



**International
Science Council**

The global voice for science

Opening the record of science: making scholarly publishing work for science in the digital era

ISC Report
February 2021



Opening the record of science: making scholarly publishing work for science in the digital era

Preface	3
Summary	3
1. Science and publishing	7
1.1 Why science matters	7
1.2 The record of science	8
1.3 Diverse publishing traditions	10
2. Principles for scientific publishing	12
2.1 Principles and their rationales	12
2.2 Responses from the scientific community	14
3. The evolving landscape of scholarly and scientific publishing	15
3.1 The commercialization of scientific publishing	15
3.2 The reader-pays model	19
3.3 The open access movement	19
3.4 The author-pays models	21
3.5 Learned society publishing	23
3.6 Institutionally-based repositories and infrastructures	24
3.7 Preprint repositories	24
3.8 'Public infrastructures' – publicly funded and scholar-led	26
3.9 Books and monographs	28
3.10 'Predatory' publishing	29
4. Publishing the data of science	29
4.1. Open data: scrutiny of the evidence for truth claims	30
4.2 Binary publication	31
4.3 The general case for open data	33
4.4 Enabling digital data sharing and reuse	33
5. Key issues: barriers to open access	35
5.1 Assessments, incentives and metrics	35
5.2 Peer review issues	37
5.3 Copyright and patents	40
5.4 Indexing	42
5.5 Costs and pricing	43
6. A changing world of science and its implications for scientific publishing	45

6.1 Trends in science	45
6.2 Open science.....	47
6.3 A critique from the Global South.....	49
6.4 Other contrary voices	50
7. Exploiting digital potential.....	50
7.1 The digital impact on the research cycle	51
7.2. Linked digital infrastructures for the research cycle	52
7.3 Monetizing the research cycle.....	53
7.4 The emergence of commercial digital platforms for science	55
7.5 Governance of digital infrastructures: perspectives from the science community	56
8. Summary assessment: how well are science and scholarship served?	57
9. Priorities for action	58
9.1 Market structures.....	59
9.2 Digital opportunities.....	60
9.3 Governance.....	61
9.4 Context and action for change	61
References.....	63
Appendix: Report production and review	74
Working Group members	74
International Science Council Staff.....	74
Review Committee	74

Preface

Many concerns have been expressed in recent decades about the extent to which scientific and scholarly publishing systems serve the needs of researchers and the public interest in accessing the results of scientific and scholarly inquiry. Such concerns were echoed in a consultation with the membership of the International Science Council (ISC), leading to its adoption of the ‘future of scientific publishing’ as a priority topic for its 2019–2021 Action Plan.

The present report is the culmination of the first phase of the resultant project. It was prepared as a discussion document in consultation with an international working group. The text was subjected to three phases of review followed by revision: involving an initial expert review group, the ISC membership including three virtual fora, and an expert team generously convened by the U.S. National Academies of Sciences, Engineering and Medicine¹, before being submitted to the ISC Governing Board for its agreement as an ISC Report.

The report is primarily directed towards the scientific community and its institutions, seeking to establish, as far as possible, a shared view of the principles and priorities of the system through which its work is disseminated, and as a precursor for action to promote beneficial change. It has a distinctive logic. It proposes a series of normative principles that should underlie the operation of scientific and scholarly publishing; describes the current publishing landscape and its trajectory of evolution; analyses the extent to which the principles are observed in practice; and identifies problematic issues that need to be addressed in realizing those principles. With a few exceptions it does not make recommendations about how to resolve problematic issues. The report will be used to set the agenda for a subsequent phase of discussion and action involving ISC members and other stakeholders.

Summary

Why science matters

Science is indispensable to the human endeavour as a fundamental part of its intellectual infrastructure. Its distinctive value derives from open scrutiny of concepts based on evidence and tested against reality, logic and the scepticism of peers. The knowledge that has been accumulated since the earliest days of scientific practice is continually refreshed, renewed and re-evaluated by new experiments, new observations and new theoretical insights, and publicly communicated in the published **record of science**. This record exposes the logic and evidence of truth claims to scrutiny, making science accessible to all who would use it through processes of widely and openly accessible publication and with the potential for innovative use in a myriad of educational, social, economic and cultural settings. Publication processes that achieve these ends and are adapted to the needs and priorities of the disciplines of science and interdisciplinary collaboration are essential to the function of science as a global public good.

¹ This does not imply endorsement.

Principles for scientific publishing

As a basis for analysing the extent to which contemporary scientific and scholarly publishing serves the above purposes, a number of fundamental principles are advocated in the belief that they are likely to be durable in the long term. They follow, in abbreviated form:

- I. There should be universal open access to the record of science, both for authors and readers.
- II. Scientific publications should carry open licences that allow reuse and text and data mining.
- III. Rigorous and ongoing peer review is essential to the integrity of the record of science.
- IV. The data/observations underlying a published truth claim should be concurrently published.
- V. The record of science should be maintained to ensure open access by future generations.
- VI. Publication traditions of different disciplines should be respected.
- VII. Systems should adapt to new opportunities rather than embedding inflexible infrastructures.

These principles have received strong support from the international scientific community as represented by the membership of the International Science Council (ISC).

The evolving landscape

As the scientific effort expanded and diversified in the later 20th century, commercial publishers progressively displaced, with some exceptions, the role of learned societies in scientific publishing by entering the market at scale, increasingly coming to dominate it and progressively driving up prices at rates in excess of inflation.

As an open access movement developed in response to the opportunities offered by the digital revolution, major publishers added an open access, author-pays option based on article processing charges (APCs) to their existing subscription (reader-pays) models. Publisher profitability has been largely based on a combination of 'high-impact journals' and large volumes of journals of lesser standing, often bundled together in inflexible 'big deals' made with universities or national research bodies. The importance of scale in determining profitability has inhibited the role of other than a few learned societies in this market, distorting it to the advantage of publishers by the confusion of customer–supplier roles and through the freedom given by universities and research institutes to authors to publish where they wish. Article Processing Charges are unlikely to resolve many of the problems of the current system, and may even entrench commercial control over the publishing market. Problems of affordability, the lag-time from submission to publication and the opportunities offered by the internet have stimulated a much richer spectrum of dissemination modes beyond the traditional journal or book. An increasingly important and timely innovation has been of repositories that make 'preprints' available prior to peer review, increasingly involving learned societies and university repositories as parts of the effort to expand access for both scholars and the public. They increase the rate at which new findings are disseminated, thereby enhancing inter-activity between researchers and providing early evidence on urgent contemporary issues, whilst 'overlay journals' offer a peer review service for preprints. Publicly funded and scholar-led publishing infrastructures have developed, in Latin America in particular, as efficient non-profit repositories that provide holistic open access systems for scholarly communication. At the same time, 'publish or perish' regimes in universities have created a massive global demand for publishing outlets that has spawned so-called 'predatory' journals that offer rapid online publication but with low publishing standards and little, if any, peer review.

Publishing the data of science

Observations and experiments that reveal novel insights into reality are first-class scientific outputs and essential parts of the record of science. They should be credited as such. Moreover, data that underpin a published truth claim must be accessible, and 'FAIR' (Findable–Accessible–Interoperable–Reusable), so that the logic of the evidence–claim connection can be scrutinized and the observation or experiment repeated, as essential parts of the process of scientific self-correction. A process of 'binary publication' is advocated, whereby when the data are too numerous to be contained in the published truth claim, the data should be concurrently 'published' in a trustworthy repository so that there are pathways to access by reviewers and readers. Protocols should be developed whereby such publications are regarded as at least equivalent in value to the traditional article, with journals requiring related evidence and data to be available as a condition of publication, an approach that could be a powerful incentive for open data sharing. There is also a general case for opening access to data that are not used in a published article. Unless the habit and the means are developed of making scientific data openly and routinely available and interoperable, the opportunity will be lost to collate and integrate data from a variety of disciplinary sources to investigate the complexity at the heart of many of the major problems that confront humanity and to which science can make a vital contribution. Management of large data volumes and their fluxes can be an onerous task. An important challenge is to embed efficient data stewardship and FAIR procedures as normal functionalities of the research cycle, as the responsibility and the cost of doing science in the digital age rather than as an optional add-on.

Barriers to open access

There are a number of key issues that impede creation of, access to, and use of the record of science.

Assessments and incentives

The use of bibliometric indices, such as journal impact factors, as proxy metrics for the performance of researchers is a convenient index of assessment but deeply flawed. Most place a relentless focus on individual achievement, thin out research support through a university's interest in high impact metrics, pressurize all to 'tick boxes' and conform, whilst they play an important role in distorting the journal publication market. There is urgent need for reform.

Peer review

Peer review is currently under considerable stress because of the sheer volume of demand, such that an incentive or reward for undertaking the task is needed. The way in which peer review is related to the increasing significance of preprints, particularly at times of crisis when there is a demand for rapid access to work that has not yet been reviewed, is a pressing issue.

Copyright

The transfer of copyright to publishers as a condition of publication is a regressive practice, particularly when it involves privatization of publicly funded research results, and where parts of the record of science are denied, by high paywalls, for use by the very people that have created it.

Indexing

Indexes to published work are important in signposting the existence of scientific knowledge. Many are owned by commercial publishers, and tend to favour their own journals and are reluctant to add new publishers. This particularly disadvantages publishing enterprises outside Europe and North America where all the 'high-impact' publishers are located, researchers from the Global South, who are unable to afford access to such publications, and the journals in those regions.

Costs and prices

The digital revolution has reduced prices in most public and private sectors, but not in large parts of the scientific publishing sector. The high prices charged for access to high-impact journals, either for authors or readers, discriminates on the basis of ability to pay against many readers, researchers and institutions, and particularly those in low- and middle-income countries. Although publishers have been reluctant to divulge their costs, there is evidence that prices for many journals are an order of magnitude higher than necessary costs, even for high-impact journals with high rejection rates. A variety of models that do not create such excessive prices have developed globally.

A changing world of science

There are major trends in science and society that create a vital context for scientific publishing. The value of science to national economies and in confronting global challenges demands more efficient processes of knowledge dissemination. The era of big data permits science to address the complexity inherent in such challenges in unprecedented ways, but requires access to and publication of data as a norm of scientific inquiry. The web has democratized information, creating both opportunities for the dissemination of scientific knowledge and problems in coping with vast networks of misinformation. These trends underline the fundamental need for scientific publishing to develop in ways that facilitate global cooperation; ensure that the richness of diverse global perspectives is drawn on in developing global solutions; create ready access to the record of science and its data to enable deeper understanding of complexity; enable open access to the record of science to citizens and other stakeholders, particularly in areas of contemporary public concern; and ensure that a scientific voice is effective in combating the global ‘infodemic’ of misinformation. The movement for a new era of open science is seen by many as a means of achieving such objectives, powerfully illustrated in the global scientific response to the COVID-19 pandemic, although it is a movement not without critics.

Exploiting digital potential

All disciplines, whether or not data-intensive, operate in a digital world where all the elements of the research process are connected or connectable in ways that permit them to be linked together as parts of a research workstream, with the possibility of digital interoperability across the ‘research cycle’. These linked digital infrastructures also provide information about the research process that can help in managing and evaluating research by researchers, universities and funders. Major commercial publishers are moving to monetize the research cycle by providing evaluation and management tools for institutions and funders, thus giving them the potential to develop a dominant position in the research system and to create science-knowledge platforms analogous to other digital platforms that are currently of concern to antitrust regulators. Should the governance of such systems be in the hands of private companies, or should they be governed from within the scientific community and its institutions to protect the public interest and those that seek to deliver it? A resounding response from the members of ISC that responded to a consultation survey was that it should be the latter. Change in the former direction is rapid, however, and any alternative options need to be undertaken with urgency.

