problems of data sharing, including challenges of common formats, standardisation of vocabularies and ontologies. Markiewicz et al. (2021) found that having common standards and a standardised vocabulary is a great benefit in data sharing. In particular, if there are common standards, this allow to more easily integrate information, and brings a larger benefit for disciplinary specific repositories, such as the OpenNeuro repository studied by Markiewicz et al. (2021) or the Genome Expression Omnibus repository studied by Xia and Liu (2013), over general repositories, such as Zenodo, Dryad or Figshare.

Reinertsen et al. (2021) reviewed Open Data initiatives using ultrasound open datasets to examine their contribution to advancement of intraoperative ultrasound-based navigation in neurosurgery. They concluded that Open Data supports new kinds of integration, as well as for algorithm optimisation and epidemiology. Two databases (BITE and RESECT) have more than 1,000 downloads and support more than 110 research publications (as of Oct 2020).

Khan et al. (2021) found that datasets from the Global Biodiversity Information Facility (GBIF) were increasingly used by researchers. The reuse of data, and citation to the data, seem to show a typical pattern familiar from citations to publications. It takes a few years before the use of data in publications becomes apparent. Ali and Dahlhaus (2022) found that many publications on Open Data in agriculture present little evidence for impact.

Other papers discuss challenges in making data reusable, as part of the Findable, Accessible, Interoperable and Re-usable (FAIR) definition (Wilkinson et al., 2016). Bishop and Collier (2022) found that researchers are influenced by “trusted brands”, such as well-known repositories, in trying to determine whether data will likely be re-usable or not.

A number of papers use theoretical models to explore the data sharing behaviour of researchers. Mueller-Langer and Andreoli-Versbach (2018) found that authors may strategically delay publications to fully enjoy the benefits of studying the data for additional papers. In their model, enforcing data sharing might have negative consequences due to the delay of relevant results, unless there are sufficient rewards associated with publication. Spiegelman (2021) studied a model of data sharing in low- and high-quality science. They find that high-quality science is more likely to share data than low-quality science, and that data sharing can then be used as a marker of high-quality science. Open Data policies, in this model, always improve the scientific quality of publications.

Piwowar and Vision (2013) found that publications that make data openly available are cited more frequently, and that a large part of that increase might be due to reuse of the original dataset.

Wiley (2021) examined the data sharing policies for a selection of medical journals. Although 60% of the studies' journals have some data sharing policy, only 17% mandate Open Data sharing. They found no difference in data sharing policies between low and high-impact journals.

3.3.2. Open Data citation advantage

Piwowar and Vision (2013) is the key study concerning the Open Data citation advantage, and is cited by most other articles for evidence of the Open Data citation advantage. They study gene expression microarray data, and find that publications that share data openly receive 9% more citations than publications that did not share data openly, while controlling for a number of relevant confounders. This contrasts an earlier finding of Piwowar et al. (2007) which found a much larger increase of 69%. They find that 6% of the citations were made in the context of data reuse, suggesting that much of the citation benefit comes from additional citations due to data reuse. The exact mechanism is not clear, but the presence of some effect seems quite convincing.

A number of other articles also study citations to publications that either share or use Open Data. AlRyalat et al. (2020) found that publications that used Open Data from BioLINCC receive an increasing number of citations. However, they do not compare their results to other publications, making it unclear whether this is a causal effect. Leitner et al. (2016) also observe that publications which share data openly are cited more frequently, but they only control for publication year, and not for a range of other confounders, thereby offering little evidence of a causal effect.

L. Zhang and Ma (2021) studied the effect of the introduction of a data sharing policy at a journal using a difference-in-differences approach. They found that the Open Data policy increased citations by 1-4 times. The estimate of receiving four times more citations seems extremely large, and quite different from the more conservative estimate of 9% (Piwowar & Vision, 2013). Nonetheless, the study corroborates the idea that there is some effect.

Kwon and Motohashi (2021) suggest that there are two competing effects of sharing data on the accrual of citations: a positive credit effect and a negative competition effect. While the credit effect leads to an increase in citations to publications that share data, it might lead to greater competition by other researchers, thereby crowding out citations to the original research. Both effects are reported to be present, where in the short-term publications that make data available are cited more frequently, due to the credit effect, but in the long-term, they are cited less frequently, due to the competition effect. This is an interesting perspective, but it is not immediately clear whether this represents a causal effect.

Raffaghelli & Manca (2023) found that reading or citing of the data does not seem to be associated with the FAIRness of the data, and that the distribution of reads and citations is very skewed, similar to citations to publications.

3.3.3. Reproducibility

One important reason for sharing data is that other researchers are then able to replicate the research. This is a large discussion in metascience, but only a few of those papers were identified via our systematic search of the literature. This gap will be filled at a later stage through snowballing.

Hardwicke et al. (2018) studied whether the introduction of a data sharing policy at a journal affected data sharing and the reproducibility of articles published in the journal. They found an increase in data-availability statements, but not all data appeared re-usable. Of the 35 articles for which data was re-usable, only 11 articles could be reproduced without author assistance, and an additional 11 could be reproduced with help. For the remaining 13 articles, at least one value could not be reproduced. Overall, 95% of the numbers produced in the articles could be reproduced (within a 10% error margin) using the open data. The actual reproduction of the figures did take considerable time and effort.

Naudet et al. (2018) analysed randomised controlled trials in BMJ and PLOS Medicine. Although only about half of the studies (17/37) shared data, most of those studies (14/17) could be reproduced. Again, the reproduction of the results required substantial effort and contact with the authors.

Nuijten et al. (2017) compared papers and journals with data sharing (policies) to those without such data sharing (policies), and studied how they differ in terms of statistical inconsistencies. They found that data sharing (policies) do not affect statistical inconsistencies (as found through statcheck), but that data sharing policies are effective at increasing data sharing.

3.3.4. Efficiency/productivity

Cannon et al. (2022) studied the effect of a new data sharing policy on submission rates, acceptance rates and peer-review times, and found no such effects. Holt et al. (2021) did find that the introduction of a data sharing policy increased the time editorial staff spent on processing a manuscript. After some modifications to the workflow these effects seem to have been mitigated.

3.3.5. Ethics and equity of data sharing

Cummings et al. (2015) studied whether Open Data sharing makes a difference to participants consenting to take part in research, but found no such influence. They do make an interesting observation that, if data is shared, participants' consent is usually only valid for the original study, while data sharing provisions allow the reuse of data for other purposes, without acquiring consent from the original participants.

Xia and Liu (2013) report a divide in the use of genomic data between the Global North and the Global South, with essentially none of the data being used by Latin America and Africa, indicating that benefits of shared data mainly accrue for the Global North.

Abebe et al. (2021) argue that those whose data is being shared might not enjoy the benefits, and that a deeper understanding of, and engagement with the local context, is necessary to make data practices more equitable. A similar argument is made by Carroll et al. (2019) who explored the intersection between data and tribal rights of indigenous people, arguing that indigenous people should regain sovereignty and governance over their data. Lastly, Bezuidenhout et al. (2017) report that access to data alone is not enough to make effective use of it. They highlight that "an emphasis on access fails to capture the social and material conditions under which data can be made usable, and the multiplicity of conversion factors required for researchers to engage with data." (Bezuidenhout et al., 2017, p. 473).

3.4. Open Methods

Regarding Open Methods, two articles were identified, relating to issues of integrity and reproducibility of science.

Bakker et al. (2020) investigated "whether the statistical power of a study is higher when researchers are asked to make a formal power analysis before collecting data" as part of a pre-registration template or institutional review guideline. Although this intervention led to higher levels of explicit statement that sample-size decisions were based on power analyses (72% versus 45% in the non-intervention sample), there was no actual difference in the planned sample sizes stated. Hence, the authors conclude that this intervention is not of itself effective in influencing sample sizes.

Next, Ebersole et al. (2020) used an experimental design to investigate whether the advance peer review of protocols for replication attempts influenced levels of replication success. They found, however, that overall "expert peer review had little impact on improving replicability across the 10 original findings we examined".

3.5. Open Code

Six articles were identified addressing this theme, five relating to efficiency gains in the production of research-enabling software via open source community efforts, and one relating to a citation advantage for publications with Open Software.

Regarding efficiency, Blasco et al. (2019) reported outcomes from three case-studies of use of open data science and development competitions for algorithm development in computational biology and bioinformatics. They found that use of open data and algorithm development competitions leads to better performing algorithms, especially through quick "exploration of the solution space" and creation of ensemble techniques through collaboration. Coetzee et al. (2020) reviewed open source software and open data initiatives in the geospatial domain, concluding that open source geospatial software and open geospatial data have changed data collection (including via crowdsourcing), processing, analysis and visualisation.

Similarly, Ratib and Rosset (2006), in their case-study of development, implementation and uptake of the open source OsiriX tool, reported that the tool's Open Source nature and embedding in a very responsive community of "specialists and professional users" equated to "software updates and new features at a rate that exceeded by far the rate of software updates in the industry". The authors do not rigorously quantify this claim, however. This, in addition to the study's focus on one case in one discipline (Biomedicine) means great care should be taken in interpreting or generalising this claim.

Mccormick et al. (2014) describe tools, methodologies, practices and usage statistics regarding the open source ITK (Insight Toolkit) software and associated community, in particular its code review feature. They report good uptake of this "reproducibility verification infrastructure" (e.g., 207 contributors, 2400 "unit tests") that uses open peer review for review of code commits, with positive effects reported (via fewer "fix-up commits" under this system).

Finally, regarding efficiency, Wallace et al. (2022) report on development of RT-Cloud, an open source cloud-based Python software package for real time fMRI experiments. They describe how "RT-Cloud has been integrated with open standards, including the Brain Imaging Data Structure (BIDS) standard and the OpenNeuro database, how it has been applied thus far, and our plans for further development and deployment of RT-Cloud in the coming years." Wallace et al. (2022) report efficiency gains in terms of ease of setup/maintenance, reduced costs and ease of scaling.

Regarding citation impact, Heumüller et al. (2020) investigated availability of software in Software Artefact papers from a Software Engineering conference. They found a slight citation advantage for papers which had made software artefacts available.

3.6. Citizen Science

While CS projects often have simultaneous impacts on citizens and science alike, in this section we restrict ourselves to documenting academic impacts (see following Section 5.2 for societal impacts of CS).

The vast majority of the 63 studies identified as relevant in this regard report on the impact of CS projects in Ecology, and especially in activities to monitor flora and fauna. The main types of academic impacts of employing CS in these monitoring activities relate to *efficiency* and *effectiveness* in data collection, as well as concerns over data quality. Academic impacts demonstrated by a range of studies are the increased temporal and spatial scales of monitoring activities enabled by CS methods, coupled with moderate to substantial cost-savings (Ashley et al., 2022; Lasky et al., 2021; Lawson et al., 2015; Roger et al., 2020).

Many state that CS can be a useful complement to existing approaches (Biraghi et al., 2022; Freihardt, 2020; Harvey et al., 2018). Although earlier literature, such as the case study by Alabri and Hunter (2010), as well as the review by Dickinson et al. (2010), reported *quality* issues with citizen generated data, more recent studies and reviews generally agree that CS projects can generate data of sufficient quality for monitoring activities (Leocadio et al., 2021; Njue et al., 2019; Quinlivan et al., 2020), with the main driver to data quality being the need to train volunteers appropriately (Adler et al., 2020; Manda et al., 2021; Yang et al., 2021).

A key question is which specific mechanisms drive the various forms of academic impact attributed to CS. While there are certain challenges related to organising large groups of volunteers, it is not surprising that more people will be able to conduct monitoring activities on larger temporal and spatial scales. It is thus clear that there exist trade-offs “between cost-effectiveness and involvement [i.e., societal impact], as well as between data richness and data quality” (Onwezen et al., 2021).

Besides the impacts on efficiency, effectiveness, and quality in data collection reported above, CS and crowdsourcing have also been suggested to offer further benefits, in leading to *novel insights* and *challenging existing approaches*, due to the diversity of backgrounds, perspectives and skills of citizens (Beck et al., 2022; MacPhail & Colla, 2020; Sauermann et al., 2020; Swan, 2012; van de Gevel et al., 2020).

3.7. Open Evaluation

Our search criteria identified eight articles relating to Open Evaluation, all on different models of open peer review, including disclosure of reviewer identities to authors (Open Identities) or a combination of Open Identities with review reports being made public (Open Reports).

Five articles investigated Open Identities, mainly investigating its effects on reviewer behaviour and review quality. Bianchi and Squazzoni (2022) constructed a theoretical model to simulate competition and status dynamics to investigate impact of Open Identities on quality of review. They found that under high levels of status-awareness and competition amongst reviewers, transparency of reviewer identities could compromise the quality and efficiency of processes. The authors themselves concede that these results of agent-based modelling are "only abstract and highly hypothetical", however. In terms of effects on quality, evidence seems to actually show neutral or positive effects. Bruce et al. (2016) carried out a systematic review and meta-analysis of the Biomedical research literature to evaluate impact of interventions to improve the quality of peer review. They found that Open Identities peer review improved review quality, did not affect time to review, but decreased rejection rates (implying that reviewers are more lenient when not anonymous).

Kowalczuk et al. (2015) undertook qualitative assessment of peer review reports using a "Review Quality Instrument" to investigate how "open or single-blind peer review models differ with regard to report quality and reviewer recommendations". They found that a comparison across two journals on similar topics from the same publisher, BMC Infectious Diseases (open peer review) and BMC Microbiology (single-blind), found quality 5% higher in the open review journal. However, examining reviews at Journal of Inflammation which used both models, there was no difference in quality across conditions.

Nonetheless, changes in behaviour due to Open Identities seem complex. Felizardo et al. (2022) conducted a survey to understand how respondents believed their review behaviour changed when not anonymous. The majority reported being more likely to write "bland and cautious" reviews that avoid issues of novelty/general interest and focus on more "objective" issues like technical concerns. Note, however, the study sample was very small (18 respondents - 12 reviewers, six authors), all from one Software Engineering workshop; hence caution is advised in generalising these results. Finally, Matsui et al. (2021) used sentiment analysis of review reports to assess (amongst other aims) how reviewer disclosure of identities (published optionally online after single-blind process) is associated with review sentiment, finding that reviews written by reviewers who choose to “sign” were “more subjective and more positive than the anonymous reviewers’ reviews". The authors contend this implies "possible social pressure from name association".

Three articles examined effects from both revealing reviewer identities and publishing reports. Van Rooyen et al. (2010) conducted a randomised controlled trial in the medical domain to examine effects on quality of reviews when informing reviewers that reviews would be published and signed with their names. Quality of reviews was assessed by two independent coders using a common instrument. They found no difference in review quality, but that reviews took longer to conduct in the “open peer review” condition.

Bravo et al. (2019) studied data from five Elsevier journals to find no significant change on referees’ willingness to review, recommendations, or time taken to review but did find that male reviewers tended to be more "constructive" under these conditions. They concluded that their findings "suggest that open peer review does not compromise the process, at least when referees are able to protect their anonymity." Hence, open peer review under these circumstances seems to show either neutral or positive effects on quality.

In addition, Zong et al. (2020) conducted a scientometric study from the PeerJ journal, where revealing identities and publishing reports are optional, to find that “articles with open peer review history could be expected to have significantly greater citation counts than articles with closed peer review history."

3.8. Open Science General

In this last section, we cover studies that demonstrated academic impacts, but did not fit the pre-specified categories, either because they addressed multiple types of OS at once, or because they took a broader perspective on OS at large. Our review identified 15 studies from general Open Science, demonstrating impact on diverse areas such as equity (4), quality (4 studies), re-use (4), trust (4), citations (3), collaboration (3), efficiency (3), integrity (3), productivity (3), reproducibility (3), and others (4).

An emerging approach to signify the application of OS practices and increase their recognition is to award ‘badges’ to journals or individual journal articles. These can display different types of adherence to OS practices, such as Open Data, Open Code etc. Schneider et al. (2022) studied the effect of badges in articles on different audiences, including scientists, students and the general public using an experimental design. They found a positive effect of badges on trust in results among scientists and students, but not in the general public. The potential cause of this effect is that openness or transparency of materials is already assumed by the public as the status quo. Given the different sampling populations across countries by Schneider et al. (2022), focussing on specific groups of students and researchers, this effect may not be generalisable without additional research.

Research on the impact of more general OS practices is often done in the form of case studies. Recent examples address in particular the impact of OS practices on COVID-19 research. Tse et al. (2020) for instance argue that Open Data, for instance by sharing protein structures, and Open Source practices, such as sharing information on potential inhibitors to COVID-19, have greatly contributed to accelerating research into COVID-19.

Similar claims are made for the use of (scientific) platforms which are based on OS practices. Anagnostou et al. (2019) provide the example of a platform, based on open repositories and digital research tools, for ageing research introduced in Ghana, which enables local researchers and policymakers to conduct research into local challenges and create insights which can be taken up in policy.

Coro (2020) provides another example in the field of the blue economy, where OS practices are found to affect the speed of execution of computational processes, collaboration and interdisciplinarity, virtual laboratories, combination of models, reuse of models, longevity of data and processes, as well as the dissemination of findings. All these case studies, however, come without a definition or measurement of impact.

Some evidence more explicitly documents direct, concrete links between OS interventions and intended results. Susanin et al. (2022) studied the effects of registered reports on adherence to OS practices. They found that "the use of data sharing, reporting of statistical form, and inclusion of startstop rules all increased following the implementation of registered reports" (Ibid., p. 1274).

Other topics investigated by case studies concern questions of efficiency and equity in research. One review article covering four case studies from Argentina (Arza & Fressoli, 2017) for instance, suggests that OS practices can support "efficiency, democratisation and social responsiveness". The authors also point out that "there are several directions of openness and they could lead to different types of benefits", implying that also partial openness can lead to benefits in terms of OS impact. The relationships between the different OS interventions and expected impacts must be carefully considered when developing models of impact pathways.

On the other hand, OS may also create negative effects. Notable warnings come from Hofmann (2022) and Ross-Hellauer et al. (2022). Both papers enumerate a range of risks associated with implementing OS practices and negative academic effects. These may include, for example, reduced quality, reduced inclusiveness and reinforced gender divides, dependence on specific skills and (digital) infrastructure. Such effects are difficult to quantify and this is not attempted in the literature. But as said above, being aware of the risks and potential negative impacts is important to reflect upon the assumptions underlying the impact pathways of OS.

The impacts reported in this section are necessarily less specific than in other sections. In the next iteration of this analysis, we will expand the evidence-base through further searches, potentially merging existing results into the more specific sections above.