Summary assessment

The above analysis leads to an assessment of the extent to which the current system serves the interests of science as reflected in the principles set out in section 2 and identifies the following needs for reform:

- I. Many business models inhibit access to the record of science by researchers and/or the public, and exclude authors from poorly-funded institutions and low- and middle-income countries.

- II. Copyright transfer to publishers inhibits access to the record of science, reducing its reuse or mining for the knowledge it contains or in response to emergencies.
- III. Peer review needs to adapt to increasing volumes of work, to diverse modes of scientific publishing and to demands for rapid access to emerging knowledge.
- IV. The data/observations underlying a published truth claim should be concurrently published, with a need for normative procedures as prerequisites for publication.
- V. The further development, federation and interoperation of digital libraries, governed by the global public interest, are important priorities for the long-term record of science.
- VI. Reflection is required about norms and priorities for open access for individual scholarly disciplines and how they can best be delivered whilst also facilitating interdisciplinary publication.
- VII. The potentials of the digital revolution for scholarly publishing have not been fully realized, and moves towards monopolistic platforms threaten innovation and the global public good.

Priorities for action

The current system of scientific and scholarly publishing is a ‘mixed economy’ of for-profit and not-for-profit operations, variously involving private sector commercial bodies, publicly funded systems and institutionally based, learned society and independent operations. We expect this mix to be maintained whilst advocating that there should be a shared view of purpose in serving the global public good by adhering to Principles I–VII in section 2. The market needs reform in ways that increase efficiency and avoid monopolistic behaviour through a more rational relationship between its customers (the scientific community) and suppliers. The opportunities offered by the digital revolution must be grasped, which involves questioning some of the assumptions that underlie a system that is still based on norms from the era of print and paper. System governance should primarily lie in the hands of the scientific community and its institutions rather than those of private companies. The ISC will work with its members, national academies, international scientific unions and associations, other regional and national science bodies and publishers to seek tractable solutions to the major problems of scientific and scholarly publishing identified by this report.

1. Science and publishing

1.1 Why science matters

Science is an indispensable part of the human endeavour. It is not a dispensable luxury. It helps us make sense of and navigate the increasingly complex world we live in. We need science for the advancement of our societies, to respond to their needs, to inform our education, improve our policies, spur innovation, address global sustainability, safeguard health and wellbeing, and as a stimulus to curiosity, imagination and wonder.

The value of science² as a distinct form of knowledge is based on open scrutiny of concepts and their evidence, tested against reality, logic and the scrutiny of peers. It is embedded in the Universal Declaration of Human Rights as ‘everyone has the right freely to participate in the cultural life of the

² Throughout this document, the word **science** is used to refer to the systematic organization of knowledge that can be rationally explained and reliably applied. It is inclusive of the natural (including physical, mathematical and life) science and social (including behavioural and economic) science domains, which represent the ISC’s primary focus, as well as the humanities, medical, health, computer and engineering sciences (ISC, 2018). It is recognized that there is no single word or phrase in English (though there are in other languages) that adequately describes this knowledge community. It is hoped that this shorthand will be accepted in the sense intended.

community, to enjoy the arts and to share in scientific advancement and its benefits'. The vision of the International Science Council (ISC) is of science as a global public good³.

1.2 The record of science

Science most effectively serves the public good if the knowledge and understanding that it creates are communicated promptly and comprehensibly into the public sphere. The processes of formal scientific publication are the prime conduits of such communication.

Central to the role of science and its communication is **the record of science**: the record of scientific knowledge and understanding from the earliest days of scientific inquiry to the present. It is continually refreshed, renewed and re-evaluated across the disciplines of science by new experiments, new observations and new theoretical insights. Perennial scrutiny is at the core of the value of science. It can invalidate, but cannot validate; it is the basis of so-called scientific self-correction. New 'truths' are provisional. In the words of Albert Einstein – 'a thousand experiments cannot prove me right, but one experiment can prove me wrong', and of Berthold Brecht – 'the aim of science is not to open the door to infinite wisdom, but to set a limit to infinite error'.

The record of science is a complex mix of novel contributions that withstand critical tests of experiment, observation or logic and that are of general or local significance, and others that are disregarded and are rarely, if ever, remembered or quoted, though sometimes rediscovered as significant insights. Contributions that fail such tests lose currency. Recognition and reuse by peers have determined how the record of science reflects the evolving framework of understanding in a discipline or about a phenomenon.

Scientific publications play an essential role in preserving and disseminating the record of science. They have, or should have, two distinctive attributes. Firstly, those who wish to find what is known about a particular topic or phenomenon need to be able to find it – no small task when about three million scientific papers are published per year. Secondly, work that is presented for publication should, at some point in the process, be 'tested against reality, logic and the scrutiny of peers' to ensure that high standards of soundness and rigour are attained. There is a spectrum of published work, from that which is so technical as to be only accessible to experts in the same field, to that which is comprehensible to any reader.

In the era of paper and print, the processes of scientific publication were relatively clear-cut. Most disciplinary areas had journals devoted to them. For books and journals, respectively, ISBN/ISSN were established as unique identifiers along with a registry of publishers. The functions of research libraries were to hold the record of science in journals, books and monographs, and to provide users with the means of navigating this knowledge space. The well-defined pathways to publication in a journal operating a peer review 'quality' threshold created a more or less formalized 'version of record'.

The digital revolution has massively disrupted this hitherto settled system. It has enhanced the potential for researchers to acquire, store and manipulate unprecedented data volumes, whilst communication and networking tools have revolutionized and diversified the ways in which the record of science is accumulated, located and disseminated. The revolution has created an increasingly

³ The vision of the ISC is to advance science as a global public good. Scientific knowledge, data and expertise must be universally accessible and their benefits universally shared. The practice of science and the opportunities for scientific education and capacity development must be inclusive and equitable (ISC, 2018). To economists, public goods have two essential properties: non-rivalrous consumption – the consumption of one individual does not detract from that of another; and non-excludability – it is difficult if not impossible to exclude an individual from enjoying the good (Stiglitz, 1999).

diverse digital ecosystem of connected or connectable research objects: digital text articles, data, software, experimental protocols, research instruments and digitally represented objects in static or video format, all with persistent unique identifiers (Box 1) that reference one another and that are findable and readable by machines. As machine-readable, metadata-described datasets are becoming more comprehensive, there is increased emphasis on datasets as primary products of research with, in some fields, the text article becoming ancillary to the data. In this context, the role of the library as the physical custodian of the record of science has changed. That record is increasingly located in the 'cloud', with the library functioning as procurer of digital content, manager of digital infrastructures and access, and as guide in navigating a new geography of knowledge. Whilst conventional journals continue to accumulate versions of record, an urgent current priority is to ensure that there is an accessible 'record of versions' that is able to capture the increasing diversity of scientific outputs, and to ensure that versions are subject to appropriate processes of peer review.

The capacity now exists to create and disseminate a much richer spectrum of research outputs than simply the traditional journal or book text. The process of publication has become less sharply defined, and information is being made publicly available in a plethora of forms: digital or audio-visual formats on personal web sites, in social media, as traditional journal articles, in databases and increasingly rarely in printed text. In many disciplines, preprints are commonly made available prior to peer review. Although these new modes make it easier in principle to become a 'publisher', a major current trend (see section 7.4) is the transformation of major publishing houses into digital information providers through the capture of data supplied by authors who publish with them. Publishing houses now increasingly compete in a market with other massive IT companies in providing data of and about science.

Scientific and scholarly publishing provides a pathway for knowledge into the public domain where it can be accessed by interested parties. As the formats of scientific publication become increasingly diverse, unambiguous indexing and directory structures have become ever more important in discovering content, irrespective of its form. Persistent identifiers and their relationships are the glue that holds the record of science together by maintaining coherence between diverse formats (Box 1).

Box 1. Persistent identifiers

Whereas a 'web address', a URL (Uniform Resource Locator) can fail, a persistent identifier (PID) reliably points to a digital entity. An ORCID identifier (Open Researcher and Contributor ID) is a persistent identifier for a person. A DOI (digital object identifier) is a persistent identifier for entities such as journal articles, books and datasets. The ORCID identifier, for example, plays a vital role in discovering the output of individual researchers. Entering a name into a web browser such as Google will produce a seemingly random and potentially confusing selection of documents relating to individuals with that name. Enter the person's ORCID, and the online forms are automatically populated with a precisely targeted academic profile and publication list. [Crossref](#) and [DataCite](#) are the main organizations that assign DOIs for these purposes in scholarly communication, working closely with other PID organizations to build trustworthy connections between identifiers.

Building a trustworthy record requires scrutiny and critique. Processes of peer review have long been developed to ensure, as far as possible, soundness and rigour in scholarly journals, books and monographs. Increasingly diverse modes of ‘publication’ make carrying out peer review a more problematic issue, but one that is particularly important in a world that needs scientific input into public policy and discourse more than ever, but where there is an ‘infodemic’ of misinformation and fake news that are readily broadcast through the World Wide Web and social media. Extending review procedures so that they are applicable to new modes and formats of publishing is an important priority. An ISC project on this issue is currently in progress.

Scientific publications should, in principle, perform a number of vital tasks⁴:

- a) Make the conclusions, and the evidence (the data⁵) on which a scientific truth claim is based, accessible to scrutiny by peer review and post-publication analysis so that method and logic can be validated or invalidated, conclusions scrutinized and any observations or experiments replicated. This process has proved a powerful means of identifying error, and is the basis of so-called ‘self-correction’ of science, the non-negotiable principle that is the bedrock of the public value of science.
- b) Preserve the record of science so that it is accessible to succeeding generations for reassessment and in reuse in further research.
- c) Enable the global community of scientists to perennially keep abreast of the development of knowledge, and thereby build on the work of earlier generations in contributing to a self-renewing, evolving body of knowledge.
- d) Form an essential part of the process whereby scientific knowledge is disseminated into wider society with the potential for innovative use in a myriad of educational, social, economic and cultural settings.
- e) Publications and their citations also enable ‘filtration’ (sorting out what is worth reading) and ‘designation’ (sorting out what is important).

These processes are fundamental to science and its function as a global public good. They form the context for this report.

However, scholarly publishing has suffered a significant distortion in turning what was meant to be a means of communication into a means of assessment. Bibliometric indicators that depend upon countable outputs of standardized format (such as journal publications) are used as proxies for research quality, although research can be meaningfully and effectively communicated in many other ‘non-standard’ forms. Scholars are thereby driven by such incentives to prioritize those publications that are perceived to further their career because of their association with a particular metric, giving ‘publication’ a meaning that is not consistent with its essential role. This is a self-reinforcing process that keeps the scholarly community locked into the measurable rather than the valuable, reducing the flexibility to adapt to new opportunities and thereby inhibiting even beneficial change.

1.3 Diverse publishing traditions

There is considerable diversity in the norms and habits of publication across the spectrum of scholarly disciplines. This diversity needs to be accommodated as publication systems evolve, with some disciplines finding creative new ways of satisfying their needs. Preferred publication modes range

⁴ Essential functions of scientific and scholarly communication have been described as registration, certification, awareness, archiving and reward (Roosendaal and Geurts, 1997).

⁵ We use ‘data’ to refer to digital or text information, images, objects, audio or film resources, all of which can be digitally represented, including all the materials, text or images found in traditional print and paper publications.

across conventional journals, books and monographs, conference proceedings, preprints, occasional reports, professional journals and an increasing variety of less formal modes of communication through blogs, videos and social media.

Journals with prepublication peer review tend to dominate in STEM disciplines and in many areas of the social sciences. The annual revenues generated from English-language STM journals are estimated to have been about \$10 billion in 2017, within a broader STM information publishing market worth some \$25.7 billion (STM, 2018). In mathematics, however, there has been a strong movement towards online preprint publication (see section 3.7), in which papers tend firstly to be made publicly available either through a preprint server or an author's webpage. Ethical standards tend to be rigorously set and applied in the biomedical disciplines (e.g. The Declaration of Helsinki or <https://www.nuffieldbioethics.org>), particularly when human subjects are involved, or when relating to publication of the results of clinical trials, for which there are regulated standards that are rigorously observed and monitored. There have been calls from these disciplines for radical reform (Raff, 2012) to counteract the bias against publication of negative results.

In computer science, although conventional journal publication is common, delivery of papers at conferences and their publication in conference proceedings has a higher status than in other STEM disciplines. They are more selective, with acceptance rates at highly regarded conferences often no more than 10%. Conference proceedings also tend to be more timely, with faster turn-round times for review and publication than journal publications. High-impact journal metrics (see section 3.2) are therefore less important in these disciplines, although some question whether this is understood by assessment bodies.

Publication norms vary greatly across the social sciences, with economics being similar to STM disciplines, monographs being important in anthropology and sociology, and fields such as public health, demography, law and health studies having professional magazines as favoured outlets. Philosophers operate double-blind reviewing, and many support free online publication of papers, but only after they have been accepted for journal publication, although that may depend on the licence that the author has signed with the publisher.

Publication in the humanities is often less time-critical than in other disciplinary areas, and work may take years to write and to publish, such that book-length publications or monographs are important and tend to carry high prestige. Authors may be disinclined to choose digital-only methods for publication (see section 3.9) although long-form publications are often costly to produce and challenging to digitize. Journals are commonly produced by university departments and rarely make a profit. Although impact factors are of limited value, there are widely shared views about the status and hierarchy of different journals.

The ability of a commercial company to recognize the value of new, external information, assimilate it and apply it to commercial ends is critical to its innovative capabilities (Cohen and Levinthal, 1990), with published papers and reports as important channels through which publicly funded research is accessed by private companies (Cohen et al., 2000). The benefits to companies are enhanced when the record of science, including datasets that are an integral part of it, is available for reuse under open licence. Although such access by the private sector is for private gain, it serves the public good when it creates employment and benefits national economies. The converse route, of openly publishing private sector research, varies between sectors. It may be published in the conventional literature, covered by patents (section 5.3) or retained as a trade secret.

2. Principles for scientific publishing

In evaluating the extent to which modern systems of scientific and scholarly publishing serve the interests of science and its role as a global public good, and whether corrective actions might be needed, it is important to establish what is required of such systems. This section proposes basic principles that need to be observed if scientific publishing is to serve the fundamental roles set out in section 1.2a–e. It is important that, as far as possible, the principles are durable, independent of changes in technologies and modes of working.

An early formulation of these principles was circulated for comment and amendment to the international scientific community, as represented by the membership of ISC, in July 2020, and further discussed in three virtual community meetings in September 2020. The outcomes of those deliberations are embedded in the following reformulations.

2.1 Principles and their rationales

Principle I: There should be universal open access to the record of science, both for authors and readers, with no barriers to participation, in particular those based on ability to pay, institutional privilege, language or geography.

The record of published science is a vital source of ideas, observations, evidence and data that provide fuel and inspiration for further enquiry, and is a profound part of the edifice of human knowledge. That record, including the back catalogues of publishers, should be regarded as a global public good, openly and perennially free to read by citizens, researchers and all societal stakeholders. It is an intrinsic part of the research process and not a separate, add-on enterprise. Authors, regardless of their circumstance, whether funded or not, should not be denied access to the process of publication on the grounds of inability to pay. Economic models should be driven by the needs of science and not by the pursuit of private gain. They should not lead to the privatization of knowledge. Many of the challenges that confront science are both local and global. The local cumulatively impacts upon the global – the global permeates the local. Global solutions therefore require global involvement, and global access to the publications of science, both by readers and authors, is of great importance. Scientific publishing should enable participation from less favoured individuals, institutions and regions, and be open to and facilitate inclusion of diverse voices in local and global scientific conversations.

Principle II: Scientific publications should carry open licences that permit reuse and text and data mining.

The progress of science depends on the ability to access and interrogate evidence and conclusions from past work. Open licences help to promote accountability and traceability, permit authors to continue to derive benefit from their work and maximize the extent to which the work can be built on by others. Yet when submitting to journals, authors may be required to transfer copyright to publishers, inhibiting reuse and preventing the use of powerful text- and data-mining algorithms to uncover patterns, relationships and solutions that may be hidden within the record of science. Such agreements should not be made. As new technologies enhance the capacity to interrogate the whole record of science to discover new knowledge, pathways to access the resources that could facilitate such discovery should be open to all, unrestricted by licensing or ability to pay (Murray-Rust et al., 2014). Privately funded research for private purposes, and research that has direct implications for national security, health and safety may require exemption from this general principle.

Principle III: Rigorous and ongoing peer review must continue to play a key role in creating and maintaining the public record of science.

Peer review processes are essential parts of the curation process in subjecting an author's scholarly work, research or ideas to the scrutiny of independent experts in the same field. The primary purposes of peer review are to ensure that unwarranted claims, fallacious interpretations or offensive views are not published, that truth claims are significant, novel and not plagiarized from elsewhere, and that the reviewed text is comprehensible and logical. Peer review is generally effective in identifying obvious errors of fact and logic, though less likely to identify errors that are more deeply embedded in complex analyses or in voluminous data, as has been revealed by recent tests of reproducibility across several disciplines (Baker, 2016). The growth in publication numbers and diversification of their formats, including preprints on platforms that use post-publication peer review (such as Wellcome Open Research or F1000R), and the need for rapid review at times of crisis, have complicated the position of review in the publishing sequence. It has increased demands on reviewers, who receive no tangible reward for their work. The large and complex data volumes increasingly used by researchers raise the issue of whether and how these should be reviewed. Formal peer review processes should be distinguished from the ongoing, long-term reviews that subject published work to the deep scrutiny and scepticism of the scientific community as part of the normal processes of science. Notwithstanding its importance, one of peer review's dangers is the potential for censorship of originality by uncompromising mainstream views that reject dissent. Kant's 'Critique of Pure Reason', one of the greatest works of philosophy, was ahead of its time and received dire reviews. A fuller review of current issues of debate is contained in section 5.2.

Principle IV: The data and observations on which a published truth claim is based should be concurrently accessible to scrutiny and supported by necessary metadata.

Data and observations supported by the metadata that make them useable, and including software, models and algorithms, must be available for scrutiny when a concept for which they provide evidence is published. If data are withheld for reasons of safety, security or privacy, there should be controlled pathways for access by reviewers and researchers. These processes are vital in enabling others to test the logic of the data/concept relationship, and in attempting to replicate the experiment or observation. Such data should be compiled and curated so as to observe FAIR data criteria⁶. This principle is essential to maintenance of the process of scientific self-correction. Adherence to it would do much to resolve the epidemic of non-reproducibility that has characterized the last decade (Begley and Ellis, 2012; Open Science Collaboration, 2015). For researchers to do otherwise should be regarded as scientific malpractice. Scientific journals should require it.

Principle V: The record of science should be maintained in such a way as to ensure open access by future generations.

The record of science is an essential part of the inheritance of humanity, and should be maintained in such a way as to ensure access by future generations. Prior to the 'digital era', the record was largely preserved and available in indexed books, monographs and journals curated in libraries. Scientific publication now occurs in a plethora of novel forms. As most libraries no longer hold large physical collections but rather manage access to many online resources, there is a danger that access to those digital resources could be lost. There is a strong case for a coordinated network of digital libraries dedicated to the preservation of the scientific record, without a sunset clause, and governed by the scientific community and its institutions and not by commercial entities. There are many regional digital libraries that have been in development for several years, such as the Egyptian Knowledge Bank (EKB) (www.ekb.eg) and the Big Ten Academic Alliance (BTAA) (www.btaa.org), which are parts of

⁶ FAIR data = Findable, Accessible, Interoperable, Reusable. See section 4.4.

international networks that might provide a basis for such a system. In addition, CLOCKSS (Controlled Lots of Copies Keep Stuff Safe – <https://clockss.org>) has developed twelve globally distributed mirror repositories that guarantee long-term preservation and that are chosen for resilience to threats from technological, economic, environmental and political failure.

Principle VI: Publication traditions of different disciplines should be respected, while at the same time recognizing the importance of inter-relating their contributions in the shared enterprise of knowledge.

The disciplines of science – natural, social, engineering and medical – and the humanities, tend to have their own principles of publication that reflect the history, values, cultures and practical norms of work of the discipline (section 1.3). They are individually valuable means of expressing contributions to learning and knowledge and part of the strength of each scholarly community’s inquiring, sceptical and analytic mindset. Although it is clear that no one size fits all, it is important that all are able to agree that contributing to the global public good is a shared purpose and that the processes of publication should avoid creating siloes between disciplines. This is particularly important for the many inter- and trans-disciplinary global challenges faced by humanity, where seeking ways of achieving data interoperability is a major priority⁷. It would be impractical to insist on common standards for publication across the disciplines of science, but it is important that journals are explicit about their standards and that they adhere to them⁸. The Transparency and Openness Promotion (TOP) guidelines (Nosek, 2015) provide a valuable template that could be the basis for such an approach.

Principle VII: Publication systems should be designed so as to continually adapt to new opportunities for beneficial change rather than embedding inflexible systems that inhibit change.

The digital revolution has created new opportunities to enhance the discovery and dissemination of new knowledge in more effective and efficient ways, and new challenges that need to be overcome. The science publication system must be able continually to adapt to and exploit new opportunities that satisfy the principles set out here, and to avoid new threats to its integrity, rather than being inflexible and unresponsive. The integration of scholarship into cyber-infrastructures has enormous benefits for science through its potential to circulate knowledge more broadly, more quickly and more interactively than systems that are still based on the assumptions of print-based publishing, where scarcity rather than ubiquity is seen as an index of value. In this digital world, scalability is essential – the technological ability to scale up or scale down dissemination without bottlenecks (Vinopal and McCormick, 2013).

2.2 Responses from the scientific community

Fifty-six of the ISC’s member organizations responded to a survey in which they were asked to score their level of support for each of the above principles on a scale from zero to ten. The regional and disciplinary distribution of respondents is as shown in Figure 1, and the average scores and their standard deviations and variances for Principles I–VII are shown in Figure 2.

⁷ The ISC’s Committee on Data (CODATA), with the support of the ISC, is currently putting in place a decade-long initiative on this topic. See: <https://codata.org/initiatives/strategic-programme/decadal-programme/>

⁸ The basis for standards could also be derived from regional and global initiatives such as DOAJ, AJOL, Latindex, Redalyc and SciELO for open access papers, and DOAB and OAPEN for open access books.

Scientific bodies		Regional Distributions	
• International Scientific Unions and Associations	34	Africa	2(2)
• National Academies	17	Americas	7(4)
• National Research Councils	2	Asia	7(2)
		Europe	16(9)
		Oceania	3(1)

Figure 1. Respondents to the survey. Numbers in parentheses show the regional locations of respondents from international scientific bodies.