3.9. Discussion

Our systematic search shows that there is a substantial body of literature (311 studies) investigating impacts of OS on academia. By far the largest share of studies is concerned with

OA, where in turn a large share of studies investigates the impact of OA on citations. There is tentative evidence for a causal effect of OA on citations, although the literature is heavily affected by issues in research design, which inhibits drawing definitive conclusions.

Besides citations, the APC-based OA business model has been found to pose a threat to equity by leading to lower geographic diversity of authors (Smith et al., 2021), thus exacerbating the exclusion of researchers from resource-poor world regions and institutions in global scientific publishing. Lastly, OA publishing has accompanied and enabled further changes in scholarly communication more broadly, with the rise of OA mega-journals, as well as forms of publishing that have been termed “predatory”. The latter have been reported to lead to the publication of lower-quality research.

Open/FAIR data has been found to have a positive impact on data reuse and on citations. Although Open Data contributes to (computational) reproducibility of published findings, actual reproduction of published results was found to require substantial effort and contact with the original authors.

Our systematic search uncovered only two articles investigating the impact of Open Methods. Both did not find a positive or negative impact of Open Methods in terms of replicability and research design.

Regarding Open Code, the main impacts reported were related to efficiency gains, with one study reporting a small citation advantage for publications with openly available software artefacts.

Research to date on the academic impact of CS finds benefits related to efficiency and effectiveness in monitoring flora and fauna. CS is often reported to provide a useful complement to existing methods of data collection, with recent studies concluding that CS projects can generate data of sufficient quality for monitoring activities. In addition, CS has been found to lead to novel insights and challenge existing approaches, due to the diversity of backgrounds, perspectives and skills of citizens.

Investigations into the effects of Open Evaluation (open identities of reviewers and/or open review reports) show mixed evidence on the effect of open identities on review quality and acceptance rates, with neutral to positive effects on review quality, but also an increase in acceptance rates which might be due to "social pressure".

In more general terms, OS was reported to increase trust in results by the research community and to lead to increases in efficiency, but also to potentially negative effects on quality and equity.

Many of the studies uncovered by our search discussed *potential* impacts of OS, often without empirically assessing *actual* impacts. This is in part due to methodological problems: while it is quite straightforward to compare effects of OA and closed access in terms of citations, it is for example much harder to assess the effect of Open Data versus closed data on reuse, since closed data that has been reused is hard to identify. Relatedly, many studies are not able to disentangle effects of OS from other potentially confounding factors, thus precluding substantive claims of causal effects. Lastly, research on the effects of OS is often conducted via case studies or based on small samples, preventing broader generalisations beyond the studied cases.

Given the context of our analysis, it is important to highlight that our results rely solely on the literature uncovered by systematic searches in academic databases. Some studies relevant to questions of academic impact have thus not been included, and will be added to later versions of the analysis.

4. Societal impact of Open Science

Authors: Nicki Lisa Cole, Eva Kormann, Tony Ross-Hellauer

Proponents of Open Science (OS) often claim that the creation of societal benefits is one of the reasons why it should be pursued and supported. Oft referenced benefits include the democratisation of science, positive impacts to learning through open outputs, empowerment through access to outputs, increased evidence-based policy-making, greater relevance of science to society's problems through collaboration and citizen science, and increased trust of the public in science. Yet to date no systematic study has examined the validity of these claims in their totality. Therefore, we set out to do so in this scoping review of evidence of societal impacts of OS.

We looked for a range of types of impact across OS in general, as well as in various aspects, including Citizen Science (CS), Open Access (OA), Open Code/Software, Open Evaluation, Open Methods, and Open/FAIR Data. The types of impact we looked for within the papers we analysed were chosen to cover all possible important aspects and differentiate between the most relevant societal issues. Categories were mostly selected to be exhaustive, but not necessarily mutually exclusive. They include the following:

- Social engagement: between citizens and scientists/other stakeholders, with scientific/project outcomes, and with the broader community
- Education and awareness: impact in the educational setting in terms of learning outcomes; knowledge acquisition about the topic of study; development of scientific thinking and skills; increase in awareness (through gains in knowledge)
- Trust: development (or breakdown) of trust between citizens and scientists or other stakeholders
- Equity: changes to equity (positive or negative) in terms of resource allocation, attention to problems related to inequality (e.g., through management of resources), impacts specifically in resource-poor or underserved settings
- Empowerment: citizens are able to access and work with government officials/policymakers due to participation or project results (empowerment through knowledge or social/cultural capital); participation builds community and change follows this
- Policy: demonstrated changes to policy, public administration or management
- SDGs: impacts specifically linked to meeting SDGs
- Gender: impacts to gender equality/inequality, or specifically gendered impacts
- Diversity: positive and negative impacts, e.g. broadening diversity of participants or creating exclusion from participation (e.g., demographics of CS overall, esp. crowdsourcing)

- Privacy/ethics: handling of sensitive data
- Health: demonstrated impacts on the health of people/public health (e.g., improvements to air quality, change in lifestyle practices, patient-informed healthcare treatment), improvement of health care (quality)
- Climate/environment: any and all impacts related to the climate and the environment (biodiversity, conservation management, ecology, pollution, etc.)
- COVID-19: societal impacts specifically related to the pandemic

Through the process of this review we found 155 papers to be in scope. Of these, the vast majority provided evidence of the societal impact of CS (137 papers, 88.4% of those reviewed), across a wide variety of types of impact (see Figure 3). Fourteen papers demonstrated societal impacts of OA, with impacts including public engagement with scientific literature, use in policy-making, and health-related outcomes. Beyond OA, our search revealed limited evidence of the societal impact of OS. We found just 2 papers that spoke to the impacts of OS in general, and 2 that demonstrated public health impacts of Open Code/Software. We found no evidence of societal impact of Open Evaluation, Open Methods, nor Open/FAIR Data.

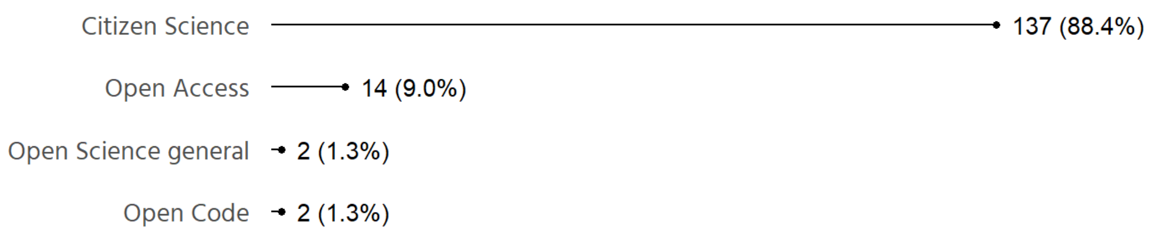

Figure 3 Number of studies reporting societal impact of Open Science by relevance to aspects of Open Science

4.1. Statistical summary

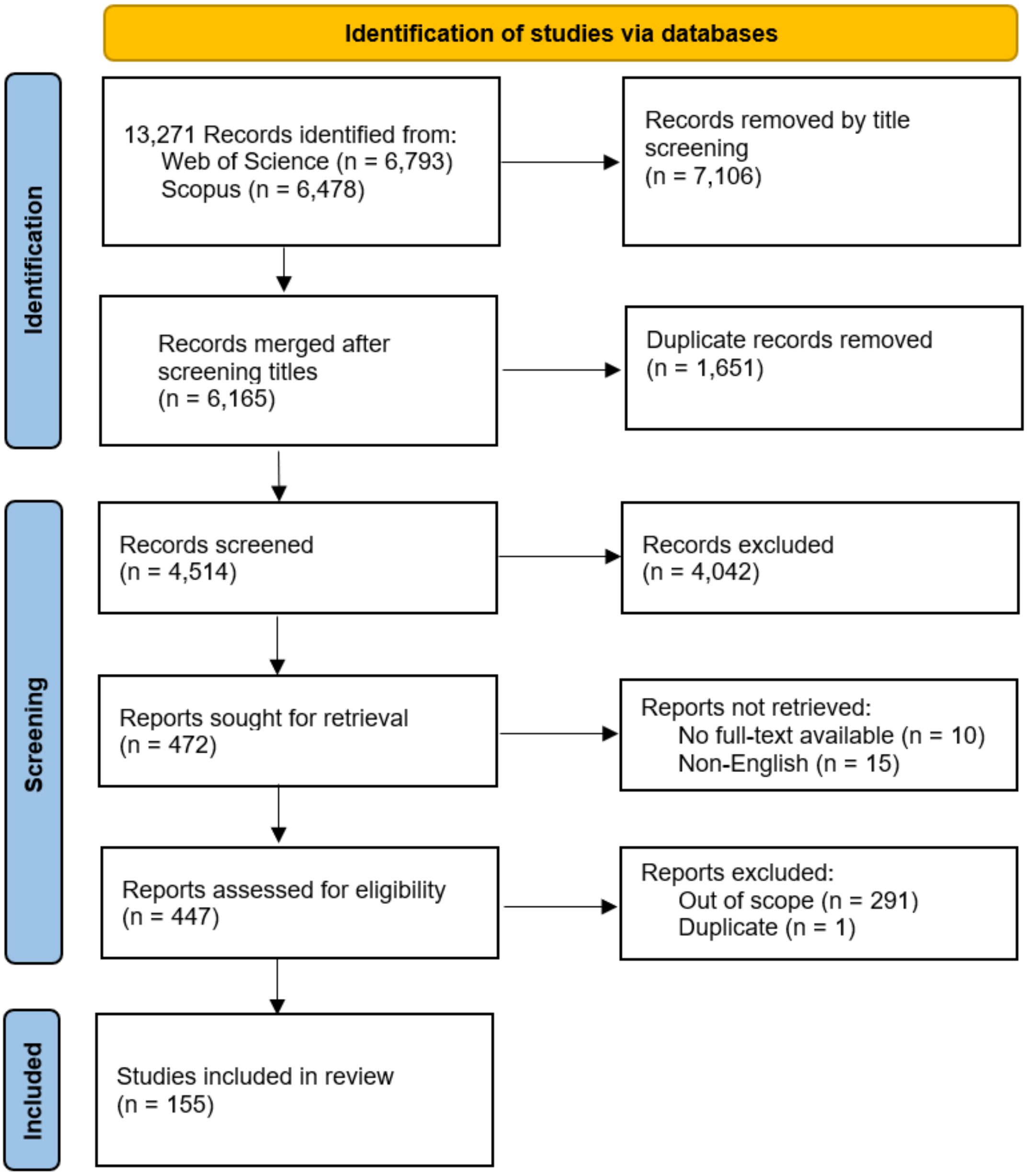

Figure 4 PRISMA diagram for scoping of societal impact

4.2. Citizen Science

CS is the practice of involving the broader public in the conduct of scientific research and is considered a key OS pathway to strengthening the relationship between science and society by directly involving the public in scientific research (European Commission, 2018), making science responsive to problems identified by communities and its outputs usable by and valuable to them (Allen, 2018), and fostering trust in science (Parthenos Project, 2019). Public participation through CS can happen at different steps of the research process or across it, from study design, to data collection, to analysis (English et al., 2018), and it can be applied across a range of disciplines and research settings.

Within the category of CS 137 papers were found to be in scope for our study. Among these, the most common areas of impact evidenced were education (65.0% of papers coded for this) and climate/environment (56.2%). Following these, more than a quarter of papers documented impact through engagement (27.7%) and policy (27.7%). Fewer papers demonstrated impact in terms of health (19.7%), empowerment (13.1%), trust (6.6%), equity (4.4%), and the sustainable development goals (SDGs) (1.5%). We identified no rigorous, quantifiable evidence of societal impact in terms of diversity, gender, or with regard to the global COVID-19 pandemic suitable for inclusion (see Figure 5).

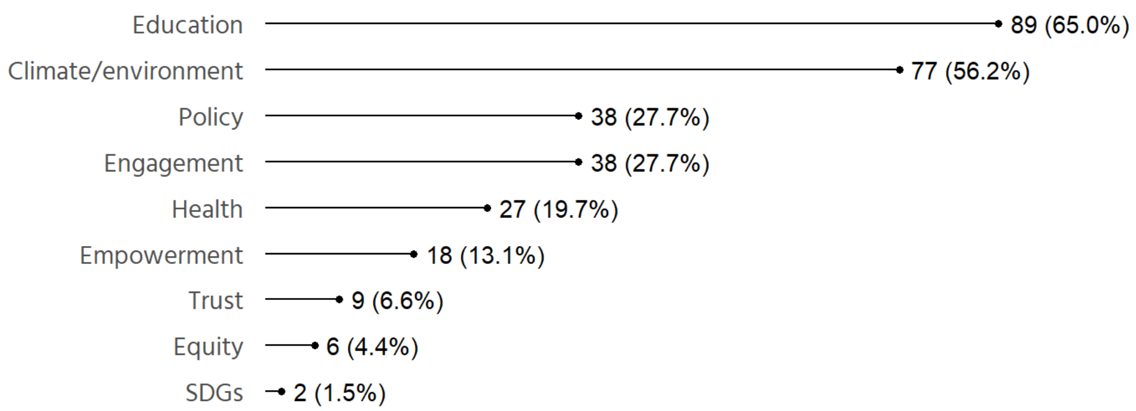

Figure 5 Number of studies reporting societal impact of Citizen Science by type of impact

4.2.1. Education and awareness

The vast majority of CS-relevant papers that were coded for societal impact through education and awareness (89 total papers) demonstrated that participation in CS leads to gains in knowledge and skills, in terms of the subject studied within the CS program or project, in terms of knowledge of science and scientific skills, generally, and even in terms of communication and organising skills (Bonney et al., 2016). Nearly all of this evidence indicates a gain in knowledge and skills, though a rare few show otherwise (Jordan et al., 2011; Meschini et al., 2021; Raddick et al., 2019; Shinbrot et al., 2022). The effects of CS participation on these outcomes were studied in a range of CS projects and programs, from CS initiatives in educational settings (from primary school through university), to crowd-sourcing, to community-based initiatives, and across the globe. Most studies in this category used a pre- and post-test methodology (typically surveys, but sometimes also interviews) to evaluate changes to participants' level of subject knowledge, understanding of science and the scientific process, scientific thinking, and/or scientific skills.

Some studies also found that participation in CS leads to greater interest in studying science and/or pursuing a scientific career (H. Cho et al., 2021; Johnson et al., 2014; Koomen et al., 2019; Lüsse et al., 2022; Rosas et al., 2022; Seifert et al., 2016; D. E. Wallace & Bodzin, 2019), though not always (Stewart et al., 2020). Additionally, Shaw (2017) found that a CS project impacted the educational setting by providing the context for development of biodiversity data collection techniques and principles that were adopted as part of curricula for courses across multiple universities.

Other broader educational impacts demonstrated by these papers include a 'multiplying effect' of knowledge gain, wherein knowledge gains by participants are 'multiplied' within the community (Frigerio et al., 2019) and the development of public awareness within communities as a result of CS activities (Costa et al., 2022; Johnson et al., 2014; Mahajan et al., 2022; Schaefer et al., 2020; Shinbrot et al., 2022).

Only one paper coded for education and awareness demonstrated a negative impact. In a review of CS applications in water science, Walker, Smigaj et al. (2021) found that knowledge gains can lead to an 'increased sensitivity to hazard', which can have a negative impact on sense of well-being and safety among participants and community members.

4.2.2. Climate and environment

Studies that investigate impact of CS on the environment and climate as one of the most pressing societal challenges have relevance to biodiversity, conservation, pollution and/or resource management (with just one paper outside of these dimensions, demonstrating impact in terms of climate resilience (Gotor et al., 2021)) (77 total papers). Across these categories, studies primarily demonstrate societal impact in terms of changes to awareness, attitudes and values, and to behaviour (though not always (Jordan et al., 2011)).[11] Studies have shown increases in awareness of, for example, human behaviour impacts on the environment and climate (Haywood et al., 2016), development of environmental stewardship values and attitudes (Ostermann-Miyashita et al., 2021), and changes to personal behaviour that support biodiversity, like gardening in ways that support rather than harm biodiversity (Deguines et al., 2020). According to Popa et al. (2022), there is evidence that certain traits among CS participants predispose them to change their behaviour as a result of participation: namely these are pre-existing 'strong environmental attitudes', and involvement in other conservation or research efforts. They also provide evidence that most behaviour changes are private or personal, rather than public-facing, and can be classed as lifestyle changes, like 'reduced consumption, recycling or picking up litter' (Popa et al., 2022).

A minority of studies documented activism related to environment or climate topics as an impact of CS (Popa et al., 2022; Ruppen & Brugger, 2022; Shinbrot et al., 2022), while slightly more demonstrated that CS programs or projects can result in community development around environmental or climate issues (Dhillon, 2017; Groulx et al., 2017; Haywood et al., 2016; Rodriguez et al., 2019; Rosas et al., 2022; Ruppen & Brugger, 2022; Sandhaus et al., 2018; Spellman et al., 2021). Shinbrot et al. (2022) also found that participation in CS can result in the development of a 'green identity' among participants.

4.2.3. Social engagement

Just over a quarter of relevant papers within CS demonstrated impact in terms of social engagement (38 papers). These include the fostering of engagement between CS participants and other stakeholders either within the research context or as a result of it (nine papers), project and participant engagement with the broader community (the majority of papers), and community-wide ripple effects.

To the first point, there is evidence that participation in CS strengthens relationships between all project stakeholders (Costa et al., 2022; Kerr, 2022; Peters et al., 2015; Rubio et al., 2021). CS projects in a school context can lead to lasting and impactful relationships between scientists and community members (Metcalfe et al., 2022), and foster new relationships and collaborations between stakeholder organisations (Van Haeften et al., 2021; J. Zhang et al., 2023). Additionally, there is evidence that CS fosters engagement of community members in policy-making processes (Bonney et al., 2016; McGreavy et al., 2016).

[11] Note that while we coded for and discussed awareness, broadly speaking, within the context of educational impacts, we track it and discuss it here when it is specific to climate and environmental issues, as a key way in which CS has societal impact in terms of climate/environment.

More evidence demonstrates that CS promotes engagement with the broader community in a variety of impactful ways. There is considerable evidence that participants of CS share their knowledge, project results and practical skills with their families, networks and communities. Studies also show that CS fosters engagement of the broader community in the programme or project and its outcomes (Haywood et al., 2016; King et al., 2021; Rodriguez et al., 2019; Rubio et al., 2021; Schläppy et al., 2017; Spellman et al., 2021).