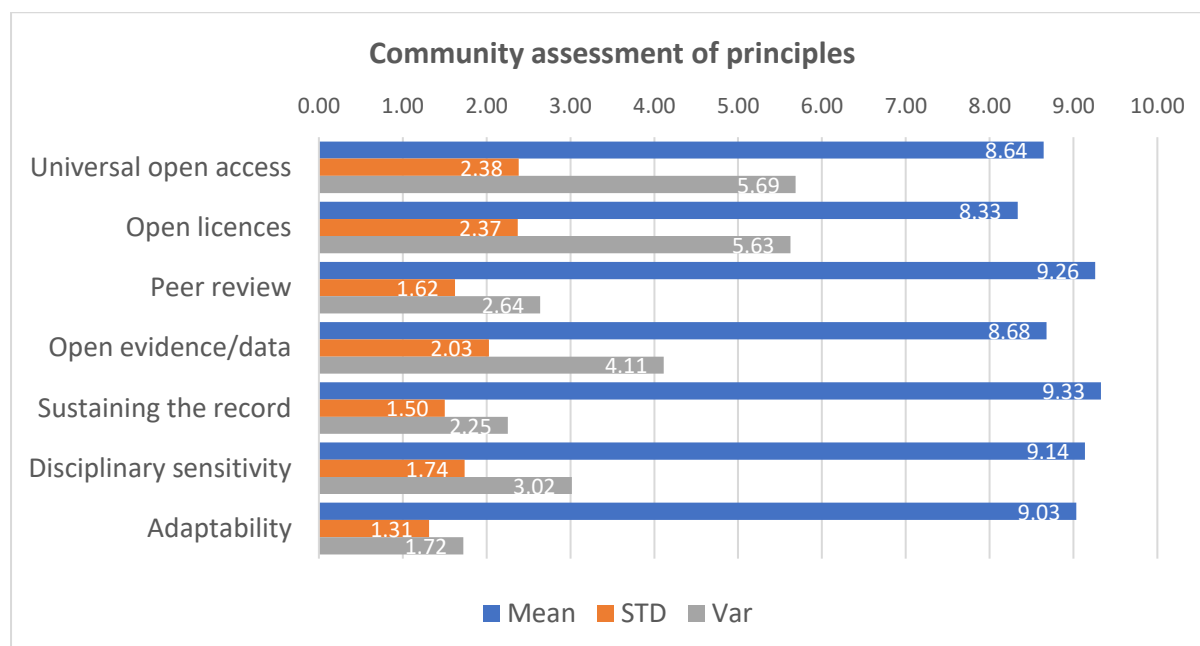


Figure 2. Survey results from ISC members, showing mean scores on a scale from 0 to 10 for the principles in section 2.1, together with standard deviation (STD) and variance (Var) ($n = 56$).

Figure 2 shows a very high level of agreement about the principles amongst respondents. It is noticeable, however, that the highest variance relates to Principles I and II, of open access and open licensing, respectively, implying that whilst most respondents were very strongly in favour of these principles, several gave low scores. The strongest support was given to peer review and to sustaining the record of science, with adaptability showing the least variance.

3. The evolving landscape of scholarly and scientific publishing

The evolution and nature of current publishing practices are now described as a basis for assessing the extent to which they are consistent with the principles outlined in section 2. The landscape has become more complex as publishing responds to the opportunities offered by the digital revolution and to the challenges of the open access movement. An ISC occasional paper on business models for scientific publishing (Gatti, 2020) has contributed to this section.

3.1 The commercialization of scientific publishing

In the first half of the 20th century, the publication of scientific and scholarly journals was largely in the hands of not-for-profit learned societies, as it had been for over 200 years. In most cases, the cost

of a journal was included in the cost of society membership. Learned societies, as formalized disciplinary communities, had unique access to researchers competent to undertake peer review and staff editorial boards, and their members provided a ready-made readership. Until relatively recently, their publishing activities were rarely profitable (Fyfe et al., 2017), but they were the primary means of disseminating new scientific knowledge and, through the approbation of their memberships, in conferring prestige on their journals.

Commercial publishers began to enter this market at scale from the middle of the 20th century, in part as a consequence of growth in the volume and diversity of scientific work and the resultant demand for more diverse and more frequent journal publications. They were more nimble in responding to this rising demand than learned societies, in part because of their capacity to raise the investment capital required to expand and diversify journal production. Many learned society journals were progressively and selectively acquired by commercial publishers. For learned societies, entering into partnership with professional publishers who would take on the production of their journals allowed them to benefit from the economies of scale and marketing capacities of the publishers. By 2004 about half of all learned societies published in partnership with a third party, either a commercial publisher or a university press (Baldwin, 2004). These 'co-publishing' agreements with established publishers replaced deficits by profits and, as discussed in section 3.4, conferred an invaluable prestige and a pre-established readership on the publisher. As learned societies were deemed to act in the interests of science, it seemed natural that scientists should continue to entrust copyright to their work to the journal, offer their services freely to serve on editorial boards and undertake onerous refereeing procedures. However, societies rarely obtained details of the publisher's business model, nor information on the level of profitability of the journal (de Knecht, 2019), yet these developments provided, in aggregate, strong support for a growing global scientific effort.

The essential concern of a profitable commercial publishing enterprise, whether 'non-open access' (section 3.2) or 'open access' (section 3.4), is to minimize unit production costs and to maximize scale. The former has, in part, been achieved by publishers seamlessly adopting the tradition established by learned societies that researchers should staff their editorial boards and undertake peer review for no remuneration. The latter has been achieved in part by the proliferation of ever larger numbers of journals produced by individual publishers and in part by the emergence of mega-journals publishing many thousands of articles each year. The importance of scale in determining profitability is the reason why learned society publishing, usually involving single or small numbers of journals, is often difficult to sustain without additional subsidy or commercial support.

Publishing at scale necessarily means being less selective in what is published. Commercial publishers have not responded to the scale imperative by significantly increasing the publication rate of their most prestigious journals, but rather by introducing new lower-quality journals, or mega-journals with less selective peer review policies.

The income from published work must also cover the costs of managing the rejection process, such that where high standards are set and rejection rates are high, maintaining high levels of profitability requires higher prices. Organizing peer review processes is administratively burdensome, and there are clearly cost advantages for a journal publisher in reducing a journal's overall rejection rate and assessment costs. Taken to its extreme this has led to the emergence of so-called 'predatory publishing' (section 3.10) in which articles are published without any significant peer review process. But we are also witnessing the emergence of numerous alternative strategies to reduce reviewing cost to publishers. Examples include introducing more specialist journals to reduce rejection rates; multi-journal publishers offering a cascade reviewing system (where a single review process is used to determine an appropriate journal for the article across a range of journals with different quality/specialization characteristics); the emergence of journals (notably mega-journals) seeking only

to ensure the research reaches adequate levels of scientific rigour; and journals that rely on post-publication peer review processes.

As we argue elsewhere (section 5.2), effective peer review is a critical component of the scholarly research cycle, and changes in the peer review processes undertaken by publishers are naturally a matter of concern for the scholarly community. However, weaknesses in the existing prepublication peer review process have been well documented and change itself is not necessarily a bad thing, providing that scientific integrity and rigour are maintained.

Journals of high perceived quality tend to have high, costly, rejection rates, but are also able to charge a high mark-up: the combination of high mark-up and large scale being a highly profitable combination⁹. An important marketing strategy for publishers in recent years has been to offer bundled 'big deals' (Box 2) to their customers (universities, their libraries and national agencies). These are inflexible in including both 'high-status' journals and those of lesser standing, although university researchers tend to cite only the high-quality fraction of the journals that their libraries are required to purchase (Shu et al., 2018). The dilemma for customers is exemplified by the demonstration that in disciplines that have been analysed, between 15% and 20% of journals are responsible for 75% of usage (The Research Information Network (RIN)). Acceptance of a bundled deal requires acceptance of a long tail of journals for which there may be little demand, but which serves the commercial interest of the publishers in selling journals for which there would otherwise be a smaller market. In practice, researchers are paying a large premium to access journals that they desire to read by also being required to buy journals that are of relatively less interest to them.

⁹ The publisher Springer-Nature struck a deal in October 2020 with the German Max Planck Digital Library in Munich that offers publication in the journal *Nature* for an APC of €9,500 (Van Noorden, 2020).

The commercial scientific and scholarly publishing market is distinctive in its convoluted relationship between supply and demand. The scientific community provides its work freely, or at its own cost, frequently gifts copyright to publishers, staffs publishers' editorial committees, provides peer reviews freely and then buys back its published work, and in most cases is legally disbarred from interrogating, through text and data mining, the very published record of science of which it is the source. This act of privatization of a predominantly publicly funded resource, without recompense and with a paywall

Box 2: Bundled deals from major commercial publishers

Large commercial publishers sell bundled online subscriptions to their entire list of academic journals at prices significantly lower than the sum of their individual prices. Bundle prices are negotiated institution-by-institution and publishers endeavour to keep them confidential, with many contracts including 'non-disclosure' clauses. Copyright ensures that publishers are able to maintain monopoly control over access to articles they publish, and so are able to price according to the buyer's willingness to pay rather than facing competitive market pressures. Bundling content allows the publisher to maximize profit margins across their portfolio of journals and increases the competitive advantage of weaker journals (that come 'free' with access to the high prestige offerings) over rival outlets. By providing electronic 'site licences' the publishers are able to ensure that access to 'back issues' of journals are part of the offering. If an institution fails to renew its licence it loses electronic access to all back issues of the journals to which it formerly had access. This contrasts with analogue purchases, where cessation of subscription leaves back copies of the journal on the library shelf. Given that increasing numbers of journals are now entirely digital, this means that all access to past publications is dependent on subscription renewal.

A lack of price transparency allows publishers to negotiate with each institution separately and so enables them to price individually to maximize the revenue they are able to obtain separately from each institution. Where information about pricing is available, prices show no systematic variation, but appear to be a result of local haggling (Tijdink et al., 2016). An efficient market is one where there is equality of information. The commercial publishers' unwillingness to reveal either their prices or costs produces a highly lucrative market for incumbent sellers, but one that is uncompetitive and highly disadvantageous for buyers.

In reaction to the disadvantages of the 'big deal' a number of universities have sought ways of avoiding the high costs that they impose on library subscriptions (e.g. Nabe and Fowler, 2012; Schonfeld, 2019).

around it that inhibits access by modern methods of knowledge discovery, may be a unique example of legal, private acquisition of a public resource. The public investment that has gone into production of research results, and the publicly funded work done by researchers for publishers, are all forgotten as a public resource is privatized at little cost to the publishers and their private investors. The largest commercial publishing operations, mostly based in Europe or North America, have almost unique profitability, commonly generating profits in excess of 20–30%, and sustained by annual price increases far in excess of inflation (Larivière et al., 2015). It is a system that can only be rationally justified if there is a balance of mutual benefit between researchers, the public interest and publishers.

Journal 'brand' plays an important role in the system. The financial value to a producer of a highly sought after – luxury – brand is considerable, and the essential element in the marketing of bundled deals. An important element of this system is the co-called 'journal impact factor' (JIF), a scientometric index calculated by Clarivate (now owned by a private equity company¹⁰) that reflects the yearly average number of citations received by articles published in the previous two years in a given journal. It has come to define a hierarchy of quality thresholds that have increasingly been used as proxies to

¹⁰ See <https://clarivate.com/news/churchill-capital-corp-and-clarivate-analytics-announce-merger-agreement/>

assess the scientific quality of an author's work, rather than assessing the quality of the work itself. It is used as a proxy for the relative importance of a journal within its field: journals with higher impact factors being generally deemed to be more important than those with lower ones. The power of the brand as a means of evaluating science has been further exacerbated by the use of the impact factor as a metric of quality by academic tenure committees and grant-funding agencies, reinforcing authors' incentives to target perceived prestige brands for publication. It persuades researchers and their institutions that it is worthwhile paying a premium for publication in a journal with a high impact factor (see section 5.1).

3.2 The reader-pays model

The dominant publishing business model through the 20th century was based on 'reader-pays', in which the publisher receives payment for access by readers – primarily through library subscriptions but also from individual subscribers to learned society journals. This model has served the industry well for many years, and continues to contribute to its income even as open access has expanded. Its primary disadvantage is that access to publications is only available to those who can afford the access fee, thereby limiting the dissemination of scientific knowledge.

A feature of the model is that because the author (or their institution) does not pay to publish, institutions can be relatively relaxed about allowing the author to select what they believe to be an appropriate publisher for their work. This author-centric decision-making process, which was established before the 20th century commercialization (Fyfe et al., 2017), reinforces a behavioural norm whereby authors maintain independent control over the choice of publishing outlets for their research findings. In practice, authors are both producers and customers in the journal market, with producers (researchers as authors) and consumers (researchers as readers) being isolated from any of the costs within the system. In economic terms it represents a fundamental market failure, with the service provider (the publisher) being relatively immune from market pressures and thereby freed to charge a high mark-up (Terry, 2005). Unlike conventional publishing systems, nothing is returned to those that arguably invest most in the process, the author and the peer reviewer. It is in essence a dysfunctional market.

High prices discriminate against access by readers in poorly-funded institutions and national science systems, and undermine the potential of scholars who might otherwise make a greater contribution to understanding and to institutional or national wellbeing. The freedom of authors to choose expensive brands has indirectly imposed considerable costs on the libraries of their home institutions, although there are many institutions with open access policies that assist authors in negotiating copyright agreements, many authors do not use them (pers. comm. M. L. Kennedy, 2020). It must be recognized that through the continued use of impact factors and willingness to transfer copyright, the academic community and its institutions are complicit in processes that ultimately inhibit access to the record of science.

3.3 The open access movement

In reaction to these trends and in awareness of the enormous potential of the digital revolution for scientific publishing, a small gathering of international scholars in Budapest in 2002 published the Budapest Open Access Initiative, the first paragraph of which reads:

An old tradition and a new technology have converged to make possible an unprecedented public good. The old tradition is the willingness of scientists and scholars to publish the fruits of their research in scholarly journals without payment, for the sake of inquiry and knowledge. The new technology is the internet. The public good they make possible is the worldwide

electronic distribution of the peer-reviewed journal literature and completely free and unrestricted access to it by all scientists, scholars, teachers, students and other curious minds. (Budapest Open Access Initiative, 2002)

This stimulated a reaction to reader-pays publishing models and the transfer of an author's copyright to publishers, for example in the Berlin Declaration of 2003:

Open access contributions must satisfy two conditions. The author(s) and right holder(s) of such contributions grant(s) to all users a free, irrevocable, worldwide, right of access to, and a licence to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship (community standards, will continue to provide the mechanism for enforcement of proper attribution and responsible use of the published work, as they do now), as well as the right to make small numbers of printed copies for their personal use.

(Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, 2003)

Such declarations have added impetus towards making scientific publications free for all. Box 3 illustrates the variety of open access models and their evolution in recent years.

Research funders and researchers have a common interest in maximizing the benefits of research by maximizing the reach of research outputs. Open access to the record of science serves this purpose, so that an increasing number of funders are moving to limit the freedom of authors to publish where they wish by requiring that they publish in open access journals as a condition of funding. These perspectives prioritize business models that provide open access to readers whilst minimizing, and in some case removing, costs to authors, recognizing that publishers need at least to cover operating costs and essential maintenance, retain a surplus sufficient for perennial technology upgrades, and weather market downturns or disasters. If articles are to be free to readers, there are four direct sources of funding: authors, institutions, private foundations, or the state.

3.4 The author-pays models

The vision of universal open access to the record of science as described above has stimulated the growth of open access enterprises and persuaded many subscription (reader-pays) publishers that they need to offer 'open access' options in adapting to this new movement. Commercial publishers in particular have adopted an author-pays model based on so-called article processing charges (APCs), which transfer costs from readers to authors. In order to maintain profit levels, many journals set APCs at rates that can still be prohibitively expensive for many (Burchardt, 2014), such that changing existing journals from subscription to APCs is unlikely to resolve many of the problems of the current system, and may even entrench commercial power (Tennant, 2019)¹¹.

Box 3. Open access operating models (Adapted from Brainard, 2021)

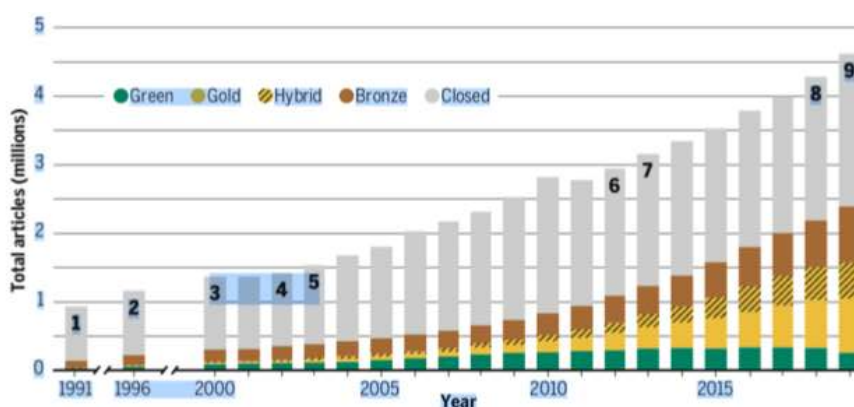
Green. Authors or publishers deposit articles in a public repository, where they are free to read. But journal embargoes can delay posting.

Gold. Articles are published with a license making them immediately free to read. Authors or institutions typically pay journals for this service. Gold journals publish only gold articles.

Hybrid. Hybrid journals offer gold open-access publication but also publish other articles behind a paywall and continue to charge for subscriptions.

Bronze. Articles are free to read on publishers' websites, but the papers are not licensed as open access, allowing publishers to place the articles behind paywalls later.

Closed. Journals keep articles behind subscription paywalls.



Staging points in the evolution of open access publishing

1. ArXiv created. A preprint server allowing free online reading of manuscripts. **2.** The *Journal of Clinical Investigation* becomes the first prominent journal to provide free online content. **3.** BioMed Central, the first open-access, for-profit scientific publisher. **4.** The Budapest Open Access Initiative. **5.** The Berlin Declaration on Open Access to Knowledge. **6.** More than 2,600 scientists vow not to publish in or referee for journals of the publisher Elsevier. **7.** The US White House Office of Science and Technology Policy requires US-funded researchers funded to make articles them open within 12 months after publication. **8.** Coalition S formed. **9.** Springer Nature and German institutions sign the largest 'transformative agreement' allowing authors to publish open access without paying per-article.

¹¹ Hindawi, an innovator in open access publishing that charges APCs, with over 200 titles and one of the fastest growing publishers, was acquired by Wiley in early 2021 for \$298 million; <https://www.businesswire.com/news/home/20210105005201/en/Wiley-Announces-the-Acquisition-of-Hindawi>).

Branding remains critically important for publishers in this model, as it is the strength of their brand that allows publishers to charge high fees to authors¹² but the mechanisms for sustaining those brands are now different. For publishers, brand matters primarily because of the impact it has on author submission decisions. It is the persistent link between journal brand and perceived research quality that permits the publishers of 'high impact' journals to maintain high profit margins, ties authors into following publishing norms they often dislike and sustains inequities in the international science system for both reader-pays and author-pays options. The concept of the high-impact journal is essential to the capacity to charge high APCs for open access journals.

Many universities have adopted institutional processes that conform to open access principles (e.g. Harvard Faculty of Arts and Sciences Open Access Policy, 2008) and there has been a partial disruption of the earlier dominance of subscription journals, resulting in the development of open access journals that conform to a greater or lesser extent to the criteria of the Berlin Declaration. They have progressively increased in market share and are now estimated to make up about 47% of all scientific journals (Piwowar et al., 2019). According to the Directory of Open Access Journals (DOAJ), 70% of open access journals do not charge APCs.

A major intervention on behalf of open access and in opposition to subscription (reader-pays) models was instituted by the European Union as 'Plan S' in 2018, with the rationale that paywalls withhold research results (including data) from the scientific community and from society as a whole in a way that 'hinder the scientific enterprise in its very foundations and hampers its uptake by society' (Science Europe, 2018). It has led to creation of 'cOAlition S', a group of international research funders that is committed to ensuring that research resulting from its members' grants must be published in open access journals or platforms or made openly and immediately available in an open access repository. It requires that subscription-based models of scientific publishing, including its so-called 'hybrid' variants, should be terminated. cOAlition S will also, by the end of 2021, issue a statement on Plan S principles as they apply to monographs and book chapters, together with related implementation guidance, and strongly encourages early sharing of research results through preprints. It was initially planned that Plan S would place a cap on APCs as a key principle, but that has since been dropped, resulting in a fear that APC prices might explode. Although Plan S, now implemented, is a step to limit the freedom of authors to choose where they place their research, it is not yet clear whether Plan S will lead to competitive price reductions when there is little control on author-pays APCs and no restraint on authors' freedom to use 'someone else's money' in choosing a highly priced APC (see section 4.2).

The shift of financial responsibility to authors has other negative consequences. Researchers in institutions or national science systems with budgetary constraints, or authors without access to external funds may be unable to publish in high-profile journals with high APCs, with authors in low- and middle-income countries being particularly discriminated against. At the same time, the use of journal impact metrics as proxies in promotion and funding assessments further disadvantages financially-stressed researchers and institutions (Houghton and Vickery, 2005; Burchardt, 2014).

The inclusion of APCs as a permitted model in Plan S is of particular concern for science in developing regions, as illustrated by the following comments from Debat and Babini (2019): that Plan S 'reveals a patronizing view of scientific sharing which translates into the control of science in the hands of rich countries and diminishes the Global South as a mere passive observer with no control beyond global commercial agreements between wealthy governments and the few large oligopolists commercial publishers', that 'we cannot emphasize more that a reasonable APC for a Global North research

¹² Nina Schönfelder (2020) shows that author charges for open access publication increase with the journal's citation measure [Source Normalized Impact per Paper (SNIP)], with greater responsiveness observed in pure open access journals than in hybrid journals.

institution will most probably be unaffordable and unreasonable for a developing region institution’ and, as pointed out by Holmwood (2018), ‘private benefit is adopting the mantle of public value and, if the advocates of commercialization succeed, the loss will be that of the public in whose name it is taking place’.

The revised implementation guidance for Plan S published in 2019 (cOAlition S, 2019) made clear that author fees were not the only model for open access, and clarified that they would not necessarily introduce a cap on such fees, but would rather require transparency about fee structure and services provided. Work is ongoing to establish the best way to share and aggregate this data. This is a crucial issue requiring urgent attention, for at a global level, the author-pays model is only open in a limited sense. It is a closed door to many authors without the level of financial support required to unlock it. It is open to some, closed to others.

3.5 Learned society publishing

Learned societies played the foundational role for the modern era of scientific and scholarly publishing dating from their creation of the first scientific journals in the 17th century. Many still publish journals, including an increasing number that publish in partnership with non-society publishers. Receiving a society journal as a part of the membership fee can be an important inducement to join a learned society, whilst the society’s name can both enhance the prestige of its journals and be enhanced by the intellectual legacy that the journals represent. Income from journal publishing can be important in supporting a society’s other activities, such as conference organization or capacity building for early-career researchers, although there is much variation in the revenues that societies receive from journal publishing. A 2014 survey by Taylor & Francis and the Association of Learned and Professional Society Publishers (Frass, 2015) found that whilst 19% of societies surveyed receive less than 5% of their revenue from publishing, 10% make nearly all of their income in this way. Some societies that have many journals and very large memberships earn up to US\$4 million per year in royalties (de Knecht, 2019).