Evidence also demonstrates that CS fosters and strengthens social ties and community. Engagement with CS can lead to stronger ties to place and connection and involvement with communities (Ekman, 2019; Evans et al., 2005; Haywood et al., 2016; King et al., 2020; Sandhaus et al., 2018; J. Zhang et al., 2023) and has been shown to increase social, human and political capital (Christoffel, 2020; Jordan et al., 2011; Walker, Smigaj, et al., 2021).

Additionally, ripple effects in communities have been documented as impacts of CS. In a review paper, Metcalfe et al. (2022) documented evidence of social and political change. Other studies have found that community action, activism and community group engagement follow CS (Dhillon, 2017; Diprose et al., 2022; Marchante & Marchante, 2016; Rosas et al., 2022; Spellman et al., 2021).

There is also evidence of negative impacts in terms of engagement. In some cases, as documented in a review by Walker, Smigaj et al. (2021), engagement may involve conflict (when findings pit the interest of one group against another, for example), the erosion of social capital (e.g., when the local knowledge of participants is not valued by researchers), and the over-burdening of the public with responsibilities that should lie with governments. Engagement may also be negative when participants experience disappointment with the project outcome (Kelly et al., 2020).

4.2.4. Policy and governance

Of the 38 papers that demonstrated policy and governance impacts (about a quarter of CS papers), the majority relate to impacts realised at local and regional levels, with fewer related to national or international policy and governance arenas. Most of these papers reported on projects within the domain of climate/environment, with a few others focused on health and infrastructure. We found evidence for a range of impacts, but the majority demonstrated the use of CS data by government agencies to monitor or manage natural resources, environmental and health risks, and the built environment (25 papers). Of these, two papers provided evidence that CS data are in use in monitoring SDG indicators (Fraisl et al., 2020; Soroye et al., 2022). In addition, some found evidence of CS leading to the development of new management techniques (Seamans, 2018; Shaw, 2017). Yet, as Peters et al. (2015) found in a study of the use of environmental CS data in New Zealand, there can be hindrances to this, like a "lack of systems in place within [...] agencies for integrating community data into environmental reporting."

Fewer papers demonstrated impact in terms of policy development (15) and just three documented CS impact in the creation of or changes to legislation (English et al., 2018; McGreavy et al., 2016; Zettler et al., 2017). According to Mahajan et al. (2022), evidence of policy impact from CS projects related to air quality is limited, and our review suggests this is a finding that can be applied to CS in general. Reporting on a study conducted on water quality in rural Maine, USA, policy impact may be hindered by political and corporate interests that conflict with the findings of CS projects (Segev et al., 2021). Fulton et al. (2019) found that getting official recognition of CS fisheries data at the national policy level in Mexico can be difficult, though it is impactful at the local level by informing the creation of "no take zones" and setting catch limits.

4.2.5. Health

We identified 27 papers demonstrating health impacts of CS. These papers are related primarily to environmental health risks (air quality, pollution, pests) and also to physical health (fitness, food and gardening, chronic disease prevention). Within this subset of papers, evidence of impact included the CS programme or project raising awareness, supporting improvements to health and safety and/or leading to behaviour choices that benefit health (like using CS air quality data to determine when to engage in outdoor activity (Mahajan et al., 2021) or choosing to cycle or walk rather than drive in order to improve air quality (Hodgkinson et al., 2022)). Evidence also demonstrates that CS effectively spreads awareness of health risks and ways to avoid them within communities, and that it can lead to changes in the lived environment that support improved health and safety. In a review paper, Walker, Smigaj et al. (2021) reported evidence that participation in some CS activities pose health and safety risks to participants (e.g., conservation monitoring in remote and/or dangerous locations). Therefore, CS might not only lead to improvements in health and safety, but also have some negative impacts on the well-being of participants.

4.2.6. Empowerment and equity

Evidence exists that CS can empower participants and communities and foster equity. About 13.1% of our CS sample of papers (18) demonstrate empowerment impacts. Evidence shows that CS data and/or project results can empower participants and community members to advocate for their interests in interaction with decision-makers, and to monitor the state of their environment. CS can also empower participants and communities to pursue and implement solutions to problems (Spellman et al., 2021) and lead to further rights and access to natural resources being granted (Chiaravalloti et al., 2022). Evidence shows that participation in CS can lead to participants developing leadership capacity and taking on leadership positions within projects and their communities (McGreavy et al., 2016), increases in self-efficacy (a person's belief in their ability to do certain things in order to achieve certain goals) (Koomen et al., 2019; Lüsse et al., 2022; Sandhaus et al., 2018; Seifert et al., 2016), and in a case documented by Hoover (2016), project training empowered participants through career development. Together, these impacts point to the democratisation of science, particularly when participants are involved in defining problems and research questions, and shaping research design.

In terms of equity, evidence shows that CS can achieve environmental justice in the context of environmental inequality, e.g., by returning rights over traditional fishing territories or improving neighbourhood infrastructures (Chiaravalloti et al., 2022; Dhillon, 2017; King et al., 2020; Rosas et al., 2022). Yet, Tubridy et al. (2022) observed that CS can in some cases "compound inequalities by transferring responsibility and blame for air pollution to those who have limited resources to address it", and Walker, Smigaj et al. (2021) documented similar evidence in their review paper. Additionally, some studies have documented that CS participant demographics overall point to inequitable participation opportunities, with wealthier and more privileged people more often targeted and better able to participate (in terms of time and resources) (Ross-Hellauer et al., 2022; Vasiliades et al., 2021; Walker, Smigaj, et al., 2021).

4.2.7. Trust and attitudes toward science

Several papers (9) demonstrated the impact of CS in terms of trust between scientists and others, and attitudes toward science in general. Bruckermann et al. (2021) and Christoffel (2020) provide evidence that CS leads to more positive attitudes toward science. Other studies have demonstrated that CS establishes trust between researchers and other stakeholders (Christoffel, 2020; Fulton et al., 2019; Jordan et al., 2011; Metcalfe et al., 2022; Nursey-Bray et al., 2018; Ramirez-Andreotta et al., 2015), that it increases trust in science (Walker, Tani, et al., 2021) and in local knowledge (Walker, Smigaj, et al., 2021). Yet, as Walker, Smigaj et al. (2021) point out in their review paper, there is also evidence that trust between researchers and other stakeholders can be damaged through CS when problems amongst stakeholders or with the project outcomes arise.

4.3. Open Access

For Open Access (OA), 14 articles were identified as being relevant regarding societal impact. Research identified on this topic mainly focuses on the areas of engagement of science with the broader public (eight of identified publications), influence on policy-making (eight papers), and health-related outcomes (three papers). Few articles were concerned with other impacts of OA on society, e.g., no studies were identified on education, or other SDGs. A lot of overlap could be found between the different areas of impact. Studies on engagement often include policy impact as measured by citations to policy documents (amongst other metrics), policy-making also covers health policy in some parts, and the finding on privacy is based within healthcare, while not directly reporting on health outcomes.

Related to the topic of greater engagement with the public, eight studies (including two reviews) investigated altmetric scores for OA in comparison to non-OA publications. Altmetrics include online mentions of publications on social media, in news outlets or on other platforms as a metric alternative to traditional citation counts (Araujo et al., 2021). These are often, not unproblematically, taken as a proxy for societal engagement. We elaborate in the discussion section of this chapter why we have reservations about them being used this way. Yet, considering the evidence surrounding Altmetric scores, Araujo et al. (2021) reported two studies in their systematic review reporting higher scores for OA publications. OA status of articles is also found to be specifically associated with increased mentions on Twitter (J. Cho, 2021; Dehdarirad et al., 2019; Tai & Robinson, 2018) and on Facebook (Dehdarirad et al., 2019), and more coverage in news outlets and on blogs (Dehdarirad et al., 2019; Tai & Robinson, 2018). Additionally, OA articles have higher rates than non-OA articles of reference on Wikipedia (Teplitskiy et al., 2017) and higher readership counts on Mendeley (J. Cho, 2021). Similar patterns could be observed for books, with OA books receiving more attention in social media networks (e.g., Twitter), in mass media, on blogs, on Wikipedia and on Mendeley compared to non-OA books (Taylor, 2020; Wei & Noroozi Chakoli, 2020). However, while findings indicate more online mentions of OA publications compared to non-OA, this does not yet show societal impact. Increased online mentions might also be caused by researchers themselves communicating with colleagues about scientific findings, which these studies did not control for. Engagement with the broader public can therefore not be inferred directly.

The impact of OA publishing on society was measured in three studies by the presence of OA publications in the reference lists of policy documents. OA books and articles were cited in policy documents more often in direct comparison to non-OA publications (Tai & Robinson, 2018; Taylor, 2020). Vilkins and Grant (2017) also report that Australian policy documents referenced OA publications proportionally more often than would be expected from the overall percentage of OA in scientific publications. The practice of making preprints openly accessible is also discussed as having negative implications for policy-making. Besançon et al. (2021) describe an enormous increase in the number of published preprints in research responding to the COVID-19 pandemic. However, they report that some preprints lacking proper reviewing had already been included in policy documents before being retracted due to quality concerns. Publishing preprints openly therefore bears the risk of unreliable or false findings being used as a basis for policy-making.

Evidence on further societal impact of OA publishing is thin and often only anecdotal evidence is presented. Regarding healthcare, one experimental study found that mental health professionals gained more knowledge when an article they were asked to read was freely accessible (Hardisty & Haaga, 2008). There were some indications that treatment recommendations were adapted more often when access to the resource was free compared to when it was not. How this relates to real clinical practice and OA, however, remains unclear. One study found medical images of transgender patients to be openly available on Google Images more often when they were published within an OA article compared to a non-OA article (Marshall et al., 2018). The authors could not obtain information on the content of informed consent from most of the studies they investigated. However, if participants are not informed about their images being publicly available as a possible outcome beforehand, strong concerns about privacy violations arise when such medical articles are published OA.

4.4. Other aspects of Open Science

For some aspects of OS, very few relevant articles were found. No articles were identified as relevant to Open Evaluation, Open Methods, or Open/FAIR Data, for example, while for Open Code/Software (2) results were also scarce. In addition, two articles were identified as broadly applicable to OS in general. We report these here.

Two articles were identified as relevant to societal impacts of OS in general. Firstly Ross-Hellauer et al. (2022) performed a scoping review of dynamics of "cumulative advantage and threats to equity" in OS, amongst which was a strand of critique concerning an "instrumentalist" logic of OS which has heretofore focused (in the view of some) more on epistemic and functional improvements than ethical and societal outcomes. Secondly, Rosman et al. (2022) examined OS's relationship to public trust in science in two studies. In the first survey study (of participants from a German general population sample), they found that OA and other OS practices are rated by the majority of participants as important and as increasing their trust in the scientists. In a second experimental vignette study, participants were presented with descriptions of research that signalled or did not signal the use of OS practices. Effects on trust were not conclusive across the two conditions, although the authors did interpret some indications of enhanced public trust when OS practices are employed.

Two relevant articles relating to Open Code and Software were identified. Bokonda et al. (2019) performed a (non-systematic) literature review to synthesise findings regarding adoption of Open Data Kit (ODK), an open source suite of tools for data collection and sharing that is free and does not require certification or a stable internet connection for usage, and is hence of particular use in developing countries. They found that this Open Source platform appeared to be most relevant in health contexts, with 11 of the 15 included papers in this area, with the remaining from agriculture (2), fisheries (1), and the "social domain" (1). They concluded from the evidence collected that ODK has been used in Kenya, Mali, India, Nigeria, Ethiopia, Madagascar, Tanzania, Mozambique and the Dominican Republic, where it had "helped to improve many health programs and systems".

Kobayashi et al. (2021) performed a narrative review of recent works related to the use of Open Source Software for COVID-19 pandemic. They found that Open Source projects including GNU Health, OpenMRS, DHIS2 and LIFE took actions enabling various activities (e.g., Contact tracing, Epidemiological reporting, Laboratory test management, Vaccination management, clinical management, Interoperability Resource (FHIR) resources, Symptom screening, Vaccine Delivery Toolkit). In addition, sequencing data was made openly available via GenBank and available for research, clinical test, drug/vaccine development, and Johns Hopkins University developed an interactive dashboard from which multiple downstream resources for "data visualisation, analysis and decision-making" were created. The authors conclude that such tools were key in enabling governments to "study the causes and the impact" of COVID, offering " a way to leverage the collective intelligence of human beings to overcome [the COVID-19] pandemic".

4.5. Discussion

Overall, our study shows that while there is quite a bit of evidence of the societal impact of OS, it is primarily contained within CS. While we found some evidence of societal impact stemming from OA, some of it (regarding Altmetrics) is questionable and therefore it is rather limited. Further, we found just a few papers suggesting societal impact from OS in general, and from Open Code/Software. We found no evidence of societal impact from Open/FAIR Data, Open Methods, or Open Evaluation. This does not mean that these aspects of OS are not producing societal impact, but rather, we can conclude that their potential societal impacts have to date not yet been studied (as far as we have found).

The findings regarding the societal impact of CS are encouraging. They support the conviction that OS (through CS) is democratising the creation of and access to scientific knowledge. The evidence that shows increases in topical and scientific knowledge and skills points to this, and it shows that these gains are not limited to participants, but that they multiply through social networks and ripple into the community. Further, the evidence related to participation and empowerment shows that CS has the power to respond to problems within communities (by centring local knowledge and community perspectives) and foster the ability of those involved to continue to be change agents after the project is over. This is promoted in part by the engagement impacts that create, foster and strengthen social ties between CS participants, community members and other stakeholders. When this occurs, communities can be catalysed to make change together and are equipped with more social and political capital to help them do so. These impacts can be beneficial to all types of communities and all types of people, but we note that they can be especially impactful, from an equity standpoint, when they occur in marginalised communities and under-resourced areas.

Our findings also show that CS has the power to address 'wicked problems' – those without singular formulations and solutions (like climate change and inequality) (Rittel & Webber, 1973) – by shifting awareness, attitudes, values and behaviour around environmental, climate, and health issues. The evidence shows that participation in CS can prompt lifestyle changes that support more sustainable and healthy ways of life, and can lead to further community action and activism around these issues – giving truth to the belief that 'knowledge is power'.

Our findings show that CS impacts on the monitoring and management of resources (primarily environmental), including indicators with relevance to achieving SDGs. While this shows a positive societal impact of CS, it also raises ethical questions about the responsibility of governments to do this work and reliance on often unpaid volunteers to fulfil this responsibility.

Beyond this, we found limited evidence that OS impacts policy-making. While the evidence shows that some CS has impacted policy-development, the evidence is slim and challenges to integrating CS data and findings into this process were documented. Similarly, our study found just three papers that demonstrated the impact of OA on policy-making. While this evidence did show that OA had a significant impact on the use of academic literature among policy-makers in Australia, it is clear that this area is under-studied, with many references to the policy impact of OS being "promissory" in nature (Reichmann & Wieser, 2022).

We also have concerns about the lack of strong evidence that OA fosters public engagement with science. While our study found a few papers that demonstrated that OA has an impact on public engagement as measured through Altmetrics, we note that these studies do not study the demographics of social media users who post and repost about OA articles, therefore one cannot rule out that this evidence of engagement is largely contained to academic social media networks. Absent demographic analysis of users (and/or network analysis), it cannot be taken as hard evidence of *societal* engagement. Additionally, we note that engagement with OA material is not evidence of it having an impact on the reader. This also appears to be an area that is largely under-studied.

We found evidence that OS can provide societal impacts in terms of health and healthcare. Our findings show that CS fosters improvements to health and safety, positive behaviour changes that support health, and raises awareness of health risks and ways to avoid them. We also found evidence that OA supports evidence-based healthcare delivery, and that Open Code/Software supports healthcare delivery in resource-poor areas and public health management. While these are encouraging findings, we emphasise that this subset of papers is quite small, which suggests again that the impact of OS on health and healthcare is an under-studied area. Additionally, we note the privacy and ethics risks that arise when OA publications contain images of study participants. Risks such as these require attention from a research ethics standpoint, given the increasing emphasis on and growth of OA publishing. Privacy violations that do occur as a result of OS practices run the risk of undermining the trust in science that OS, in particular CS, is found to foster.

Regarding Open/FAIR Data, our finding is that there is no evidence of it having societal impact. This, however, is shaped by our exclusion criteria. We found, through our screening process, considerable evidence of the societal impact of OGD, which we ruled as out of scope, because our focus is on understanding the societal impacts of Open (scientific/academic) Data. In addition to ruling out data generated and shared openly by governments and their institutions, we excluded papers that focused on data shared by organisations like NASA, the United States Geological Survey, the Brazilian Institute for Space Research, and COVID data released by public health ministries. We classed the data released by organisations like these as OGD because they are government organisations. However, it is possible that some of these data were generated through collaboration with academic researchers, and therefore may not be purely OGD. We intend to review this subset of data in the next phase of this research process to re-evaluate its validity to our research. Nonetheless, we are still faced with a lack of evidence of societal impact from (academic) Open/FAIR data. This may be in part because of the challenges that arise when one wishes to study it. How would one track the societal impact of a dataset released by the average academic? How would one compare the societal impact of Open Data with closed data? Questions like these deserve thoughtful consideration.

In sum, there is considerable evidence within academic literature of the societal impact of OS, but it is almost entirely derived from studies of the impact of CS. A few studies focused on OA and Open Code/Software also show some positive (and some negative) societal impacts. We conclude from this evidence that aspects of OS other than CS need to be widely and deeply examined in terms of societal impact. We expect that we will find more evidence of societal impact, across OS aspects, when we conduct snowballing and examine grey literature in the next phase of our research process.

5. Economic impact of Open Science

Authors: Istvan Karasz, Sofia Liarti, Lena Tsipouri, Silvia Vignetti

Economic benefits from OS occur thanks to more effective or efficient business research affecting firm level, entire sectors as well as the economy as a whole. They occur via lowering costs and increasing speed of knowledge dissemination, to the extent that OA contributes to the incorporation of the openly offered scientific results into direct productivity gains and/or knowledge and capacity building. Open/FAIR data and Open Methods can be invaluable for avoiding duplication of experiments. In certain cases, they generate sufficient aggregate data so that together with Open Evaluation and Assessment they allow companies to benchmark and better assess their own performance, generating an indirect contribution to economic benefits via better strategy. Open Source Software is the most direct contribution to productivity increase thanks to free of charge inputs to firms' research and new products and services. CS does not have a direct impact but indirectly contributes to economic life to the extent that it improves academic results.