In response to the call for open access, learned societies have generally moved into publishing hybrid open access journals, funded through APCs (Wise and Estelle, 2019). Moving away from hybrid journals and to full open access (as required by cOAlition S funders, and as envisaged by the principles outlined in this paper) are proving challenging for many societies, particularly smaller ones that operate resource-intensive publishing processes with single or limited numbers of journals and without the easy access to finance of large publishing companies.

There is a renewed interest in the role of learned societies in knowledge creation and dissemination, with initiatives designed to support them in sustainably adapting to open access publishing modes and developing their influence in a thriving scholarly research ecosystem. The Society Publishers Accelerating Open Access and Plan S (SPA-OPS) initiative¹³ seeks to identify routes by which learned society publishers might successfully transition to open access and align with Plan S. PeerJx offers societies the use of its online journal production platforms in sustaining their publishing activities as part of a wider learned society community¹⁴.

Societies are deeply engaged in the preprint domain (see section 3.7), and research libraries are working with societies as part of the scientific community’s efforts to expand access for both scholars

¹³ See

https://wellcome.figshare.com/collections/Society_Publishers_Accelerating_Open_access_and_Plan_S_SPA-OPS_project/4561397

¹⁴ See <https://peerj.com/blog/post/115284883221/partner-with-peerj-to-build-a-new-ecosystem-for-society-publishing/>

and the public. As the international representative for international scientific unions and associations, the ISC strongly supports the enhancement of the contribution of learned societies to the scholarly publishing landscape.

3.6 Institutionally-based repositories and infrastructures

There is a long history of universities financially supporting publishing operations – often through university presses and library publishers. Several such presses play major roles in the production of scholarly books and monographs, primarily of importance in the humanities and social sciences. Universities, university libraries and related institutions have relatively recently acquired new roles in response to the imperatives of the open science movement, through the establishment of online archives or repositories. In satisfying the principle of ‘free-to-read’, in the so-called ‘green’ open access mode (Box 2), an author may self-archive in an institutional, or specialist, online archive a copy of an article that has been published in a journal that agrees to this mode. The ‘green’ mode can apply whether or not the original journal article was funded through an APC. In addition, there are a growing number of cases that do not rely on payment at any point from either reader or author, as part of a business model that is based on an institutionally-funded digital publishing infrastructure. In a digital world, free from the restrictions of paper and print, such developments offer a major challenge and opportunity to universities in particular, to re-invigorate their historical roles as both creators and disseminators of knowledge.

There are many not-for-profit, scholar-led initiatives to publish open access journals, and university libraries that organize online editorial processes for their open access journals and journal collections using an Open Journal System (OJS-PKP) platform. There are a growing number of cooperative and collaborative initiatives between libraries and academic presses to manage journal collections, such as the Open Library of Humanities, Coalition Publica and OpenEdition Journals. A number of major open access initiatives are currently under way that offer alternative approaches where universities and research institutions have a leading role (Rodrigues, 2020). As noted in section 3.5, many universities have long supported open access by hosting post-prints on their servers (in the ‘green’ open access mode). This has been built on by the Confederation of Open Access Repositories (COAR), which involves university libraries in collecting, preserving and providing access to their research outputs in institutional repositories, usually in university libraries, whilst research communities continue to undertake certification and quality control using traditional peer review. In 2016, COAR launched the Next Generation Repositories initiative, which aims to position research institutions and their repositories as the foundation for a distributed, globally networked infrastructure for scholarly communication, on top of which layers of value-added services are to be offered.

3.7 Preprint repositories

Repositories may also act as archives that contain studies that have not yet been peer-reviewed or published in an academic journal. Some are hosted by university libraries and some by subject-specific repositories. The leading preprint archives tend to perform an initial screening of submissions to ensure that they are authentic scientific papers (e.g. arXiv, bioRxiv, medRxiv and SSRN) but without strict formatting guidelines (Box 4). Preprint repositories aim to change the way research is published and ideas are disseminated. Researchers can access the manuscripts, use them for their own research and share comments. They are more cost-efficient than traditional publishing, and offer a more open forum for improvement and critique of a work so that novel truth claims can be subject to a wider critique than conventional peer review. They short-circuit the process of publicizing an idea, permit authors to stake a claim to it, circumvent the expense and delay of publication and permit authors to avoid so-called ‘predatory journals’ (see section 3.10). When authors are comfortable with their

manuscripts, they can then find an appropriate journal for publication. Some have a direct transfer service to specified established journals.

An alternative means of enhancing final preprints is through an overlay journal, which avoids some of the costs associated with conventional journal publication (see section 5.5). The overlay journal is an important development in which the journal manages the peer review function on final preprint versions of an article in an online repository. Refereed and accepted versions are then republished with the repository and registered with a digital object identifier (DOI) that certifies that the article is the final version (Scholastica, 2019). arXiv for example, as a major repository, has served as a natural launchpad for many of today's overlay journals.

Box 4. Preprint servers

The number of preprint repositories and the disciplines they serve have expanded significantly since the early 1990s. Leading preprint servers now include [ArXiv](#) covering a wide range of fields, but particularly in mathematics and physics, [SocArXiv](#) in the social sciences, [BioRxiv](#) in biology, [EngrXiv](#) in engineering, [ChemRxiv](#) in chemistry, [PsyArXiv](#) in psychological sciences, [LawArXiv](#) in legal scholarship and [EarthArXiv](#) in Earth science fields. They are also developing outside the hitherto dominant Europe–North America nexus of publishing, in Indonesia ([INA-Rxiv](#)), India ([IndiaRxiv](#)) and Africa ([AfricArxiv](#)). Many preprint repositories are funded on a not-for-profit basis by institutions, and where this is done by universities, they tend to be extensions of pre-existing publishing activities. There has also been significant growth in the number of preprint servers owned by commercial publishers, such as Elsevier's SSRN and FirstLook, or Springer's InReview (Schonfeld, 2020).

The role of preprints has been highlighted by the COVID-19 pandemic. This has created urgent demands from policy-makers, politicians and the public for access to potentially relevant scientific knowledge, irrespective of its state of readiness for publication, where the slowness of conventional peer review has inhibited its availability. The speed with which science has responded to these demands has benefitted greatly from the well-developed procedures for access to preprints developed over recent decades, although the crisis has drawn attention to the need for new forms of urgent, agile peer review (Rovenskaya et al., 2020) to mitigate the damage that can be done by faulty analysis or logic, and the risk of preprint articles being seized and politicized by commentators wishing to spread ideologically motivated positions without due care for the reproducibility of the scientific claims. Social media have also been one of the primary discovery channels for preprint articles, where the latter have been an important bridge for the public desire to know and understand. Social media channels also became the venue for rapid 'crowd-sourced' peer review of preprints (Vlasschaert et al., 2020). However, given the rapid spread of false news on social media (Vosoughi et al., 2018), there is cause for concern about how preprint articles that are later debunked by the scientific community could be used to spread misinformation via social media.

Modern digital technologies can offer powerful economies of scale that facilitate cost-efficient processes. For example, the Open Science Foundation, working at scale to use shared infrastructure and open-source software that permits authors to include data, code and other supplementary information, have forecast that by late 2020 it will cost \$300,000 to post 33,650 papers, at an average cost of \$6.81 per paper (Mellor, 2020). However, developing business models that ensure long-term financial sustainability in not-for-profit repositories has often proved difficult, as many preprint services are operated by volunteer academic groups and dependent on one-time funding or grants from foundations. Moreover, a successful service requires a holistic technical and managerial approach. It is not sufficient merely to have readily available repository architectures.

Many major research funders (e.g. Wellcome Trust, Gates Foundation and US National Institute of Health) recognize the value and endorse the practice of preprints. Many major publishers have also entered the preprint business, such that the majority encourage researchers to use preliminary preprinting prior to submission in their journals, and Crossref allows preprints to be allocated DOIs. Foundations concerned to support the move towards open science have formed the Open Research Funders Group (ORFG; www.orfg.org), a partnership of 16 philanthropic organizations committed to the open sharing of research outputs as a means of accelerating the pace of discovery, reducing information-sharing gaps, encouraging innovation and promoting reproducibility.

The conversation has therefore moved on from whether to develop preprints, to how best to embed them as distinctive and valuable parts of the scientific record with the following:

- sustainable business models that satisfy the diversity of regional and disciplinary needs;
- transparent and systematic screening procedures both before and after submission;
- clear emphasis that preprints are preliminary reports, as yet uncertified by peer review, and not to be reported by the news media as established conclusions.

Realizing the potential value of repositories and avoiding the pitfalls will depend upon the following:

- the extent to which they are taken up by the international research community as valuable elements in the acquisition and dissemination of knowledge;
- the extent to which they are consistent with priorities for rigour, global inclusion and equity;
- the extent to which a distributed, flexible, but globally networked infrastructure that is responsive to regional and disciplinary needs can be realized.

As established institutions with deep commitments to scholarship and networked internationally, the universities are an ideal basis for such networks. COAR is spearheading efforts to promote the design of a distributed global networked infrastructure for repositories, on top of which layers of value-added services can be added, including open access overlay journals able to take texts from existing preprint services, subject them to editorial review and, in some cases, full peer review (Ferwerda et al., 2017). However, given that academic libraries have suffered budget cuts in recent years and that lack of financial resources is reported as the biggest barrier to investing in online resources (Frederick and Wolff-Eisenberg, 2019), it is clear that securing additional long-term funding for digital infrastructure remains a challenge.

Preprints have the potential to offer a fundamental opportunity to evolve scientific publishing in ways that avoid the drawbacks of conventional open access publishing, and that – with appropriate licensing – avoid its capture by commercial interests for commercial benefit (McKenzie, 2017). A key issue is that of sustainable financing, with systems such BioRxiv having the potential to show the way (Schneider, 2019).

3.8 ‘Public infrastructures’ – publicly funded and scholar-led

Notwithstanding the reduction of production and dissemination costs inherent in digital systems, the costs of system installation and maintenance is significant, and a deterrent to the entry of new and innovative groups that lack the investment capacity of major commercial publishers that are able to raise investments from financial markets. This problem has been overcome by a number of initiatives that have created cost-effective ‘public infrastructures’ that significantly lower barriers to entry, and

have a substantial impact by encouraging and sustaining a large and diverse range of mission-led initiatives¹⁵.

Latin America has led the way, through inter-governmental agreements (see <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/access-by-region/latin-america-and-the-caribbean/>) in making the transition to open access through the development of non-profit repositories as the preferred locations for research publications and research data (Babini, 2020). It has developed into the most advanced open access system of scholarly communications based on the percentage of research publications available through publicly funded initiatives (Alperin, 2015).

In 2012, nine Latin American public science and technology agencies (Argentina, Brazil, Colombia, Costa Rica, Chile, Ecuador, El Salvador, Mexico and Perú) agreed to develop national systems of repositories to coordinate funding, training and national and regional cooperation. They also created a Federated Network of Institutional Repositories of Scientific Publications, known as La Referencia with the presidency rotating between the national participants. La Referencia boosts interoperability agreements in the region, with its regional harvester having 1,431,703 full-text, peer-reviewed articles, theses and research reports. La Referencia observes the OpenAIRE interoperability guidelines, and is an active member of COAR, working together with the participation of repositories worldwide towards an international network of repositories, and functionality for next-generation repositories (OpenAIRE Guidelines, 2015).