The systematic scoping generated enough publications but after screening only 13 remained, which were explicitly referring to economic impact (see Figure 7). The number is so limited because most papers reviewed were excluded, because they are referring to economic evidence using a theoretical rationale, which explains why the academic and societal impacts eventually can turn into economic benefits There is, however, seldom quantitative corroboration of this rational. As a consequence, most of the originally identified papers had to be rejected. Some of them were using the term "economic" were rejected because they referred to efficiency gains generated by the charging models of publishers, which are extensively analysed under academic impact. Others used OS as a synonym to Open Innovation claiming that industry collaborations limit firms' potential to appropriate new knowledge; in this sense companies consider any collaborations as *open science*.[12]

The 13 remaining papers provide only scarce evidence regarding the effect of OS on the economic impact of research. Most of the papers addressing economic impacts are those that refer to OA. The eight papers in this category were selected because they discuss how the new models may affect the business sector and thus innovation. There were six Open/Fair Data papers which refer to the relevance of data generation for industry, the role crowdfunding can play for research, the perception of specific countries (Canada, India) and an initiative at city-level. There are five papers on OS, as a general term, one of which is the only thorough

[12] Although, note that this does not fit our definition.

economic literature review from the past, one relevant for new medicines and two which see business collaboration models as a path to openness as a source of economic gains. One each in the areas of CS and Open Evaluation refer to an agro-food project and the experience in Taiwan (see Figure 6).

Yet, in most of these papers there is no hard evidence but rather a discussion on the perception of types of positive or negative, direct or indirect impacts, the mechanisms producing an OS impact, specific enabling and/or inhibiting factors (drivers and barriers) and methods and/or indicators employed in the literature. What was hardly found was explicit quantitative, generalisable evidence.

What we were looking for, namely the effect of OS, in the form of business impact (turnover, profits), labour income (or job creation/destruction), cost savings and fiscal impact (taxes, incentives) was missing. Hardly any specific indicators and no counterfactual analyses were found.

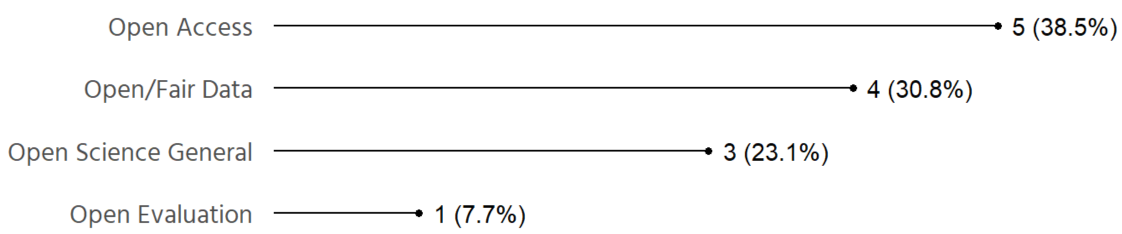

Figure 6 Number of studies reporting economic impact of Open Science by relevance to aspects of Open Science

5.1. Statistical Summary

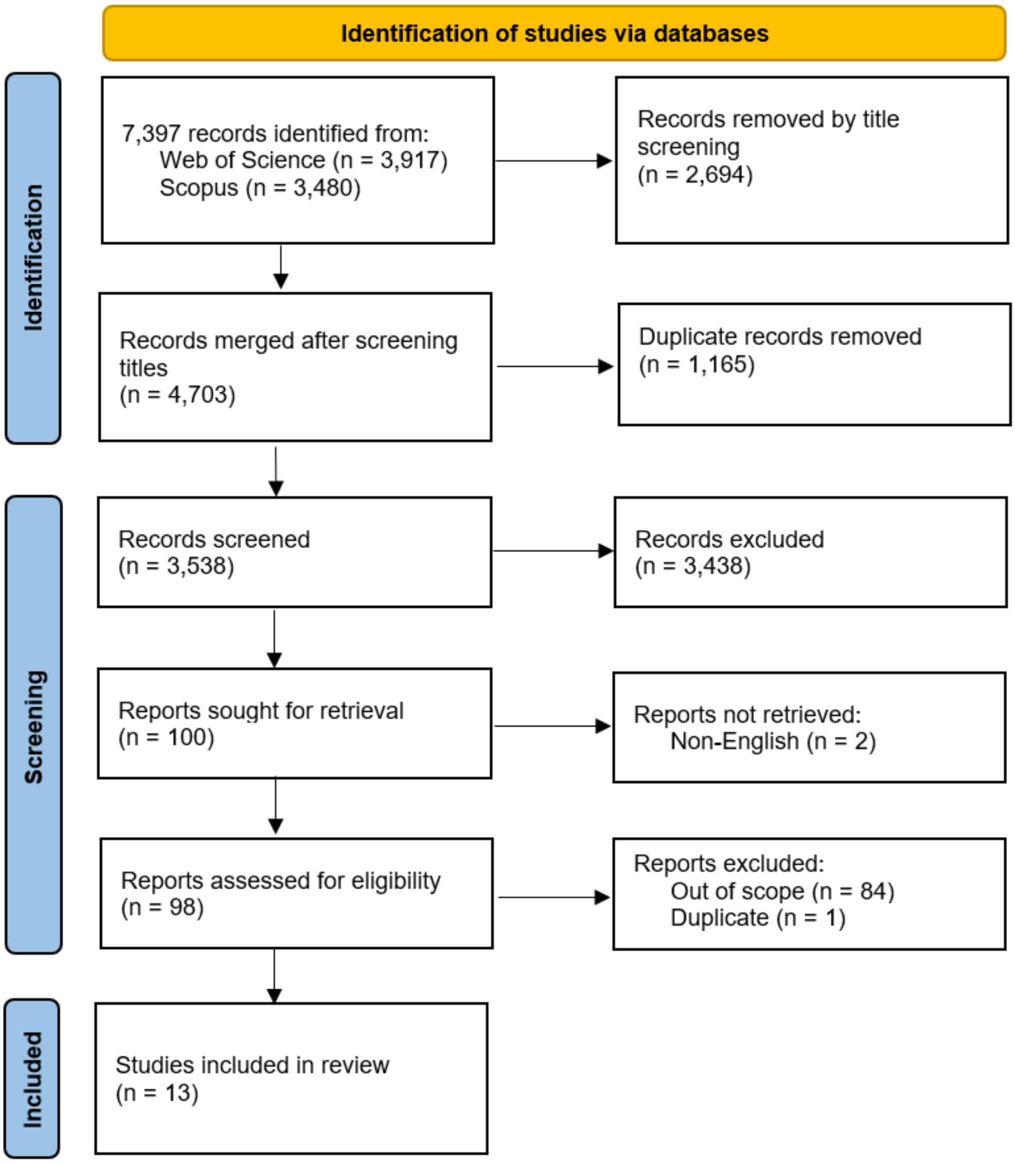

Figure 7 PRISMA chart for identification of economic studies

5.2. Types and mechanisms of Open Science Impact

A thorough review of economic impacts of OS published in 2019 remains the most comprehensive stock taking and a good explanation why quantifying economic evidence is so challenging (Fell, 2019). There is indicative evidence that OA to findings/data can lead to savings in access costs, labour costs and transaction costs. There are examples of OS enabling new products, services, companies, research and collaborations. Modelling studies suggest higher returns to R&D if OA permits greater accessibility and efficiency of use of findings. The paper identifies cost-benefit analyses, which provide an indication into what different levels of accessibility/efficiency improvements might mean for value in years to come and specific case studies in the health sector where results are positive but not generalisable. Other aspects include cost savings (easier access to knowledge) and reduced transaction costs (Fell, 2019).

Since Fell's paper there is very limited identifiable, concrete, measurable economic impact of OS and there are good reasons for that. Only a few papers have attempted either measuring OS impacts compared to closed science. Unlike academic and societal impacts, where tangible indicators can be obtained (citations, number of citizens involved, etc.), in economics assumptions and hypotheses need to be made if they are expected to lead to concrete numbers. This is why existing evidence is based on interviews, surveys, inference based on existing costs, and modelling approaches (Fell, 2019). However, the business sector has never been an enthusiastic proponent of publishing research results, as suggested by investigating academic publications and patenting after collaborative industry-academic projects (Bikard et al., 2019).

The main argument in terms of economic impact of OS is that the cost and time for accessing new knowledge is reduced and this facilitates both research and diffusion of new knowledge contributing directly to more and better research hence productivity increases. Productivity and efficiency gains that OS creates are of crucial importance and related data provides evidence that businesses can include these gains into their business models, in case they are exploiting the advantages of OS. The reduced number of duplications and data and text mining account for a significant reduction of costs that can be built in the annual planning. This is equally relevant to private and public research entities (Fell, 2019).

There is no confirmation of significant economic impact of open data as pointed out by a recent study in Canada (Boudreau, 2021). Drawing on an analysis of government measures and community practices in the field of data reuse, the study shows that the benefits of Open Data appear to be inconclusive in terms of economic growth.

Another area of importance for economic impact is the role of openness for new sources of funding research. Academic Crowdfunding was studied in Japan as a means used by research candidates to raise funds for a variety of projects online from other citizens. The study examined the characteristics of funded projects on four Japanese platforms that facilitate academic crowdfunding, which can be exploited at a later stage for business creation. Correlation is high between the number of backers and the amount raised but not with the achievement rate. Most of the projects requested for academic crowdfunding are in biology, art and design, and physics. Such information was claimed to be useful not only for future research candidates but also for universities and research institutions to design their financial support for researchers (Ikkatai & Ono, 2018).

Indirect economic impact can come from better research policies. OA along with AI technology can be used for the evaluation of the impact in different fields and thus improve policy making and contribute to higher productivity emanating from better designed research in the business sector. In Taiwan, Chen et al. (2017) used an analysis for cross-referencing secondary (published literature, research results and news information) and Open Data with big data mining. According to the results of this study, the use of text mining techniques in Open Data could help in policy planning.

One argument was found at the territorial level presenting a business plan expecting high economic benefits via OS (Arnone et al., 2016): The Lazio Pulse initiative was created based on the argument that a dynamic ecosystem of public and private actors for improving Research and Innovation, based on the value and knowledge generated by cross-disciplinary Open Scientific Data would be highly beneficial to the area, which through open innovation, can contribute to the national and international competitiveness of the Lazio Region by supporting growth of new businesses.

The main plausible theoretical argument that OS reduces cost and time of accessing research results, therefore it improves research productivity of the business sector avoiding duplication and speeding up research. Access to the research literature is key for innovative enterprises (Tennant et al., 2016).

Although quantitative generalised conclusions cannot be drawn, the literature points at two main directions where there is more evidence as analysed in the two sections below:

- The relevance of openness in relation to APC models is analysed in more detail in the next section. If different business models show small differences in expenditure (Bruns et al., 2020) then the impact of OA is unlikely to be very high. However, if the cost for access is significantly reduced more and better research might be expected in the

business sector. More research is needed, as access is less well-understood (Tennant et al., 2016).
- The literature focuses on a few sectors only. Health, medicine and biosciences are the most frequently encountered because of early regulation by funders and high interest in clinical trials results. Data-driven sectors and the agri-food industry are also touched upon.

5.3. Impact of Open Science for business models: potential repercussions of APC models

In general, there is a lack of literature on the overall direct impact of OS to business R&I decisions with some exceptions, mostly related to OA and article processing charges. The analysis of OA pricing was out of scope during the economics literature review; included were only papers in the case where publishers' new and different APC pricing models were conceived as impacting stakeholders' behaviour in economic decision-making. Similarly, papers were included, which highlighted a developmental economic perspective and how new business models of publishers may change global inequalities through the enlargement of scientific dissemination and the potential inclusion or exclusion of research outputs from the Global South.

Compared to direct costing models where subscribers pay for the articles, in case of OS, different OA models are implemented that require a reverse and more complex logic as well as a different approach. In the case of OA, APC requires researchers or their hosting organisations to pay for their articles. Shifting costs from readers to authors, however, can have several side effects that need to be addressed from the side of policymakers, publishers and of other key stakeholders.

The benefits of the lower costs for accessing research results are a main benefit analysed above. There is however a reverse side of the coin. One of the key issues is the unequal burden that APC sets on research intensive organisations while it reduces expenses by a significant amount, posed on research consuming organisations. For this reason, several organisations decided on cost sharing models (Bruns et al., 2020). This affects not only academic research but business research as well, since companies have very different research intensities.

5.4. Sectoral evidence

Scientific literature is a source of strategic knowledge and information that paves the way for new ideas to be explored in industrial research and innovation activities. Malfunctioning of the traditional publishing models, for example delays and bias in data publishing, can affect the productivity of private research, as discussed in Harding (2017).

We found, however, little empirical evidence of positive benefits of OA and Open Data on industry. One interesting piece of work is the one discussed by Bryan and Ozcan (2021). They performed a natural experiment to examine how the OA mandate of NIH for biomedical research in 2008 affected industry use of academic research results looking at in-text patent citations. The authors examined more than a hundred thousand articles in 43 medical and biotech journals and US patents applications since 2005 that were published as of 2015 and estimated patent citation propensity. They built a model of the invention production function and estimated a difference-in-difference in patent citation propensity of open vs non-open articles, using in-text citations. The results of the analysis show that, after the OA mandate, patents cite NIH-funded research 12 to 27% more often. This provides for a quantitative measure of the effect of OA on patent citations. Interestingly, academic citations see no change, however (see above Section 4.2.1 on the Open Access citation advantage). Despite the work not attempting to place an economic value on the increase in citations induced by OA, the model suggests that a potential social loss (in the economic sense, where placing a social value means translating into a monetary value with some economic evaluation) can be associated with expensive subscriptions to academic journals.

Similarly, to OA, the benefits of Open Data on industry are also not well documented. As noted by van Vlijmen et al. (2020), the availability of an increasing amount of Open Data in principle offers a unique opportunity for data-intensive industries. Private companies rely on the analysis of large amounts of data and research results to carry out their activities and, in particular, develop new products and services.

A notable example is the biopharmaceutical industry, a sector with a longstanding tradition of collaboration with the life science research community and the public sector on data related, for example, to health records, imaging, genome sequencing and others. Drug discovery requires strong modelling capacity since drug properties, for example, efficacy and toxicity in man, cannot be predicted well. Computational technology increases the possibility to use large amounts of data in a way that was not conceivable before. Though not systematic, anecdotal evidence is available about the value of machine-actionable data and textual sources for the biopharma industry, as indicated in the following. For example, more accurate results are possible by linking scientific results with clinical and experimental evidence. As reported in van Vlijmen et al. (2020), a previous study presents a machine learning model which can predict the efficacy of a particular drug for a specific disease with an increase of 12% points of accuracy on earlier state-of-the-art models. This study made use of an existing commercial solution, Euretos

AI Platform[13], which integrates 250 public data sources with proprietary data in a harmonised way. This tool can also assist clients in hypothesis generation with predictive models providing novel target insights.

The health industries have a higher need for speed (Harding, 2017). New medicines for many diseases, in particular neurodegenerative disorders, are not forthcoming, despite patient demands and billions of dollars spent on biomedical research globally. Traditional publishing methods in biomedical sciences are generally slow and sometimes disseminate manuscripts, without the inclusion of primary data, to a privileged audience affiliated to institutions which can afford publication subscription costs. To overcome this barrier to progressive scientific endeavours, many researchers are championing the use of preprints, transparent subject-relevant data repositories, OA journals and open lab notebooks in an effort to more effectively and efficiently communicate their research to a wider audience. Based on his experience on research output on Huntington's disease in real-time through an open lab notebook, Harding indicated the value of OS for new medicines.

Examples beyond the biopharma case are also available. As reported by McManamay and Utz (2014), openly available databases for fisheries science have been on the rise in the recent past. In this paper the authors list almost twenty different regional, national and international databases providing information on, for example, marine and freshwater fish species, biological and physicochemical data for fisheries management, stock assessment results of commercially exploited populations of marine organisms and many others. This wealth of data can support fisheries science; it also inspires the fisheries industry to adopt novel approaches and take informed decisions of prioritising restoration or preventative management actions.

The use of Open Data is associated with costs as well, since manipulating and managing large datasets requires specific individual skills and computational capacities (van Vlijmen et al., 2020). Regular data production and consumption calls for shared principles and standards, especially in view of large collaborations also in private-public mode. While in the past companies relied on proprietary data and internal codes of practices and standards, the increased use of public data sources requires interoperability and imposes time-consuming tasks to data users.

In a generalised absence of wide agreements on metadata standards and ontologies, there is an increasing number of projects and initiatives supporting interoperability and reusability of data. Examples of initiatives promoting FAIR principles in the biopharma are, among many others, the FairPlus project[14] which includes 21 partners from academia and industry or the

[13] www.euretos.com
[14] https://fairplus-project.eu/

Innovative Medicine Initiative[15]. Such initiatives indicate the increasing interest and need of the industry to invest in open data management. In the face of the challenge to promote FAIR data principles, van Vlijmen et al. (2020) indicate a critical need to create a market space for professional organisations tackling the problem of data reliability and interoperability and providing tools and services for high-quality data generation and use in R&D.

5.5. Discussion

Summarising the findings of the literature review on economic impact the most obvious conclusion is that there is very limited, hardly comparable, and generalisable evidence on the influence of OS on the economy. Much of the literature deals with theoretical arguments and not evidence-based numbers on impact obtained from indicators or counterfactual analyses. Our findings are based on case studies and an earlier review of the literature, since data allowing for empirical modelling are rare.

Economic arguments are concerned with the cost of accessing research results. APC variations are dealt with in the context of academic impacts, but the easiness of Open Access is contemplated as potentially having a direct impact on business research. The theoretical reasoning is that cost and time saved may be decisive for technology management in companies. If this is the case, efficiency and effectiveness of research is influenced in the business sector and this in turn leads to speeding up innovation. Unfortunately, this reasoning is mainly theoretical. The only evidence found in a Canadian case study is inconclusive.

In sectoral terms biotech and health have been researched more than other areas. After the open access mandate, patents cite NIH-funded research 12 to 27% more often, this being clear evidence that OS is contributing to innovation. Biopharmaceutical research, typically a data-intensive industry, is found to benefit from Open Data for drug discovery. This indicates that Open Data offers a unique opportunity for data-intensive industries. The slow process of traditional publishing methods is an inhibiting factor that OS can address. Finally, using the fisheries case there is evidence that the wealth of open data not only supports science but gives firms the opportunity to adopt novel approaches and take informed decisions for their strategies.

Openness was further suggested to affect economics via increasing research funding, improving R&I policies and making a difference at regional level, if OS is locally promoted. Funding can be influenced by openness, as suggested by an example of research crowdfunding in Japan. Preferred projects were in biology, art and design, and physics. Supporting individual

[15] https://www.imi.europa.eu

researchers may have an impact on new business creation. Conversely, research in Taiwan claims that economic benefits can occur because open science combined with sophisticated big data analytics can lead to better R&I policies. Finally, one territorial case study was found which bases a multi-actor local collaboration on the assumption that their open data will increase local competitiveness.

Evidence is lacking in Open Source Software, a case where one could imagine significant economic benefits for new firm creation, and in Citizen Science, which can offer data that can enrich the array of evidence needed by companies to calibrate their business decisions.

Indicative areas of new research include

- case studies differentiating the impact of the incremental publishing cost for large versus small businesses;
- the impact on research intensive companies compared to less research-intensive ones;
- the potential of a mutually reinforcing role of open innovation and open science; using non-patent literature in patents in more disciplines that biotech;
- surveys on the sources of radically new products.

Conversely, macro-economic modelling would need to be based on assumptions in sectoral terms that are not yet available.

In terms of impact of research on innovation creation at the firm level some positive evidence is found but much more is needed. We expect evidence from case studies and grey literature to be valuable in this direction.

6. Discussion and future plans

In answer to the question "What evidence exists in the literature regarding the effect of Open Science (OS) on (1) academic, (2) societal, and (3) economic impact of research?", this deliverable details a range of evidence derived from a systematic search of academic databases.

We here reflect on the larger picture revealed by this evidence, contextualise our results by reflecting on knowledge gaps, and detail the path ahead for the PathOS team to complete this study through further snowballing and grey literature search.

6.1. Summary of Findings

Figure 8 below shows the distribution of studies according to their relevance to elements of Open Science, and across types of impact. As can be seen, coverage is highly variable across these domains.

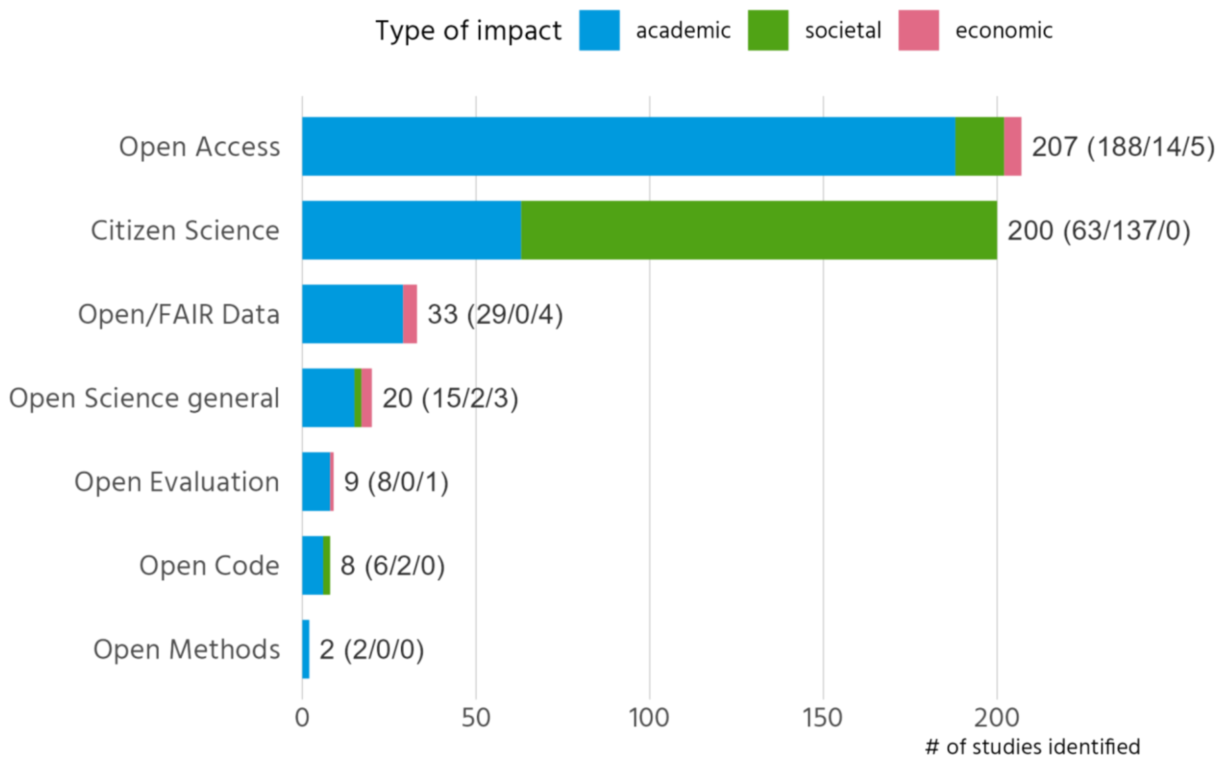

Figure 8 Overview of identified studies across aspects of Open Science and types of impact

On **academic impact**, we identified a substantial body of relevant literature (311 studies). A large proportion of this literature examines Open Access, and especially its effect upon impact as measured via citations. There seems to be an "Open Access citation advantage" where OA literature accrues higher citations than closed access literature, although across the large number of papers addressing this issue there are often methodological issues and failures to control for biases or confounding effects.

More negatively, a growing literature indicates the APC model of OA publishing is fostering effects of exclusion for authors from less--resourced regions and institutions unable to pay fees for publication, and that predatory publishing can threaten the quality of the research literature. FAIR and Open Data are associated with data reuse and a citation advantage for associated papers, but its role in fostering (computational) reproducibility seems less significant than expected as reproduction of results often requires substantial contact with original authors. Open Code and Software seem impactful regarding efficiency gains in software development, and may also increase citations of associated papers. Citizen Science demonstrates impact in the efficiency and scope of data collection, while data quality is sometimes of issue. Open peer review shows neutral to positive effects on review quality.

Regarding **societal impact**, we identified 155 papers as relevant to our inclusion criteria. Of these, the vast majority (137) concerned Citizen Science, across a wide variety of types of societal impact including educational, engagement and empowerment benefits for participants and their communities (supporting the claim that CS democratises science), and the creation of data for use in governmental monitoring and administering of environments and natural resources.

Outside of CS, evidence of the societal impact of OS is limited. Just fourteen papers demonstrated societal impacts of OA, including public engagement with scientific literature, use in policy-making, and health-related outcomes. However, we note that the evidence of public engagement with scientific literature is measured with Altmetrics, and this measure does not include any demographic evidence of social media users, therefore truly public engagement, as opposed to academic engagement, cannot actually be verified. Beyond this, our search revealed very little evidence of societal impact. Especially relevant is a paucity of evidence regarding the policy impact of OS (a recurrent claim in OS advocacy), and a complete lack of evidence of the societal impact of Open/FAIR Data. We note, however, that our exclusion of Open Government Data and how we defined this may have overly narrowed the scope of our research (excluding research from NASA, for example, and other government-funded scientific agencies around the world). In the next phase of this research we will review the inclusion/exclusion criteria for Open/FAIR data and likely reassess the literature gathered so far and apply new, wider criteria, going forward.

Finally, regarding **economic impact**, we identified only 13 papers as relevant. The number is so limited because most papers reviewed were excluded, because they rely on a theoretical rationale, explaining why the academic and societal impacts eventually can turn into economic benefits. There is, however, seldom quantitative corroboration of this rationale. Evidence was most prevalent from the biomedical and health domains. Some evidence gives positive indications of the potential of OA and Open/FAIR data to power economic activity, but this is still largely without rigorous quantification. To wit, we identified just one study that found an increase in patent citations of OA literature from a specific funder in the USA (Bryan & Ozcan, 2021).

6.2. Evidence gaps

Taken together, these results indicate that OS is developed enough for a major research literature on its effects to have bloomed. Yet the literature we have identified is heavily skewed towards the investigation of specific impacts (e.g., citation impact of OA and educational effects of CS), while in other areas much less exists (e.g., Open Code/Software, Open Methods, economic impact).

Before discussing these gaps, we should first remind the reader that this deliverable reports only on the first stage of the literature review process. In the next stage, leading to PathOS Milestone 1.3, we will complement the database search results synthesised here with further snowballing and web search for grey literature (see next subsection). For many aspects of OS impact, this will add substantially to the picture here. As an illustration, Fell's (2019) rapid review of economic impact identified 21 studies, of which 18 were from grey literature sources.

In addition, the coverage of our results suggests certain blindspots in our search strategy, most prominently a lack of evidence regarding preprints and open methods. At the same time, our strict inclusion criteria which set items/topics, for example, Open Government Data, as out of scope may have meant that initiatives which blurred the lines between that and Open/FAIR data within research were excluded (see reflections on this in Section 5.5). Seeking to address these gaps will be a major priority in the next phase.

Beyond this, however, the relative lack of evidence in some areas may reflect the timelines upon which the various aspects of OS have accrued. OA now has a rich history, with more than 20 years since the Budapest declaration (Chan et al., 2002) which crystalized and accelerated that movement. Meanwhile, the term "Open Science" has only been in common parlance for the last decade or so, and many aspects of OS (open methods, open peer review) have only begun to really gather steam in that time. These newer OS practices might not yet be mature enough or at sufficient scales of uptake for longer-term impacts to yet be discoverable, let alone visible or fully studied, and systematic/national/global methods to be applied. Yet lack of maturity in impact assessment should not preclude studies, which should rather use different approaches to collect evidence, including case studies and other tools. More support for such investigations is to be recommended.

8.3. Causality and impact assessment

Quantifying the impact of OS, and in particular, disentangling effects of OS from other potentially confounding factors, is difficult. This difficulty stems from multiple issues, which can be summarised into conceptual and data-related issues. Conceptually, there are many factors contributing to certain impacts like citations and reproducibility in academia, engagement and awareness in society, or economic activity. In addition, the chain of impact from a given Open Science practice to a certain impact can be long, which "dilutes" the initial effect. Forming a clear understanding of all potential factors and formalising this into an empirical approach able to retrieve causal claims is a complex task, which has not been undertaken by most of the included studies.

This is also connected to a sparsity of data in certain areas, and the conceptual issue of missing counterfactual scenarios: if no data have been shared, it is difficult to assess the impact the non-shared data had and to compare it with shared data, because there are substantial practical limitations to detecting reuse of data that is not openly available. It seems telling that the most thematically coherent cluster of research on academic impact concerns the OA citation advantage, a body of work which relies on fairly standardised scientometric methods and readily available datasets (although there is room for improvement as results may be affected by some failures to control for relevant factors). It seems to us that this clustering of research on some topics is at least suggestive of the "streetlight effect", the observational bias whereby we tend to search where it is easiest to look[16]. This could lead to a skewed understanding or partial solutions because the investigation is limited to the 'easiest' or most visible areas, not necessarily the most pertinent or meaningful ones.

A related issue is contrasting the size of effects with the amount of available literature. Even though a lot of research has been conducted on a certain area, such as the OA citation advantage, this does not imply that the size of this effect is stronger than any other areas under study. It is thus important not to mistake the availability of broad literature on a certain topic for the substantive importance in terms of actual effects.

As we have highlighted throughout the deliverable (see especially section 3.2.1), many studies included in our review do not permit drawing causal claims due to inadequate research designs. A case in point is the literature on the OA citation advantage. A substantial portion of the

[16] The name comes from an old joke about a confused person searching for lost keys under a streetlight, not because that's where the keys were lost, but because that's where the light is.

included studies conducted very basic investigations, such as comparing impact factors between closed and OA journals. This ignores that (gold) OA and closed access journals differ systematically in other regards than their OA status. During the earlier phases of the OA movement (roughly up until the late 2010s), OA journals were often newly founded journals, whereas many of the prestigious journals of a given field kept the closed access model. Comparing their impact factors is thus not only comparing the effect of OA, but also comparing prestige and the broader publishing culture in each field. We thus must assume that findings from such studies are substantially biased.

Another prominent area where some of the reviewed studies did not adequately address causality are studies on data quality in CS. The literature shows concern about whether CS data is of sufficient quality to serve a particular purpose. When considering the evidence more broadly, it emerges that the determining factor is not the involvement of citizens per se, but whether studies have adequate methods and protocols for the task at hand. Citizens, with appropriate protocols and training, deliver data of equal quality to trained scientists. The main issue in such cases is thus a lack of appropriately developed and validated research designs .

8.4. Next steps

In the concluding part of our search, currently under completion, we are searching for further relevant literature not identified via the database search using (1) snowball searching of citation lists and (2) online search for grey literature from bodies likely to have produced relevant reports such as research funders, research-performing organisations, academic publishers, student coalitions, and international bodies.[17] During this process, screening and assessment/selection of studies for relevance will take place during the search. All steps will be recorded, including number of results and how many pages of results were screened per source. Data-charting will subsequently be applied to relevant reports, and the results integrated with the foregoing analysis.

Combined analyses are currently being compiled and will be made available to the community shortly via three preprints (on academic, societal and economic impact) which together present a complete picture of Open Science impact. In addition, the underlying literature will be shared via an open Zotero library for direct reference by the community.[18]

[17] We identify the following sources as particularly relevant in this regard: Google Scholar, OpenAIRE, CORDIS (esp. SWAFS) and particular project websites, Overton, EU Publications Office https://op.europa.eu/en/home, Science Europe, EUA, National Academy of Sciences, OECD https://stip.oecd.org/, JISC, Centre for Open Science (OSF), Open Research Funders Group, UKRI, and UNESCO.

[18] https://pathos-project.eu/os-impact-evidence-library

7. References

Abbasi, Z., Shekofteh, M., Shahbodaghi, A., & Kazemi, E. (2019). Citation indicators' comparison of LIS open access and subscription publications based on Scopus. *Global Knowledge, Memory and Communication*, *68*(4/5), 288–299. https://doi.org/10.1108/GKMC-02-2018-0016

Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S. L., & Sadagopan, S. (2021). Narratives and Counternarratives on Data Sharing in Africa. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 329–341. https://doi.org/10.1145/3442188.3445897

Adler, F. R., Green, A. M., & Şekercioğlu, Ç. H. (2020). Citizen science in ecology: A place for humans in nature. *Annals of the New York Academy of Sciences*, *1469*(1), 52–64. https://doi.org/10.1111/nyas.14340

Alabri, A., & Hunter, J. (2010). Enhancing the Quality and Trust of Citizen Science Data. *2010 IEEE Sixth International Conference on E-Science*, 81–88. https://doi.org/10.1109/eScience.2010.33

Ali, B., & Dahlhaus, P. (2022). The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review. *Agriculture*, *12*(2), 309. https://doi.org/10.3390/agriculture12020309

Allen, B. L. (2018). Strongly Participatory Science and Knowledge Justice in an Environmentally Contested Region. *Science Technology and Human Values*, *43*(6), 947–971. https://doi.org/10.1177/0162243918758380

Alordiah, C. O., Owamah, H. I., Ogbinaka, E. J. A., & Alordiah, M. O. (2021). NIGERIA'S LOW CONTRIBUTION TO RECOGNIZED WORLD RESEARCH LITERATURE: CAUSES AND REMEDIES. *Accountability in Research*, *28*(8), 471–491. https://doi.org/10.1080/08989621.2020.1855984

AlRyalat, S. A., El Khatib, O., Al-qawasmi, O., Alkasrawi, H., al Zu'bi, R., Abu-Halaweh, M., alkanash, Y., & Habash, I. (2020). The impact of the National Heart, Lung, and Blood Institute data: Analyzing published articles that used BioLINCC open access data. *F1000Research*, *9*, 30. https://doi.org/10.12688/f1000research.21884.1

Anagnostou, A., Taylor, S. J. E., Groen, D., Suleimenova, D., Anokye, N., Bruno, R., & Barbera, R. (2019). Building Global Research Capacity in Public Health: The Case of a Science Gateway for Physical Activity Lifelong Modelling and Simulation. *2019 Winter Simulation Conference (WSC)*, 1067–1078. https://doi.org/10.1109/WSC40007.2019.9004845

Araujo, A. C., Vanin, A. A., Nascimento, D. P., Gonzalez, G. Z., & Costa, L. O. P. (2021). What are the variables associated with Altmetric scores? *Systematic Reviews*, *10*(1), 193. https://doi.org/10.1186/s13643-021-01735-0

Arnone, S., Buoni, L., Manieri, A., Mathieu, P.-P., Mazzitelli, G., & Spagnoli, F. (2016). *UNIVERSITY, INDUSTRY AND RESEARCH COOPERATION: THE LAZIO PULSE INITIATIVE*. 3808–3814.

https://doi.org/10.21125/iceri.2016.1902

Arza, V., & Fressoli, M. (2017). Benefits of Open Science: An Analytical Framework Illustrated with Case Study Evidence from Argentina. *Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices*, 153–173. https://doi.org/10.3233/978-1-61499-769-6-153

Asai, S. (2019). Changes in revenue structure of a leading open access journal publisher: The case of BMC. *Scientometrics*, *121*(1), 53–63. https://doi.org/10.1007/s11192-019-03200-1

Asai, S. (2021). Author Choice of Journal Type Based on Income Level of Country. *Journal of Scholarly Publishing*, *53*(1), 24–34. https://doi.org/10.3138/jsp.53.1.03

Ashley, E. A., Pattengill-Semmens, C. V., Orr, J. W., Nichols, J. D., & Gaydos, J. K. (2022). Documenting fishes in an inland sea with citizen scientist diver surveys: Using taxonomic expertise to inform the observation potential of fish species. *Environmental Monitoring and Assessment*, *194*(3), 227. https://doi.org/10.1007/s10661-022-09857-1

Bakker, M., Veldkamp, C. L. S., van den Akker, O. R., van Assen, M. A. L. M., Crompvoets, E., Ong, H. H., & Wicherts, J. M. (2020). Recommendations in pre-registrations and internal review board proposals promote formal power analyses but do not increase sample size. *PLOS ONE*, *15*(7), e0236079. https://doi.org/10.1371/journal.pone.0236079

Beck, S., Brasseur, T.-M., Poetz, M., & Sauermann, H. (2022). Crowdsourcing research questions in science. *Research Policy*, *51*(4), 104491. https://doi.org/10.1016/j.respol.2022.104491