Scientific Electronic Library Online (SciELO) is an international network of over 1,700 journals, primarily, but not exclusively, based in Latin America. There are several important technological features of this network. From the outset it has been designed as a modular platform, developed using open-source software and adopting open standards and protocols between operational units¹⁶. It can support innovative initiatives that may be based on different technological approaches, and the economies of scale permit technologies to be redeveloped whilst mitigating against participants being locked-in to outmoded technologies and approaches (Principle VII, section 2.1).

Important elements in this system are university-based publishing systems for peer-reviewed, open-access journals. Redalyc (Red de Revistas Científicas de América Latina y el Caribe, España y Portugal) is one such system that works with open-access social science journals from Latin America, the Caribbean, Spain and Portugal. It currently has 1,294 active open-access, peer-reviewed journals published across all disciplines. Fewer than 5% of these journals charge APCs, but recently Redalyc has decided not to accept journals that charge APCs. Thousands of authors have created profiles in Redalyc, linked to ORCID when available. Bibliometric and scientometric indicators are provided at publication, institution, country and discipline levels. Redalyc also manages the Latin American Council of Social Sciences' (CLACSO) collection of 930 peer-reviewed social science and humanities journals (387,018 full texts) available in CLACSO's digital repository (Red de Bibliotecas Virtuales de Ciencias Sociales), which also provides open access to CLACSO members journals, books, working documents,

¹⁵ In recent years we have seen initiatives such as Invest in Open Infrastructure (<https://investinopen.org>) and SCOSS (<https://scoss.org/>) advocating for coordinated funding of these infrastructures, and direct funding for initiatives developing such infrastructures from public research funding bodies such as Research England (e.g. COPIM; <https://www.copim.ac.uk/>) and the European Funding Council (e.g. OPERAS; <https://operas.hypotheses.org/>). The ORFG (www.orfg.org), a partnership of 16 philanthropic organizations, has also stated its commitment to the open sharing of research outputs.

¹⁶ It should be noted that open-source code is not in itself sufficient to create an open platform. Android is built using open-source code; however, it has been effectively controlled by Google through technical dependencies and closed protocols between some key components of the system (Google Play Services).

research reports, theses and multimedia, comprising over 110,000 full, open-access texts, with more than two million downloads a month (CLACSO, 2019).

Regional digital libraries such as Redalyc, CLACSO and SciELO have become particularly important for the Global South, where knowledge production has had a low visibility in traditional indexing services. Latindex (Online Regional Information System for Scientific Journals from Latin America, the Caribbean, Spain and Portugal), which started in 1998 as the Latindex Directory, provides the necessary basic information about journals in the region.

Disciplinary repositories are also present in the global open-access landscape, such as the UN information system for agriculture, AGRIS (<http://agris.fao.org>), which offers open access to three million full-texts.

Existing repositories are also complemented by innovations such as Pubfair, a framework for open-access publishing, which enriches a variety of research outputs (including preprints, data and software), managed by repositories or other data providers, with additional services that support quality control, dissemination and discovery. The aim is to provide publishing services that enable researchers to share a wide array of research outputs, to support trusted evaluation and assessment processes, and to empower research communities, funders, institutions and scholarly societies to create novel dissemination channels. As indicated above (section 3.4), there is currently concern that initiatives such as cOAlition S could disrupt these regional solutions (Debat and Babini, 2019).

3.9 Books and monographs

Books and monographs are important vehicles of publication in the humanities and social sciences in particular. However, sales of monographs are declining (Ferwerda et al., 2017), and publishers are currently rethinking their models (Open Access Directory: <http://oad.simmons.edu/>) in ways that favour transition to open access for book-length publications based on Book Processing Charges (BPCs). There are a growing number of experiments in university and academic publishing houses, as well as other scholar-led collaborative initiatives, with an emphasis on the social sciences and humanities (Adema and Schmidt, 2010; Universities UK Open Access and Monographs Group, 2019). Examples include university press collections of open access books, the OAPEN-Online Library and Publication Platform, COPIM (Community-led Open Publication Infrastructures for Monographs), OpenBooks Publishers, OpenEdition Books, SciELO Books, TOME (Toward an Open Monograph Ecosystem) and CLACSO's Open Access Books – some, but not all, charge BPCs. To improve the discoverability of high quality academic open-access books, the Directory of Open Access Books has been developed. It requires books to be subjected to independent and external peer review prior to publication, and to be made available under an open access licence (such as a Creative Commons licence).

A fundamental dilemma has long been how to develop a business model for open-access monographs that works at scale for publishers, libraries and scholars. Novel economic models for monographs and books, are being developed, such as 'Opening the Future' (<https://www.openingthefuture.net/>). It is particularly relevant to many mid-sized university presses. It preserves print, presents a low risk and is affordable for libraries whilst avoiding charging authors. Most importantly, it is a model that scales dynamically: as membership grows, books are made open access as soon as a press hits the revenue threshold, meaning that it is not an 'all or nothing' approach (Eve, 2020). Within social science and humanities publishing, the TOME initiative advocates a model in which open-access monograph publishing costs are met by institutionally funded book subsidies.

3.10 'Predatory' publishing

In recent decades there has been strong growth of published research output contained, by 2018, in over 33,000 peer-reviewed English-language journals (STM, 2018). The growth rate in science and engineering alone has been about 4% per year, rising from 1.8 to 2.6 million articles per year between 2008 and 2018 (White, 2019). The change has been associated with an increase in numbers of researchers worldwide (a growth of 21% between 2007 and 2013) (UNESCO, 2015) and incentives for university researchers that are interpreted as 'publish or perish' (Mandke, 2019). This lucrative, demand-driven market has also spawned journals with low publishing standards, little if any peer review and offering rapid online publication. They have been characterized as 'predatory', defined as 'entities that prioritize self-interest at the expense of scholarship and are characterized by false and misleading information, deviation from best editorial and publications practices, a lack of transparency, and/or the use of aggressive and indiscriminate solicitation practices' (Grudniewicz et al., 2019). Digital production and author-pays models (the reader-pays market is limited) are exploited by predatory publishers to enhance market penetration (Siler, 2020).

Researchers from developing countries are particularly prone to publish in predatory journals (Xia et al., 2014). A 2015 study (Shen and Björk, 2015) showed that 75% of authors in predatory journals are from Asia and Africa. Although it is presumed that many authors are misled into publishing in predatory journals by carefully targeted advertisements, there is evidence that this can be an effective personal advancement strategy and a deliberate and rational choice where assessments only use the numbers of publications as the bibliometric criterion (Seethapathy et al., 2015).

Those who are misled should be aware that publishing in such journals has very little impact on science (Singh Chawla, 2020). In one study, 60% of articles published in predatory journals received no citations over the five-year period following publication (Brainard, 2020). The harm that such journals do is to use up the time and resources of academics who might otherwise be better employed, and to contribute to the long tail of inconsequential research. An important priority therefore is to find ways in which such actions are not incentivized, and that potential authors and readers are directed away from such journals. It is hoped that an ongoing project on predatory publishing by the InterAcademy Partnership (IAP)¹⁷ will point the way to such solutions.

4. Publishing the data of science

Making observations and undertaking experiments that enlarge understanding are highly creative acts, and the data derived from them are often at least as important as the text publication that reports the insight. They are first-class scientific outputs (Callaghan et al., 2012) and as essential parts of the record of science should be as accessible as conventional publications.

Most, but not all, of the data of science are acquired as a consequence of scientific inquiry through experiments, observations or surveys. In many fields, particularly in the social sciences, medical science and humanities, much of the data that are used in research may be derived from government statistical surveys, health systems, commercial activities, social media platforms or from other public or private sources. If used for scientific purposes, methods of collection, sample representativeness and ethical standards for access and use need to be rigorously assessed, at which point they become part of the corpus of scientific data; see for example the NiST Research Data Framework (RDaF), <https://www.nist.gov/programs-projects/research-data-framework-rdaf>.

¹⁷ See: <https://www.interacademies.org/news/launch-new-iap-project>.

4.1. Open data: scrutiny of the evidence for truth claims

The association of experimental or observational data with a truth claim in a published article for which it provides the evidence is an essential element of empirically based scientific inquiry. Prior to recent decades, there have been relatively few areas of science where the volume of data was so large that it could not be included as part of the publication that relied upon it. The digital revolution has so expanded the volumes, fluxes and disciplinary diversities of the data available to and used by researchers that there are now increasing instances where data cannot readily be contained within the confines of even a digital-only article. This mostly concerns the natural, medical and engineering sciences, but also increasing sectors of the social sciences and humanities (for example, in large digital databases of language, and in discourses from online sources such as Twitter, Instagram and social messaging). The fundamental reason why the article and the related data must be associated is to permit scrutiny of the evidence for a truth claim and the logic of their relationship, which forms the basis of the principle of self-correction and the maintenance of scientific rigour. However, the difficulty of doing so with increasingly large and complex datasets, or through a desire to withhold data for whatever reason (see Box 5), have created a situation in which data and metadata are not routinely available alongside a published truth claim. Such an omission undermines the capacity to subject the logic of the claim to scrutiny. It has contributed to the so-called crisis of replication by making it impossible to test the replicability or even the honesty of a published truth claim (Baker, 2016; Miyakawa, 2020). It is an omission that must and can be corrected (National Academies of Sciences, Engineering and Medicine, 2019). Publishing the data is as important, and sometimes more important, than publishing the written text. The case for openness of the evidence and data that are relied upon to support a scientific argument have been powerfully made. Charles Darwin made the case: ‘False facts are highly injurious to the progress of science, for they often endure long; but false views, if supported by some evidence, do little harm, for everyone takes a salutary pleasure in proving their falseness; and when this is done, one path towards error is closed and the road to truth is often at the same time opened’ (Darwin, 1871). The impact of published ideas may be ephemeral, the impact of facts is not. Yet our modern perspective seems to be the reverse. The published concept is preserved and indexed, but whether the same happens to the related data tends to be left to the whim of the author.

Richard Feynman, one of the 20th century’s great physicists, was unequivocal about the open exposure of evidence relating to a scientific truth claim (Box 5). Although written from the perspective of a natural scientist, we see no reason why a similar logic, appropriately formulated, should not apply to all the disciplines of scholarly inquiry. Observing these structures would do much to resolve the issues of reproducibility that have plagued many disciplines in the last decade and poor research practices such as p-hacking (selective reporting of those statistical analyses that show significant results) (Nuzzo, 2014), poorly designed sampling and ill-defined training sets for machine learning. These issues are more common than they should be and play into opportunities to criticize science more generally. It has been argued that the preregistration of data analytic procedures prior to examination of data could do much to avoid inappropriate data selectivity and improve reproducibility (Nosek et al., 2018).

Box 5: ‘The first principle is that you must not fool yourself’.

Richard Feynman – Commencement address at CalTech in 1974

‘There is an idea that we all hope you have learned in studying science in school – we never explicitly say what this is, but just hope that you catch on by all the examples of scientific investigation. It’s a kind of scientific integrity, a principle of scientific thought that corresponds to a kind of utter honesty – a kind of leaning over backwards. For example, if you’re doing an experiment, you should report everything that you think might make it invalid – not only what you think is right about it; other causes that could possibly explain your results; and things you thought of that you’ve eliminated by some other experiment, and how they worked – to make sure the other fellow can tell they have been eliminated. Details that could throw doubt on your interpretation must be given, if you know them. You must do the best you can – if you know anything at all wrong, or possibly wrong – to explain it. If you make a theory, for example, and advertise it, or put it out, then you must also put down all the facts that disagree with it, as well as those that agree with it. In summary, the idea is to try to give **all** the information to help others to judge the value of your contribution, not just the information that leads to judgment in one particular direction or another.’