Besançon, L., Peiffer-Smadja, N., Segalas, C., Jiang, H., Masuzzo, P., Smout, C., Billy, E., Deforet, M., & Leyrat, C. (2021). Open science saves lives: Lessons from the COVID-19 pandemic. *BMC Medical Research Methodology*, *21*(1), 117. https://doi.org/10.1186/s12874-021-01304-y

Bezuidenhout, L. M., Leonelli, S., Kelly, A. H., & Rappert, B. (2017). Beyond the digital divide: Towards a situated approach to open data. *Science and Public Policy*, *44*(4), 464–475. https://doi.org/10.1093/scipol/scw036

Bianchi, F., & Squazzoni, F. (2022). Can transparency undermine peer review? A simulation model of scientist behavior under open peer review. *Science and Public Policy*, *49*(5), 791–800. https://doi.org/10.1093/scipol/scac027

Bianchini, C., Cosentino, C., Paci, M., & Baccini, M. (2020). Open Access Physical Therapy Journals: Do Predatory Journals Publish Lower-Quality Randomized Controlled Trials? *Archives of Physical Medicine and Rehabilitation*, *101*(6), 969–977. https://doi.org/10.1016/j.apmr.2019.12.012

Bikard, M., Vakili, K., & Teodoridis, F. (2019). When Collaboration Bridges Institutions: The Impact of University–Industry Collaboration on Academic Productivity. *Organization Science*, *30*(2), 426–445. https://doi.org/10.1287/orsc.2018.1235

Biraghi, C. A., Carrion, D., & Brovelli, M. A. (2022). Citizen Science Impact on Environmental Monitoring towards SDGs Indicators: The CASE of SIMILE Project. *Sustainability*, *14*(13), 8107. https://doi.org/10.3390/su14138107

Bishop, B. W., & Collier, H. (2022). Fitness for use of data: Scientists' heuristics of discovery and reuse behaviour framed by the FAIR Data Principles. *Information Research: An International Electronic Journal*, *27*(3). https://doi.org/10.47989/irpaper942

Björk, B.-C., Kanto-Karvonen, S., & Harviainen, J. T. (2020). How Frequently Are Articles in Predatory Open Access Journals Cited. *Publications*, *8*(2), 17. https://doi.org/10.3390/publications8020017

Björk, B.-C., & Solomon, D. (2015). Article processing charges in OA journals: Relationship between price and quality. *Scientometrics*, *103*(2), 373–385. https://doi.org/10.1007/s11192-015-1556-z

Blasco, A., Endres, M. G., Sergeev, R. A., Jonchhe, A., Macaluso, N. J. M., Narayan, R., Natoli, T., Paik, J. H., Briney, B., Wu, C., Su, A. I., Subramanian, A., & Lakhani, K. R. (2019). Advancing computational biology and bioinformatics research through open innovation competitions. *PLOS ONE*, *14*(9), e0222165. https://doi.org/10.1371/journal.pone.0222165

Bokonda, P. L., Ouazzani-Touhami, K., & Souissi, N. (2019). Open Data Kit: Mobile Data Collection Framework For Developing Countries. *International Journal of Innovative Technology and Exploring Engineering*, *8*(12), 4749–4754. https://doi.org/10.35940/ijitee.L3583.1081219

Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science*, *25*(1), 2–16. https://doi.org/10.1177/0963662515607406

Boudreau, C. (2021). Reuse of open data in Quebec: From economic development to government transparency. *International Review of Administrative Sciences*, *87*(4), 855–869. https://doi.org/10.1177/0020852319884628

Bravo, G., Grimaldo, F., López-Iñesta, E., Mehmani, B., & Squazzoni, F. (2019). The effect of publishing peer review reports on referee behavior in five scholarly journals. *Nature Communications*, *10*(1), 322. https://doi.org/10.1038/s41467-018-08250-2

Bruce, R., Chauvin, A., Trinquart, L., Ravaud, P., & Boutron, I. (2016). Impact of interventions to improve the quality of peer review of biomedical journals: A systematic review and meta-analysis. *BMC Medicine*, *14*(1), 85. https://doi.org/10.1186/s12916-016-0631-5

Bruckermann, T., Greving, H., Schumann, A., Stillfried, M., Börner, K., Kimmig, S. E., Hagen, R., Brandt, M., & Harms, U. (2021). To know about science is to love it? Unraveling cause–effect relationships between knowledge and attitudes toward science in citizen science on urban wildlife ecology. *Journal of Research in Science Teaching*, *58*(8), 1179–1202. https://doi.org/10.1002/tea.21697

Bruns, A., Rimmert, C., & Taubert, N. (2020). Who Pays? Comparing Cost Sharing Models for a Gold Open Access Publication Environment. *Journal of Library Administration*, *60*(8), 853–874. https://doi.org/10.1080/01930826.2020.1820275

Bryan, K. A., & Ozcan, Y. (2021). The Impact of Open Access Mandates on Invention. *The Review of Economics and Statistics*, *103*(5), 954–967. https://doi.org/10.1162/rest_a_00926

Cannon, M., Kelly, A., & Freeman, C. (2022). Implementing an Open & FAIR data sharing policy—A case study in the earth and environmental sciences. *Learned Publishing*, *35*(1), 56–66. https://doi.org/10.1002/leap.1442

Carroll, S. R., Rodriguez-Lonebear, D., & Martinez, A. (2019). Indigenous Data Governance: Strategies from United States Native Nations. *Data Science Journal*, *18*(1), 31. https://doi.org/10.5334/dsj-2019-031

Chan, L., Cuplinskas, D., Eisen, M., Friend, F., Genova, Y., Guedon, J.-C., Hagemann, M., Harnad, S., Johnson, R., Kupryte, R., La Manna, M., Rev, I., Segbert, M., de Souza, S., Suber, P., & Velterop, J. (2002, February 14). *Budapest Open Access Initiative*. Budapest Open Access Initiative. https://web.archive.org/web/20171213093708/http://www.budapestopenaccessinitiative.org/read

Chen, S.-C., Chen, Y.-C., & Chen, W.-L. (2017). Data Mining Techniques vs. Policy Development: Evaluating Advanced Applied Technological Policies and Emerging Communication Technology. *2017 IEEE International Symposium on Multimedia (ISM)*, 469–474. https://doi.org/10.1109/ISM.2017.93

Chen, X., Bharti, N., & Marsteller, M. R. (2021). Use of Bibliometrics Data to Understand the Citation Advantages of Different Open Access Categories in Covid-19 Related Studies. *Proceedings of the Association for Information Science and Technology*, *58*(1), 410–414. https://doi.org/10.1002/pra2.469

Chiaravalloti, R. M., Skarlatidou, A., Hoyte, S., Badia, M. M., Haklay, M., & Lewis, J. (2022). Extreme citizen science: Lessons learned from initiatives around the globe. *Conservation Science and Practice*, *4*(2). https://doi.org/10.1111/csp2.577

Cho, H., Low, R. D., Fischer, H. A., & Storksdieck, M. (2021). The STEM Enhancement in Earth Science "Mosquito Mappers" Virtual Internship: Outcomes of Place-Based Engagement with Citizen Science. *Frontiers in Environmental Science*, *9*, 682669. https://doi.org/10.3389/fenvs.2021.682669

Cho, J. (2021). Altmetrics of Highly Cited Research Papers in Social Science. *Serials Review*, *47*(1), 17–27. https://doi.org/10.1080/00987913.2021.1882652

Christoffel, R. (2020). THIRTEEN. How Participation in Citizen Science Projects Impacts Individuals. In *THIRTEEN. How Participation in Citizen Science Projects Impacts Individuals* (pp. 185–188). University of California Press. https://doi.org/10.1525/9780520960473-017

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2021). The open access advantage for studies of human electrophysiology: Impact on citations and Altmetrics. *International Journal of Psychophysiology*, *164*, 103–111. https://doi.org/10.1016/j.ijpsycho.2021.03.006

Clements, J. C., Daigle, R. M., & Froehlich, H. E. (2018). Predator in the Pool? A Quantitative Evaluation of Non-indexed Open Access Journals in Aquaculture Research. *Frontiers in Marine Science*, *5*, 106. https://doi.org/10.3389/fmars.2018.00106

Coetzee, S., Ivánová, I., Mitasova, H., & Brovelli, M. (2020). Open Geospatial Software and Data:

A Review of the Current State and A Perspective into the Future. *ISPRS International Journal of Geo-Information*, *9*(2), 90. https://doi.org/10.3390/ijgi9020090

Coro, G. (2020). Open science and artificial intelligence supporting blue growth. *Environmental Engineering and Management Journal*, *19*(10), 1719–1729. Scopus.

Costa, I. A., Morais, C., & Monteiro, M. J. (2022). Citizen science through schools: The importance of interpersonal relationships. *Cuadernos.Info*, *52*, Article 52. https://doi.org/10.7764/cdi.52.39405

Craig, I., Plume, A., Mcveigh, M., Pringle, J., & Amin, M. (2007). Do open access articles have greater citation impact?A critical review of the literature. *Journal of Informetrics*, *1*(3), 239–248. https://doi.org/10.1016/j.joi.2007.04.001

Crosby, C. J., Arrowsmith, J. R., & Nandigam, V. (2020). Zero to a trillion: Advancing Earth surface process studies with open access to high-resolution topography. In *Developments in Earth Surface Processes* (Vol. 23, pp. 317–338). Elsevier. https://doi.org/10.1016/B978-0-444-64177-9.00011-4

Cummings, J. A., Zagrodney, J. M., & Day, T. E. (2015). Impact of Open Data Policies on Consent to Participate in Human Subjects Research: Discrepancies between Participant Action and Reported Concerns. *PLOS ONE*, *10*(5), e0125208. https://doi.org/10.1371/journal.pone.0125208

De Filippo, D., & Mañana-Rodríguez, J. (2020). Open access initiatives in European universities: Analysis of their implementation and the visibility of publications in the YERUN network. *Scientometrics*, *125*(3), 2667–2694. https://doi.org/10.1007/s11192-020-03705-0

Deguines, N., Princé, K., Prévot, A.-C., & Fontaine, B. (2020). Assessing the emergence of pro-biodiversity practices in citizen scientists of a backyard butterfly survey. *Science of The Total Environment*, *716*, 136842. https://doi.org/10.1016/j.scitotenv.2020.136842

Dehdarirad, T., Didegah, F., & Didegah, A. (2019). Social media visibility of open access versus non-open access articles: A case study of Life Sciences & Biomedicine. *17TH INTERNATIONAL CONFERENCE ON SCIENTOMETRICS & INFORMETRICS (ISSI2019), VOL I*, 762–769. https://research.chalmers.se/en/publication/515327

Dekker, R., Karasz, I., & Stoy, L. (2023). PathOS - D1.1 Open Science Intervention Logic (Version v2). Zenodo. https://doi.org/10.5281/zenodo.10418573

Dhillon, C. M. (2017). Using citizen science in environmental justice: Participation and decision-making in a Southern California waste facility siting conflict. *Local Environment*, *22*(12), 1479–1496. https://doi.org/10.1080/13549839.2017.1360263

Dickinson, J. L., Zuckerberg, B., & Bonter, D. N. (2010). Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics*, *41*(1), 149–172. https://doi.org/10.1146/annurev-ecolsys-102209-144636

Diprose, G., Greenaway, A., & Moorhouse, B. (2022). Making Visible More Diverse Nature Futures through Citizen Science. *Citizen Science: Theory and Practice*, *7*(1), 6. https://doi.org/10.5334/cstp.442

Dorta-González, P., González-Betancor, S. M., & Dorta-González, M. I. (2017). Reconsidering the gold open access citation advantage postulate in a multidisciplinary context: An analysis of the subject categories in the Web of Science database 2009–2014. *Scientometrics*, *112*(2), 877–901. https://doi.org/10.1007/s11192-017-2422-y

Dorta-González, P., & Santana-Jiménez, Y. (2018). Prevalence and citation advantage of gold open access in the subject areas of the Scopus database. *Research Evaluation*, *27*(1), 1–15. https://doi.org/10.1093/reseval/rvx035

Ebersole, C. R., Mathur, M. B., Baranski, E., Bart-Plange, D.-J., Buttrick, N. R., Chartier, C. R., Corker, K. S., Corley, M., Hartshorne, J. K., IJzerman, H., Lazarević, L. B., Rabagliati, H., Ropovik, I., Aczel, B., Aeschbach, L. F., Andrighetto, L., Arnal, J. D., Arrow, H., Babincak, P., … Nosek, B. A. (2020). Many Labs 5: Testing Pre-Data-Collection Peer Review as an Intervention to Increase Replicability. *Advances in Methods and Practices in Psychological Science*, *3*(3), 309–331. https://doi.org/10.1177/2515245920958687

Eger, T., Mertens, A., & Scheufen, M. (2021). Publication cultures and the citation impact of open access. *Managerial and Decision Economics*, *42*(8), 1980–1998. https://doi.org/10.1002/mde.3429

Ekman, K. (2019). INFORMAL LEARNING IN AN ONLINE CITIZEN SCIENCE COMMUNITY. *EDULEARN19 Proceedings*, 5144–5151. https://doi.org/10.21125/edulearn.2019.1272

English, P. B., Richardson, M. J., & Garzón-Galvis, C. (2018). From Crowdsourcing to Extreme Citizen Science: Participatory Research for Environmental Health. *Annual Review of Public Health*, *39*(1), 335–350. https://doi.org/10.1146/annurev-publhealth-040617-013702

Erfanmanesh, M., & Teixeira da Silva, J. A. (2019). Is the soundness-only quality control policy of open access mega journals linked to a higher rate of published errors? *Scientometrics*, *120*(2), 917–923. https://doi.org/10.1007/s11192-019-03153-5

European Commission. (2018, July 26). *Horizon Europe programme analysis*. Research and Innovation, European Commission. https://research-and-innovation.ec.europa.eu/strategy/support-policy-making/shaping-eu-research-and-innovation-policy/evaluation-impact-assessment-and-monitoring/horizon-europe-programme-analysis_en

Evans, C., Abrams, E., Reitsma, R., Roux, K., Salmonsen, L., & Marra, P. P. (2005). The Neighborhood Nestwatch Program: Participant Outcomes of a Citizen-Science Ecological Research Project. *Conservation Biology*, *19*(3), 589–594. https://doi.org/10.1111/j.1523-1739.2005.00s01.x

Eysenbach, G. (2010). 10 Years Experience with Pioneering Open Access Publishing in Health Informatics: The Journal of Medical Internet Research (JMIR). *MEDINFO 2010*, 1329–1333. https://doi.org/10.3233/978-1-60750-588-4-1329

Ezema, I. J. (2021). Gold Route Open Access Journals in Engineering and Technology: Analysis of Research Impact and Implications for Scholarly Communication. *Issues in Science and Technology Librarianship*, *96*. https://doi.org/10.29173/istl43

Felizardo, K., Flach, C., Pereira, R., & Oliveirajr, E. (2022). Surveying the Audience Effect in Open Peer Review of a Software Engineering Workshop. *Proceedings of the XXXVI Brazilian Symposium on Software Engineering*, 446–451. https://doi.org/10.1145/3555228.3555271

Fell, M. J. (2019). The Economic Impacts of Open Science: A Rapid Evidence Assessment. *Publications*, *7*(3), 46. https://doi.org/10.3390/publications7030046

Fraisl, D., Campbell, J., See, L., Wehn, U., Wardlaw, J., Gold, M., Moorthy, I., Arias, R., Piera, J., Oliver, J. L., Masó, J., Penker, M., & Fritz, S. (2020). Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*, *15*(6), 1735–1751. https://doi.org/10.1007/s11625-020-00833-7

Freihardt, J. (2020). Can Citizen Science using social media inform sanitation planning? *Journal of Environmental Management*, *259*, 110053. https://doi.org/10.1016/j.jenvman.2019.110053

Frigerio, D., Puehringer-Sturmayr, V., Neuböck-Hubinger, B., Gegendorfer, G., Kotrschal, K., & Hirschenhauser, K. (2019). Monitoring public awareness about the endangered northern bald ibis: A case study involving primary school children as citizen scientists. *PeerJ*, *7*, e7569. https://doi.org/10.7717/peerj.7569

Fulton, S., López-Sagástegui, C., Weaver, A. H., Fitzmaurice-Cahluni, F., Galindo, C., Fernández-Rivera Melo, F., Yee, S., Ojeda-Villegas, M. B., Fuentes, D. A., & Torres-Bahena, E. (2019). Untapped Potential of Citizen Science in Mexican Small-Scale Fisheries. *Frontiers in Marine Science*, *6*, 517. https://doi.org/10.3389/fmars.2019.00517

Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *PLoS ONE*, *5*(10), e13636. https://doi.org/10.1371/journal.pone.0013636

Gaulé, P., & Maystre, N. (2011). Getting cited: Does open access help? *Research Policy*, *40*(10), 1332–1338. https://doi.org/10.1016/j.respol.2011.05.025

Gotor, E., Pagnani, T., Paliwal, A., Scafetti, F., van Etten, J., & Caracciolo, F. (2021). Smallholder Farmer Engagement in Citizen Science for Varietal Diversification Enhances Adaptive Capacity and Productivity in Bihar, India. *Frontiers in Sustainable Food Systems*, *5*, 726725. https://doi.org/10.3389/fsufs.2021.726725

Gray, R. J. (2020). Sorry, we're open: Golden open-access and inequality in non-human biological sciences. *Scientometrics*, *124*(2), 1663–1675. https://doi.org/10.1007/s11192-020-03540-3

Groulx, M., Brisbois, M. C., Lemieux, C. J., Winegardner, A., & Fishback, L. (2017). A Role for Nature-Based Citizen Science in Promoting Individual and Collective Climate Change Action? A Systematic Review of Learning Outcomes. *Science Communication*, *39*(1), 45–76. https://doi.org/10.1177/1075547016688324

Hall, L. M., & Hendricks, A. E. (2020). High-throughput analysis suggests differences in journal false discovery rate by subject area and impact factor but not open access status. *BMC Bioinformatics*, *21*(1), 564. https://doi.org/10.1186/s12859-020-03817-7

Harding, R. J. (2017). Open Science and Accelerating Discovery in Rare and Neglected Diseases.

Proceedings of the 21st International Conference on Electronic Publishing, 1–5.

Hardisty, D. J., & Haaga, D. A. F. (2008). Diffusion of treatment research: Does open access matter? *Journal of Clinical Psychology*, *64*(7), 821–839. https://doi.org/10.1002/jclp.20492

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., Henry Tessler, M., Lenne, R. L., Altman, S., Long, B., & Frank, M. C. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal *Cognition*. *Royal Society Open Science*, *5*(8), 180448. https://doi.org/10.1098/rsos.180448

Harrison, P. W., Lopez, R., Rahman, N., Allen, S. G., Aslam, R., Buso, N., Cummins, C., Fathy, Y., Felix, E., Glont, M., Jayathilaka, S., Kadam, S., Kumar, M., Lauer, K. B., Malhotra, G., Mosaku, A., Edbali, O., Park, Y. M., Parton, A., … Apweiler, R. (2021). The COVID-19 Data Portal: Accelerating SARS-CoV-2 and COVID-19 research through rapid open access data sharing. *Nucleic Acids Research*, *49*(W1), W619–W623. https://doi.org/10.1093/nar/gkab417

Harvey, G. K. A., Nelson, T. A., Paquet, P. C., Ferster, C. J., & Fox, C. H. (2018). Comparing citizen science reports and systematic surveys of marine mammal distributions and densities. *Biological Conservation*, *226*, 92–100. https://doi.org/10.1016/j.biocon.2018.07.024

Haywood, B. K., Parrish, J. K., & Dolliver, J. (2016). Place-based and data-rich citizen science as a precursor for conservation action: Citizen Science and Conservation Action. *Conservation Biology*, *30*(3), 476–486. https://doi.org/10.1111/cobi.12702

Heumüller, R., Nielebock, S., Krüger, J., & Ortmeier, F. (2020). Publish or perish, but do not forget your software artifacts. *Empirical Software Engineering*, *25*(6), 4585–4616. https://doi.org/10.1007/s10664-020-09851-6

Hodgkinson, I. R., Mousavi, S., & Hughes, P. (2022). New development: Citizen science—discovering (new) solutions to wicked problems. *Public Money & Management*, *42*(2), 133–136. https://doi.org/10.1080/09540962.2021.1967630

Hofmann, B. (2022). Open Science Knowledge Production: Addressing Epistemological Challenges and Ethical Implications. *Publications*, *10*(3), 24. https://doi.org/10.3390/publications10030024

Holt, J., Walker, A., & Jones, P. (2021). Introducing a data availability policy for journals at IOP Publishing: Measuring the impact on authors and editorial teams. *Learned Publishing*, *34*(4), 478–486. https://doi.org/10.1002/leap.1386

Hoover, E. (2016). 'We're not going to be guinea pigs;' Citizen Science and Environmental Health in a Native American Community. *Journal of Science Communication*, *15*(01), A05. https://doi.org/10.22323/2.15010205

Houghton, J. (2009). Exploring the costs and benefits of alternative publishing models. *Rethinking Electronic Publishing: Innovation in Communication Paradigms and Technologie*, 207–238.

Ikkatai, Y., & Ono, E. (2018). Exploring Characteristics of Academic Crowdfunding in Japan. *2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI)*, 458–461.

https://doi.org/10.1109/IIAI-AAI.2018.00098

Jerčić Martinić-Cezar, I., & Marušić, A. (2019). Completeness of reporting in abstracts of randomized controlled trials in subscription and open access journals: Cross-sectional study. *Trials*, *20*(1), 669. https://doi.org/10.1186/s13063-019-3781-x

Johnson, M. F., Hannah, C., Acton, L., Popovici, R., Karanth, K. K., & Weinthal, E. (2014). Network environmentalism: Citizen scientists as agents for environmental advocacy. *Global Environmental Change*, *29*, 235–245. https://doi.org/10.1016/j.gloenvcha.2014.10.006

Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R., & Ehrenfeld, J. G. (2011). Knowledge Gain and Behavioral Change in Citizen-Science Programs: Citizen-Scientist Knowledge Gain. *Conservation Biology*, *25*(6), 1148–1154. https://doi.org/10.1111/j.1523-1739.2011.01745.x

Kelly, R., Fleming, A., Pecl, G. T., von Gönner, J., & Bonn, A. (2020). Citizen science and marine conservation: A global review. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1814), 20190461. https://doi.org/10.1098/rstb.2019.0461

Kerr, S. (2022). Citizen science and deep mapping for climate communication: A report on CHICC. *Journal of Community Archaeology & Heritage*, 1–8. https://doi.org/10.1080/20518196.2022.2051139

Khan, N., Thelwall, M., & Kousha, K. (2021). Measuring the impact of biodiversity datasets: Data reuse, citations and altmetrics. *Scientometrics*, *126*(4), 3621–3639. https://doi.org/10.1007/s11192-021-03890-6

King, A., King, D., Banchoff, A., Solomonov, S., Ben Natan, O., Hua, J., Gardiner, P., Goldman Rosas, L., Rodriguez Espinosa, P., Winter, S., Sheats, J., Salvo, D., Aguilar-Farias, N., Stathi, A., Akira Hino, A., Porter, M., & On behalf of the Our Voice Global Citizen Science Research Network. (2020). Employing Participatory Citizen Science Methods to Promote Age-Friendly Environments Worldwide. *International Journal of Environmental Research and Public Health*, *17*(5), 1541. https://doi.org/10.3390/ijerph17051541

King, A., Odunitan-Wayas, F., Chaudhury, M., Rubio, M., Baiocchi, M., Kolbe-Alexander, T., Montes, F., Banchoff, A., Sarmiento, O., Bälter, K., Hinckson, E., Chastin, S., Lambert, E., González, S., Guerra, A., Gelius, P., Zha, C., Sarabu, C., Kakar, P., … on behalf of the Our Voice Global Citizen Science Research Network. (2021). Community-Based Approaches to Reducing Health Inequities and Fostering Environmental Justice through Global Youth-Engaged Citizen Science. *International Journal of Environmental Research and Public Health*, *18*(3), 892. https://doi.org/10.3390/ijerph18030892

Kobayashi, S., Falcón, L., Fraser, H., Braa, J., Amarakoon, P., Marcelo, A., & Paton, C. (2021). Using Open Source, Open Data, and Civic Technology to Address the COVID-19 Pandemic and Infodemic. *Yearbook of Medical Informatics*, *30*(01), 038–043. https://doi.org/10.1055/s-0041-1726488

Koomen, M. H., Hedenstrom, M. N., Moran, M., & Oberhauser, K. S. (2019). I didn't know what real science was or what it could be: Citizen science and interest in stem education and careers. In *Enhancing STEM Motivation through Citizen Science Programs* (pp. 243–269).

Scopus.

Kowalczuk, M. K., Dudbridge, F., Nanda, S., Harriman, S. L., Patel, J., & Moylan, E. C. (2015). Retrospective analysis of the quality of reports by author-suggested and non-author-suggested reviewers in journals operating on open or single-blind peer review models. *BMJ Open*, *5*(9), e008707. https://doi.org/10.1136/bmjopen-2015-008707

Kwon, S., & Motohashi, K. (2021). Incentive or disincentive for research data disclosure? A large-scale empirical analysis and implications for open science policy. *International Journal of Information Management*, *60*, 102371. https://doi.org/10.1016/j.ijinfomgt.2021.102371

Langham-Putrow, A., Bakker, C., & Riegelman, A. (2021). Is the open access citation advantage real? A systematic review of the citation of open access and subscription-based articles. *PLOS ONE*, *16*(6), e0253129. https://doi.org/10.1371/journal.pone.0253129

Lansingh, V. C., & Carter, M. J. (2009). Does Open Access in Ophthalmology Affect How Articles Are Subsequently Cited in Research? *Ophthalmology*, *116*(8), 1425–1431. https://doi.org/10.1016/j.ophtha.2008.12.052

Lasky, M., Parsons, A., Schuttler, S., Mash, A., Larson, L., Norton, B., Pease, B., Boone, H., Gatens, L., & Kays, R. (2021). Candid Critters: Challenges and Solutions in a Large-Scale Citizen Science Camera Trap Project. *Citizen Science: Theory and Practice*, *6*(1), 4. https://doi.org/10.5334/cstp.343

Lawson, B., Petrovan, S. O., & Cunningham, A. A. (2015). Citizen Science and Wildlife Disease Surveillance. *EcoHealth*, *12*(4), 693–702. https://doi.org/10.1007/s10393-015-1054-z

Leitner, F., Bielza, C., Hill, S. L., & Larrañaga, P. (2016). Data Publications Correlate with Citation Impact. *Frontiers in Neuroscience*, *10*. https://doi.org/10.3389/fnins.2016.00419

Leocadio, J. N., Ghilardi-Lopes, N. P., Koffler, S., Barbiéri, C., Francoy, T. M., Albertini, B., & Saraiva, A. M. (2021). Data Reliability in a Citizen Science Protocol for Monitoring Stingless Bees Flight Activity. *Insects*, *12*(9), 766. https://doi.org/10.3390/insects12090766

Lüsse, M., Brockhage, F., Beeken, M., & Pietzner, V. (2022). Citizen science and its potential for science education. *International Journal of Science Education*, *44*(7), 1120–1142. https://doi.org/10.1080/09500693.2022.2067365

MacPhail, V. J., & Colla, S. R. (2020). Power of the people: A review of citizen science programs for conservation. *Biological Conservation*, *249*, 108739. https://doi.org/10.1016/j.biocon.2020.108739

Maddi, A., & Sapinho, D. (2021). Article Processing Charges based publications: To which extent the price explains scientific impact? *18th International Conference on Scientometrics & Informetrics*. ISSI 2021, Leuven.

Mahajan, S., Chung, M.-K., Martinez, J., Olaya, Y., Helbing, D., & Chen, L.-J. (2022). Translating citizen-generated air quality data into evidence for shaping policy. *Humanities and Social Sciences Communications*, *9*(1), 122. https://doi.org/10.1057/s41599-022-01135-2

Mahajan, S., Luo, C.-H., Wu, D.-Y., & Chen, L.-J. (2021). From Do-It-Yourself (DIY) to Do-It-Together (DIT): Reflections on designing a citizen-driven air quality monitoring framework in Taiwan. *Sustainable Cities and Society*, *66*, 102628. https://doi.org/10.1016/j.scs.2020.102628

Manda, A. K., Etheridge, J. R., Grace-McCaskey, C., Allen, T. R., & Howard, R. (2021). Are Groundwater Level Data Collected by Citizen Scientists Trustworthy? A Cautionary Tale. *Groundwater*, *59*(1), 123–130. https://doi.org/10.1111/gwat.13019

Marchante, E., & Marchante, H. (2016). Engaging Society to Fight Invasive Alien Plants in Portugal—One of the Main Threats to Biodiversity. In P. Castro, U. M. Azeiteiro, P. Bacelar-Nicolau, W. Leal Filho, & A. M. Azul (Eds.), *Biodiversity and Education for Sustainable Development* (pp. 107–122). Springer International Publishing. https://doi.org/10.1007/978-3-319-32318-3_8

Markiewicz, C. J., Gorgolewski, K. J., Feingold, F., Blair, R., Halchenko, Y. O., Miller, E., Hardcastle, N., Wexler, J., Esteban, O., Goncavles, M., Jwa, A., & Poldrack, R. (2021). The OpenNeuro resource for sharing of neuroscience data. *ELife*, *10*, e71774. https://doi.org/10.7554/eLife.71774

Marshall, Z., Brunger, F., Welch, V., Asghari, S., & Kaposy, C. (2018). Open Availability of Patient Medical Photographs in Google Images Search Results: Cross-Sectional Study of Transgender Research. *Journal of Medical Internet Research*, *20*(2), e70. https://doi.org/10.2196/jmir.8787

Matsui, A., Chen, E., Wang, Y., & Ferrara, E. (2021). The impact of peer review on the contribution potential of scientific papers. *PeerJ*, *9*, e11999. https://doi.org/10.7717/peerj.11999

McCormick, M., Liu, X., Jomier, J., Marion, C., & Ibanez, L. (2014). ITK: Enabling reproducible research and open science. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00013

McGreavy, B., Calhoun, A. J. K., Jansujwicz, J., & Levesque, V. (2016). Citizen science and natural resource governance: Program design for vernal pool policy innovation. *Ecology and Society*, *21*(2), art48. https://doi.org/10.5751/ES-08437-210248

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., & Yarkoni, T. (2016). How open science helps researchers succeed. *ELife*, *5*, e16800. https://doi.org/10.7554/eLife.16800

McManamay, R. A., & Utz, R. M. (2014). Open-Access Databases as Unprecedented Resources and Drivers of Cultural Change in Fisheries Science. *Fisheries*, *39*(9), 417–425. https://doi.org/10.1080/03632415.2014.946128

Meerpohl, J. J., Wolff, R. F., Antes, G., & von Elm, E. (2011). Are pediatric Open Access journals promoting good publication practice? An analysis of author instructions. *BMC Pediatrics*, *11*(1), 27. https://doi.org/10.1186/1471-2431-11-27

Meschini, M., Prati, F., Simoncini, G. A., Airi, V., Caroselli, E., Prada, F., Marchini, C., Machado

Toffolo, M., Branchini, S., Brambilla, V., Covi, C., & Goffredo, S. (2021). Environmental Awareness Gained During a Citizen Science Project in Touristic Resorts Is Maintained After 3 Years Since Participation. *Frontiers in Marine Science*, *8*, 584644. https://doi.org/10.3389/fmars.2021.584644

Metcalfe, A. N., Kennedy, T. A., Mendez, G. A., & Muehlbauer, J. D. (2022). Applied citizen science in freshwater research. *WIREs Water*, *9*(2), e1578. https://doi.org/10.1002/wat2.1578

Moed, H. F. (2007). The effect of "open access" on citation impact: An analysis of ArXiv's condensed matter section. *Journal of the American Society for Information Science and Technology*, *58*(13), 2047–2054. https://doi.org/10.1002/asi.20663

Mueller-Langer, F., & Andreoli-Versbach, P. (2018). Open access to research data: Strategic delay and the ambiguous welfare effects of mandatory data disclosure. *Information Economics and Policy*, *42*, 20–34. https://doi.org/10.1016/j.infoecopol.2017.05.004

Mueller-Langer, F., & Watt, R. (2018). HOW MANY MORE CITES IS A $3,000 OPEN ACCESS FEE BUYING YOU? EMPIRICAL EVIDENCE FROM A NATURAL EXPERIMENT: EFFECT OF OPEN ACCESS ON CITATIONS. *Economic Inquiry*, *56*(2), 931–954. https://doi.org/10.1111/ecin.12545

Naudet, F., Sakarovitch, C., Janiaud, P., Cristea, I., Fanelli, D., Moher, D., & Ioannidis, J. P. A. (2018). Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: Survey of studies published in *The BMJ* and *PLOS Medicine*. *BMJ*, k400. https://doi.org/10.1136/bmj.k400

Nielsen, M. (2013). *Reinventing Discovery: The New Era of Networked Science*. Princeton University Press.

Njue, N., Stenfert Kroese, J., Gräf, J., Jacobs, S. R., Weeser, B., Breuer, L., & Rufino, M. C. (2019). Citizen science in hydrological monitoring and ecosystem services management: State of the art and future prospects. *Science of The Total Environment*, *693*, 133531. https://doi.org/10.1016/j.scitotenv.2019.07.337

Nuijten, M. B., Borghuis, J., Veldkamp, C. L. S., Dominguez-Alvarez, L., van Assen, M. A. L. M., & Wicherts, J. M. (2017). Journal Data Sharing Policies and Statistical Reporting Inconsistencies in Psychology. *Collabra: Psychology*, *3*(1), 31. https://doi.org/10.1525/collabra.102

Nursey-Bray, M., Palmer, R., & Pecl, G. (2018). Spot, log, map: Assessing a marine virtual citizen science program against Reed's best practice for stakeholder participation in environmental management. *Ocean & Coastal Management*, *151*, 1–9. https://doi.org/10.1016/j.ocecoaman.2017.10.031

OECD (2000), Glossary Of Evaluation And Results Based Management Terms, Working Party On Aid Evaluation, 33rd Meeting, Paris, 22-23 November 2000

Onwezen, M. C., Bouwman, E. P., & van Trijp, H. C. M. (2021). Participatory Methods in Food Behaviour Research: A Framework Showing Advantages and Disadvantages of Various Methods. *Foods*, *10*(2), 470. https://doi.org/10.3390/foods10020470

Ostermann-Miyashita, E., Pernat, N., & König, H. J. (2021). Citizen science as a bottom-up approach to address human–wildlife conflicts: From theories and methods to practical implications. *Conservation Science and Practice*, *3*(3). https://doi.org/10.1111/csp2.385

Papin-Ramcharan, J., & Dawe, R. A. (2006). The Other Side of the Coin for Open Access Publishing – A Developing Country View. *Libri*, *56*(1). https://doi.org/10.1515/LIBR.2006.16

Parthenos Project. (2019, February 11). *Benefits of Citizen Science*. PARTHENOS. https://training.parthenos-project.eu/sample-page/citizen-science-in-the-digital-arts-and-humanities/what-is-citizen-science/potentials-and-challenges-of-citizen-science/

Pastorino, R., Milovanovic, S., Stojanovic, J., Efremov, L., Amore, R., & Boccia, S. (2016). Quality Assessment of Studies Published in Open Access and Subscription Journals: Results of a Systematic Evaluation. *PLOS ONE*, *11*(5), e0154217. https://doi.org/10.1371/journal.pone.0154217

Peters, M., Eames, C., & Hamilton, D. (2015). The use and value of citizen science data in New Zealand. *Journal of the Royal Society of New Zealand*, *45*(3), 151–160. https://doi.org/10.1080/03036758.2015.1051549

Peterson, G. M. (2013). Characteristics of retracted open access biomedical literature: A bibliographic analysis. *Journal of the American Society for Information Science and Technology*, *64*(12), 2428–2436. https://doi.org/10.1002/asi.22944

Piwowar, H., Day, R., & Fridsma, D. (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE*, *2*(3), e308. https://doi.org/10.1371/journal.pone.0000308

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

Piwowar, H., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *PeerJ*, *1*, e175. https://doi.org/10.7717/peerj.175

Popa, C. L., Dontu, S. I., Savastru, D., & Carstea, E. M. (2022). Role of Citizen Scientists in Environmental Plastic Litter Research—A Systematic Review. *Sustainability*, *14*(20), 13265. https://doi.org/10.3390/su142013265

Quinlivan, L., Chapman, D. V., & Sullivan, T. (2020). Applying citizen science to monitor for the Sustainable Development Goal Indicator 6.3.2: A review. *Environmental Monitoring and Assessment*, *192*(4), 218. https://doi.org/10.1007/s10661-020-8193-6

Raddick, M. J., Prather, E. E., & Wallace, C. S. (2019). Galaxy zoo: Science content knowledge of citizen scientists. *Public Understanding of Science*, *28*(6), 636–651. https://doi.org/10.1177/0963662519840222

Raffaghelli, J. E., & Manca, S. (2023). Exploring the social activity of open research data on ResearchGate: Implications for the data literacy of researchers. *Online Information Review*, *47*(1), 197–217. https://doi.org/10.1108/OIR-05-2021-0255

Ramirez-Andreotta, M. D., Brusseau, M. L., Artiola, J., Maier, R. M., & Gandolfi, A. J. (2015). Building a co-created citizen science program with gardeners neighboring a superfund site: The Gardenroots case study. *International Public Health Journal*, *7*(1), 13. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4420190/

Ratib, O., & Rosset, A. (2006). Open-source software in medical imaging: Development of OsiriX. *International Journal of Computer Assisted Radiology and Surgery*, *1*(4), 187–196. https://doi.org/10.1007/s11548-006-0056-2

Reichmann, S., & Wieser, B. (2022). Open science at the science–policy interface: Bringing in the evidence? *Health Research Policy and Systems*, *20*(1), 70. https://doi.org/10.1186/s12961-022-00867-6

Reinertsen, I., Collins, D. L., & Drouin, S. (2021). The Essential Role of Open Data and Software for the Future of Ultrasound-Based Neuronavigation. *Frontiers in Oncology*, *10*, 619274. https://doi.org/10.3389/fonc.2020.619274

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/BF01405730

Rodriguez, N. M., Arce, A., Kawaguchi, A., Hua, J., Broderick, B., Winter, S. J., & King, A. C. (2019). Enhancing safe routes to school programs through community-engaged citizen science: Two pilot investigations in lower density areas of Santa Clara County, California, USA. *BMC Public Health*, *19*(1), 256. https://doi.org/10.1186/s12889-019-6563-1

Roger, E., Tegart, P., Dowsett, R., Kinsela, M. A., Harley, M. D., & Ortac, G. (2020). Maximising the potential for citizen science in New South Wales. *Australian Zoologist*, *40*(3), 449–461. https://doi.org/10.7882/AZ.2019.023

Rosas, L. G., Rodriguez Espinosa, P., Montes Jimenez, F., & King, A. C. (2022). The Role of Citizen Science in Promoting Health Equity. *Annual Review of Public Health*, *43*(1), 215–234. https://doi.org/10.1146/annurev-publhealth-090419-102856

Rosman, T., Bosnjak, M., Silber, H., Koßmann, J., & Heycke, T. (2022). Open science and public trust in science: Results from two studies. *Public Understanding of Science*, *31*(8), 1046–1062. https://doi.org/10.1177/09636625221100686

Ross-Hellauer, T. (2017). What is open peer review? A systematic review. *F1000Research*, *6*, 588. https://doi.org/10.12688/f1000research.11369.2

Ross-Hellauer, T., Reichmann, S., Cole, N. L., Fessl, A., Klebel, T., & Pontika, N. (2022). Dynamics of cumulative advantage and threats to equity in open science: A scoping review. *Royal Society Open Science*, *9*(1), 211032. https://doi.org/10.1098/rsos.211032

Rubio, M. A., Triana, C., King, A. C., Rosas, L. G., Banchoff, A. W., Rubiano, O., Chrisinger, B. W., & Sarmiento, O. L. (2021). Engaging citizen scientists to build healthy park environments in Colombia. *Health Promotion International*, *36*(1), 223–234. https://doi.org/10.1093/heapro/daaa031

Ruppen, D., & Brugger, F. (2022). “I will sample until things get better – or until I die.” Potential and limits of citizen science to promote social accountability for environmental pollution. *World Development*, *157*, 105952. https://doi.org/10.1016/j.worlddev.2022.105952

Sandhaus, S., Ramírez-Andreotta, M. D., Kilungo, A., Wolf, A. M., Sandoval, F., & Henriquez, P. (2018). Combating Climate Injustices: An Informal Science and Popular Education Approach to Addressing Environmental Health Disparities. *Pedagogy in Health Promotion*, *4*(4), 260–269. https://doi.org/10.1177/2373379917751476

Sauermann, H., Vohland, K., Antoniou, V., Balázs, B., Göbel, C., Karatzas, K., Mooney, P., Perelló, J., Ponti, M., Samson, R., & Winter, S. (2020). Citizen science and sustainability transitions. *Research Policy*, *49*(5), 103978. https://doi.org/10.1016/j.respol.2020.103978

Schaefer, T., Kieslinger, B., & Fabian, C. M. (2020). Citizen-Based Air Quality Monitoring: The Impact on Individual Citizen Scientists and How to Leverage the Benefits to Affect Whole Regions. *Citizen Science: Theory and Practice*, *5*(1), 6. https://doi.org/10.5334/cstp.245

Schläppy, M.-L., Loder, J., Salmond, J., Lea, A., Dean, A. J., & Roelfsema, C. M. (2017). Making Waves: Marine Citizen Science for Impact. *Frontiers in Marine Science*, *4*, 146. https://doi.org/10.3389/fmars.2017.00146

Schneider, J., Rosman, T., Kelava, A., & Merk, S. (2022). Do Open-Science Badges Increase Trust in Scientists Among Undergraduates, Scientists, and the Public? *Psychological Science*, *33*(9), 1588–1604. https://doi.org/10.1177/09567976221097499

Schönfelder, N. (2020). Article processing charges: Mirroring the citation impact or legacy of the subscription-based model? *Quantitative Science Studies*, *1*(1), 6–27. https://doi.org/10.1162/qss_a_00015

Seamans, G. S. (2018). Grassroots citizen science in urban spontaneous vegetation. *Arboriculture and Urban Forestry*, *44*(2), 116–128. Scopus.

Segev, T., Harvey, A. P., Ajmani, A., Johnson, C., Longfellow, W., Vandiver, K. M., & Hemond, H. (2021). A case study in participatory science with mutual capacity building between university and tribal researchers to investigate drinking water quality in rural Maine. *Environmental Research*, *192*, 110460. https://doi.org/10.1016/j.envres.2020.110460

Seifert, V. A., Wilson, S., Toivonen, S., Clarke, B., & Prunuske, A. (2016). Community Partnership Designed to Promote Lyme Disease Prevention and Engagement in Citizen Science. *Journal of Microbiology & Biology Education*, *17*(1), 63–69. https://doi.org/10.1128/jmbe.v17i1.1014

Shaw, B. J. (2017). Citizen Science – Harnessing the Expertise of Farmers to Monitor Biodiversity in Austrian Meadows. In C. Bieling & T. Plieninger (Eds.), *The Science and Practice of Landscape Stewardship* (1st ed., pp. 99–100). Cambridge University Press. https://doi.org/10.1017/9781316499016.010

Shinbrot, X. A., Jones, K. W., Solomon, J., & Escobedo, M. R. (2022). Examining the influence of citizen science participation on individual volunteers in the global South: A case study of hydrologic monitors in Veracruz, Mexico. *The Journal of Environmental Education*, *53*(1), 6–21. https://doi.org/10.1080/00958964.2021.1984193

Siler, K., Haustein, S., Smith, E., Larivière, V., & Alperin, J. P. (2018). Authorial and institutional stratification in open access publishing: The case of global health research. *PeerJ*, *6*, e4269. https://doi.org/10.7717/peerj.4269

Smith, A. C., Merz, L., Borden, J. B., Gulick, C. K., Kshirsagar, A. R., & Bruna, E. M. (2021). Assessing the effect of article processing charges on the geographic diversity of authors using Elsevier's "Mirror Journal" system. *Quantitative Science Studies*, *2*(4), 1123–1143. https://doi.org/10.1162/qss_a_00157

Soroye, P., Edwards, B. P. M., Buxton, R. T., Ethier, J. P., Frempong-Manso, A., Keefe, H. E., Berberi, A., Roach-Krajewski, M., Binley, A. D., Vincent, J. G., Lin, H.-Y., Cooke, S. J., & Bennett, J. R. (2022). The risks and rewards of community science for threatened species monitoring. *Conservation Science and Practice*, *4*(9), e12788. https://doi.org/10.1111/csp2.12788

Sotudeh, H., Arabzadeh, H., & Mirzabeigi, M. (2019). How do self-archiving and Author-pays models associate and contribute to OA citation advantage within hybrid journals. *The Journal of Academic Librarianship*, *45*(4), 377–385. https://doi.org/10.1016/j.acalib.2019.05.004

Spellman, K. V., Cost, D., & Villano, C. P. (2021). Connecting Community and Citizen Science to Stewardship Action Planning Through Scenarios Storytelling. *Frontiers in Ecology and Evolution*, *9*, 695534. https://doi.org/10.3389/fevo.2021.695534

Spezi, V., Wakeling, S., Pinfield, S., Creaser, C., Fry, J., & Willett, P. (2017). Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. *Journal of Documentation*, *73*(2), 263–283. https://doi.org/10.1108/JD-06-2016-0082

Spiegelman, E. (2021). Esteemed Colleagues: A Model of the Effect of Open Data on Selective Reporting of Scientific Results. *Frontiers in Psychology*, *12*, 761168. https://doi.org/10.3389/fpsyg.2021.761168

Stewart, K., Castaneda, D. I., & Azadeh Bolhari, P. E. (2020). *Work in Progress: Citizen scientists' description of an engineer*. *2020-June*. Scopus.

Suber, P. (2012). *Open Access*. The MIT Press. https://library.oapen.org/handle/20.500.12657/26065

Sugimoto, C. R., Larivière, V., Ni, C., & Cronin, B. (2013). Journal acceptance rates: A cross-disciplinary analysis of variability and relationships with journal measures. *Journal of Informetrics*, *7*(4), 897–906. https://doi.org/10.1016/j.joi.2013.08.007

Susanin, A., Boyar, A., Costello, K., Fraiman, A., Misrok, A., Sears, M., & Hildebrandt, T. (2022). Rigor and reproducibility for data analysis and design in the study of eating disorders. *International Journal of Eating Disorders*, *55*(10), 1267–1278. https://doi.org/10.1002/eat.23774

Swan, M. (2012). Crowdsourced Health Research Studies: An Important Emerging Complement to Clinical Trials in the Public Health Research Ecosystem. *Journal of*

Medical Internet Research, *14*(2), e46. https://doi.org/10.2196/jmir.1988

Tai, T. C., & Robinson, J. P. W. (2018). Enhancing Climate Change Research With Open Science. *Frontiers in Environmental Science*, *6*, 115. https://doi.org/10.3389/fenvs.2018.00115

Tanhua, T., Pouliquen, S., Hausman, J., O'Brien, K., Bricher, P., de Bruin, T., Buck, J. J. H., Burger, E. F., Carval, T., Casey, K. S., Diggs, S., Giorgetti, A., Glaves, H., Harscoat, V., Kinkade, D., Muelbert, J. H., Novellino, A., Pfeil, B., Pulsifer, P. L., … Zhao, Z. (2019). Ocean FAIR Data Services. *Frontiers in Marine Science*, *6*, 440. https://doi.org/10.3389/fmars.2019.00440

Taylor, M. (2020). An altmetric attention advantage for open access books in the humanities and social sciences. *Scientometrics*, *125*(3), 2523–2543. https://doi.org/10.1007/s11192-020-03735-8

Teixeira da Silva, J. A., & Kimotho, S. G. (2022). Signs of divisiveness, discrimination and stigmatization caused by Jeffrey Beall's "predatory" open access publishing blacklists and philosophy. *The Journal of Academic Librarianship*, *48*(3), 102418. https://doi.org/10.1016/j.acalib.2021.102418

Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, Chris. H. J. (2016). The academic, economic and societal impacts of Open Access: An evidence-based review. *F1000Research*, *5*, 632. https://doi.org/10.12688/f1000research.8460.1

Teplitskiy, M., Lu, G., & Duede, E. (2017). Amplifying the impact of open access: Wikipedia and the diffusion of science. *Journal of the Association for Information Science and Technology*, *68*(9), 2116–2127. https://doi.org/10.1002/asi.23687

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., … Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Tse, E. G., Klug, D. M., & Todd, M. H. (2020). Open science approaches to COVID-19. *F1000Research*, *9*, 1043. https://doi.org/10.12688/f1000research.26084.1

Tubridy, F., Mölter, A., Lennon, M., & Pilla, F. (2022). Citizen science and environmental justice: Exploring contradictory outcomes through a case study of air quality monitoring in Dublin. *Local Environment*, *27*(5), 622–638. https://doi.org/10.1080/13549839.2022.2068143

van de Gevel, J., van Etten, J., & Deterding, S. (2020). Citizen science breathes new life into participatory agricultural research. A review. *Agronomy for Sustainable Development*, *40*(5), 35. https://doi.org/10.1007/s13593-020-00636-1

Van Haeften, S., Milic, A., Addison-Smith, B., Butcher, C., & Davies, J. M. (2021). Grass Gazers: Using citizen science as a tool to facilitate practical and online science learning for secondary school students during the COVID-19 lockdown. *Ecology and Evolution*, *11*(8), 3488–3500. https://doi.org/10.1002/ece3.6948

van Rooyen, S., Delamothe, T., & Evans, S. J. W. (2010). Effect on peer review of telling reviewers that their signed reviews might be posted on the web: Randomised controlled trial. *BMJ*, *341*(nov16 2), c5729–c5729. https://doi.org/10.1136/bmj.c5729

van Vlijmen, H., Mons, A., Waalkens, A., Franke, W., Baak, A., Ruiter, G., Kirkpatrick, C., da Silva Santos, L. O. B., Meerman, B., Jellema, R., Arts, D., Kersloot, M., Knijnenburg, S., Lusher, S., Verbeeck, R., & Neefs, J.-M. (2020). The Need of Industry to Go FAIR. *Data Intelligence*, *2*(1–2), 276–284. https://doi.org/10.1162/dint_a_00050

van Vlokhoven, H. (2019). The effect of open access on research quality. *Journal of Informetrics*, *13*(2), 751–756. https://doi.org/10.1016/j.joi.2019.04.001

Vasiliades, M. A., Hadjichambis, A. Ch., Paraskeva-Hadjichambi, D., Adamou, A., & Georgiou, Y. (2021). A Systematic Literature Review on the Participation Aspects of Environmental and Nature-Based Citizen Science Initiatives. *Sustainability*, *13*(13), 7457. https://doi.org/10.3390/su13137457

Vilkins, S., & Grant, W. J. (2017). Types of evidence cited in Australian Government publications. *Scientometrics*, *113*(3), 1681–1695. https://doi.org/10.1007/s11192-017-2544-2

von Gönner, J., Herrmann, T.M., Bruckermann, T. et al. (2023). Citizen science's transformative impact on science, citizen empowerment and socio-political processes. Socio Ecol Pract Res 5, 11–33. https://doi.org/10.1007/s42532-022-00136-4

Walker, D., Smigaj, M., & Tani, M. (2021). The benefits and negative impacts of citizen science applications to water as experienced by participants and communities. *WIREs Water*, *8*(1). https://doi.org/10.1002/wat2.1488

Walker, D., Tani, M., Gyawali, N., Chapagain, P., Davids, J., Ghimire, A., Maharjan, M., Parajuli, B., Prajapati, R., Regmi, S., Kumar, R., Shakya, P., & Upadhyay, S. (2021). Citizen Science Water Projects in Nepal: Participant Motivations And the Impacts of Involvement. *Water Alternatives*, *14*, 664–689.

Wallace, D. E., & Bodzin, A. M. (2019). Using mobile learning and authentic practice in citizen science contexts to foster stem interest in high school students. In *Enhancing STEM Motivation through Citizen Science Programs* (pp. 271–298). Scopus.

Wallace, G., Polcyn, S., Brooks, P. P., Mennen, A. C., Zhao, K., Scotti, P. S., Michelmann, S., Li, K., Turk-Browne, N. B., Cohen, J. D., & Norman, K. A. (2022). RT-Cloud: A cloud-based software framework to simplify and standardize real-time fMRI. *NeuroImage*, *257*, 119295. https://doi.org/10.1016/j.neuroimage.2022.119295

Wei, M., & Noroozi Chakoli, A. (2020). Evaluating the relationship between the academic and social impact of open access books based on citation behaviors and social media attention. *Scientometrics*, *125*(3), 2401–2420. https://doi.org/10.1007/s11192-020-03678-0

Wiley, C. (2021). Data Sharing: An Analysis of Medical Faculty Journals and Articles. *Science & Technology Libraries*, *40*(1), 104–115. https://doi.org/10.1080/0194262X.2020.1781740

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B.

(2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.18

Xia, J., & Liu, Y. (2013). Usage Patterns of Open Genomic Data. *College & Research Libraries*, *74*(2), 195–207. https://doi.org/10.5860/crl-324

Yang, J., Xing, D., & Luo, X. (2021). Assessing the performance of a citizen science project for monitoring urban woody plant species diversity in China. *Urban Forestry & Urban Greening*, *59*, 127001. https://doi.org/10.1016/j.ufug.2021.127001

Young, J. S., & Brandes, P. M. (2020). Green and gold open access citation and interdisciplinary advantage: A bibliometric study of two science journals. *The Journal of Academic Librarianship*, *46*(2), 102105. https://doi.org/10.1016/j.acalib.2019.102105

Zettler, E. R., Takada, H., Monteleone, B., Mallos, N., Eriksen, M., & Amaral-Zettler, L. A. (2017). Incorporating citizen science to study plastics in the environment. *Analytical Methods*, *9*(9), 1392–1403. https://doi.org/10.1039/C6AY02716D

Zhang, J., Chen, S., Cheng, C., Liu, Y., & Jennerjahn, T. C. (2023). Citizen science to support coastal research and management: Insights from a seagrass monitoring case study in Hainan, China. *Ocean & Coastal Management*, *231*, 106403. https://doi.org/10.1016/j.ocecoaman.2022.106403

Zhang, L., & Ma, L. (2021). Does open data boost journal impact: Evidence from Chinese economics. *Scientometrics*, *126*(4), 3393–3419. https://doi.org/10.1007/s11192-021-03897-z

Zhang, L., Wei, Y., Huang, Y., & Sivertsen, G. (2021). *The prevalence and impact of different types of open access articles from China and USA*. 18th International Conference of the International Society for Scientometrics and Informetrics, Leuven.

Zong, Q., Xie, Y., & Liang, J. (2020). Does open peer review improve citation count? Evidence from a propensity score matching analysis of PeerJ. *Scientometrics*, *125*(1), 607–623. https://doi.org/10.1007/s11192-020-03545-y