4.2 Binary publication

For cases where the relevant data and related metadata cannot be accommodated within a text publication, a formal ‘binary-publication’ system is essential if openness and rigour are to be sustained (Box 6). One part of binary publication would be a paper, monograph or book that uses the data. Its twin would be concurrent publication of the data in a trusted data repository, in the publisher’s data repository or in a specialist data journal that manages data deposition, and where data related to a text can be accessed via a reference in the paper. The ideal of course is that the text and the data should be mutually digitally available and interoperable. It is a process that has long been technically possible, but one that has not been widely developed. It exemplifies the importance of Principle VII in section 2.

The demands of open-data deposition can be stringent (see section 4.4), but could be incentivized if data were formally published, citable and widely regarded as equivalent in standing to a conventional text paper. The processes associated with data deposited carrying a persistent DOI in a trustworthy repository are summarized in section 4.4, but it would also be important to identify the data with an appropriate title and abstract to underline its status as a first-order output of scientific inquiry that is equivalent to that of a text publication. Some datasets are perennially updated, a process that could be recognized and cited as a cumulative contribution to science. Although the binary publications described here are generally the most important outputs of research, section 7.2 describes how other elements of the ‘research cycle’ are increasingly accessible and potentially publishable. Creative ways of developing the data-publication space are illustrated in the development of ‘nanopublications’ (http://nanopub.org/guidelines/working_draft/). A nanopublication is the smallest unit of publishable information. It can be about anything, for example a relation between a gene and a disease or an opinion. It can be readily contained in a FAIR data (see section 4.4) and machine interpretable framework and can be disseminated as data, as an independent publication with or without an accompanying research article. Because they can be attributed and cited, nanopublications provide incentives for researchers to make their data available in standard formats that improve data accessibility and interoperability.

Box 6: Binary publication of article and related data, an example

The Nature journal *Scientific Data* requires that authors deposit their data in a recommended data repository as part of the manuscript submission process. Manuscripts will not otherwise be sent for review. The datasets must be made available to editors and referees at the time of submission, and must be shared with the scientific community as a condition of publication. The publisher does not host data, but asks authors to submit datasets to an appropriate, publicly accessible data repository. Data should be submitted to discipline-specific, community-recognized repositories where possible, or to generalist repositories if no suitable community resource is available. If data have not been deposited in a repository prior to manuscript submission, authors can upload their data to Figshare or the Dryad Digital Repository during the submission process. Data may also be deposited in these resources temporarily if the main host repository does not support confidential peer review. The ultimate repositories must meet the journal's requirements for data access, preservation and stability. The journal provides a 'date-stamped archive of our recommended repository list', which is available for use under the CC-BY licence (see Box 7). Recommended repositories and standards that are indexed by FAIRsharing, can be also be viewed and filtered via the *Scientific Data* FAIRsharing collection.

There are situations in which open access to data is difficult or even inappropriate: particularly where access would prejudice privacy, safety, security (Journal Editors and Authors Group, 2003), where it has potential for harmful dual use, where the data has been created for commercial purposes or where there are legitimate concerns about access to the data of indigenous populations. Concern for the data rights of indigenous peoples is an increasingly important issue, reflected in the CARE principles (<https://www.gida-global.org/care>). These include the right to create value from indigenous data in ways that are grounded in indigenous worldviews and realize opportunities within the knowledge economy. Where the data of study has been acquired from third parties, as is common in economics, other areas of the social sciences, and in medicine, researchers may be required to sign confidentiality agreements about data use, or may only have access to aggregated data. In such cases it may not be possible to host the data separately from the source, or to verify the integrity of the data or of an 'aggregation process' (Cai et al., 2019). These issues complicate the process of data publication and reuse. Careful thought is needed about the research norms that should be applied in such cases. Increasing trends for public-private partnerships involving business frequently require special arrangements whereby restrictions are placed on precisely what may be openly published whilst holding back commercially sensitive data (see section 5.3). In many of the instances where data privacy is an important issue, particular attention needs to be paid to processes of data governance (Royal Society and British Academy, 2017).

Nonetheless, the scientific case for scrutiny by peer reviewers and researchers in these circumstances remains. In such cases it is important that (a) the data are retained somewhere, (b) there are pathways to access for referees and researchers and (c) there are rich metadata and FAIR criteria are satisfied (see section 4.4). The access pathway should be regulated but must exist so that access cannot be unreasonably withheld. There is an urgent need for the scientific community and publishers to work together to establish protocols for such pathways. This is particularly important for personalized data, where there are currently no fool-proof methods of de-identification, although in some cases 'safe havens', where data are maintained in a password-protected independent network, can be controlled by a steering committee to ensure that publication does not reveal sensitive data (Royal Society, 2012).

4.3 The general case for open data

The evolving perspectives of the open science movement have extended beyond the frame discussed above to embrace broader horizons of open data to include the data that are not logically required in publications. The case has been powerfully made over the last two decades in a series of influential reports (Royal Society, 2012; Science International, 2015; National Academies of Sciences, Engineering and Medicine, 2018).

There are pragmatic benefits to this approach. Firstly, subsequent users may find value that the data originator has not seen¹⁸. Secondly, unless the habit and the means are developed of making scientific data openly and routinely available¹⁹ and interoperable, the opportunity will be lost to collate and integrate data from a variety of disciplinary sources to investigate complex systems to which individual datasets contribute partial perspectives (e.g. <https://codata.org/initiatives/strategic-programme/decadal-programme/>). Exploiting such opportunities will depend upon incentivizing the habit of openness by ensuring that originators are credited in the ways described in section 4.2 and ensuring that data is FAIR (see section 4.4) in the repositories and cloud systems that increasingly hold them. Appropriate metadata are crucial in implementing policies that govern the arrangement, attribution, provenance, naming, description, representation, administration, access controls, retention, disposition, integrity and replication of digital objects.

4.4 Enabling digital data sharing and reuse

Digital data, whether ‘born-digital’ or digitized, are most valuable when they are ‘machine actionable’, allowing computational systems to find, access, interoperate and reuse data without human intervention. These are necessary attributes without which researchers would be unable to do justice to the volume, complexity and flux of much of the data increasingly available to them. As data are increasingly recognized as first-order publishable or published scientific outputs, whether or not associated with a published text article, rigorous data management is essential. The management of research data is usefully framed by the concept of the research data life-cycle involving a well-defined series of procedures (<https://www.reading.ac.uk/RES/rdm/about/res-rdm-lifecycle.aspx>; <https://www.nist.gov/programs-projects/research-data-framework-rdaf>) and illustrated in Figure 3.

¹⁸ The UK Economics and Social Research Council has stipulated that funding applicants who require specific data series must give evidence that the series does not exist before being funded to collect new data.

¹⁹ The International Union of Crystallography has issued a ringing challenge to the scientific community: ‘We urge the worldwide community of scientists, whether publicly or privately funded, always to have the starting goal to divulge fully all data collected or generated in experiments’ (Hackert et al., 2016).

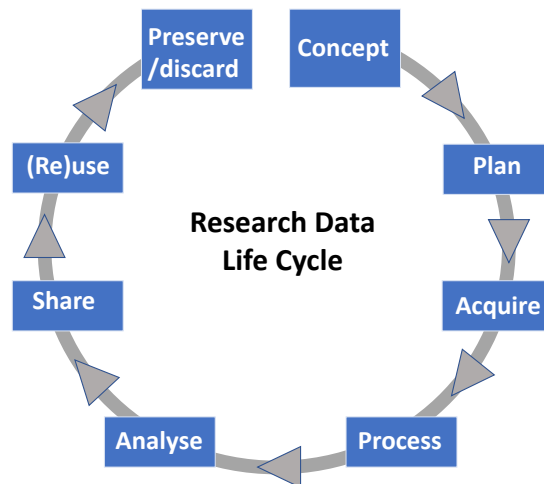


Figure 3. Schematic diagram illustrating the research data life cycle.

The costs of managing data through their life-cycle vary greatly depending upon their volume, complexity or the rate at which they stream through acquisition systems when real time analysis is involved. A 2020 report from the US National Academies of Sciences, Engineering and Medicine on life-cycle costs in biomedical science provides a conceptual framework for cost-effective decision-making about and management of the cycle. The cost of data stewardship should be regarded as a cost of doing science in the digital age rather than an optional add-on. For data-intensive science, the framework is increasingly organized at institutional, disciplinary/thematic, national or regional levels rather than at the level of individuals or groups. The ISC World Data System, for example, is building worldwide ‘communities of excellence’ for scientific data services by certifying member organizations – holders and providers of data or data products from wide-ranging fields – for their use of internationally recognized standards (<https://www.worlddatasystem.org>). It seeks to create the building blocks of a searchable common infrastructure, from which a data system that is both interoperable and distributed can be created.

If datasets are to be shared, reused and open to scrutiny by reviewers and other researchers, they need to be ‘intelligently open’ (Royal Society, 2012). The procedures that enable this have been formalized as FAIR – Findable–Accessible–Interoperable–Reusable (Wilkinson et al., 2016; <https://www.go-fair.org/fair-principles/>), as follows:

Findable: Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services. They should be assigned globally unique persistent identifiers, described with rich metadata, provided with the DOI of the data they describe and (meta)data should be registered/indexed in a searchable resource.

Accessible: The user needs to find the required data and how can they be accessed, possibly involving authentication and authorization. Data and metadata should be retrievable by their identifier using a standard communication protocol that is free and universally implementable and allows for authentication and authorization where necessary, and metadata should be accessible even when the data are no longer available.

Interoperable: Data usually need to be integrated with other data, and need to interoperate with applications or workflows for analysis, storage and processing. Data and metadata should be described and curated using a formal, accessible, shared and broadly applicable language for knowledge representation, and used with vocabularies that follow FAIR principles.

Reusable: The ultimate goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings. They should be richly described with plurality of accurate and relevant attributes, released with a clear and accessible data usage licence, associated with detailed provenance and meet domain-relevant community standards.

These are challenging requirements, but necessary if researchers and societal collaborators are to be able to share and reuse data and to combine diverse data series in ways that have the potential to reveal deep structure in the many inherently complex, interdisciplinary problems that science is called on to confront. Satisfying them is a vital step in exploiting the potential of the digital revolution and one that should be recognized and rewarded in the ways suggested in section 4.3. The use of FAIR priorities and procedures are being strongly promoted by GO-FAIR, a bottom-up, stakeholder-driven, self-governing consortium working to support researchers and institutions in implementing FAIR procedures (<https://www.go-fair.org/go-fair-initiative/>). An important challenge is to embed such procedures as a functionality of the research cycle (see Figure 4) in ways that are readily and routinely useable, creating added value to both the data originator and other potential users (e.g. <https://www.go-fair.org/events/international-fair-convergence-symposium/>).

There is increasing use of FAIR principles as necessary enabling tools, often building on established processes of data sharing. For example, particle physicists tend to share data within consortia attached to particular experiments. Some areas of the social sciences, particularly those concerned with longitudinal data, have a long history of data repositories, with a great deal of reuse (Dunning et al., 2017). For the disciplines that have successfully implemented FAIR principles, such data have become an essential part of their research infrastructures, widely used by the community in its daily research work. They include the ESFRI infrastructures in the humanities, DARIAH for arts and humanities, CEESDA in the social sciences and CLARIN, a language resource for humanities and social sciences.

There is as yet little general agreement about the principles of data availability and access that should be adopted by journals (PLoS ONE, 2019; Stall et al., 2019) and palpable reluctance from many authors to submit to the necessary disciplines. Journal editors and referees have a vital role in ensuring good practice for the link between data and publication. The editor-in-chief of *Molecular Brain* recently commented that amongst 180 manuscripts submitted since early 2017, 41 needed authors to provide raw data. Of these, 21 were withdrawn, indicating that requiring raw data drove away more than half of the manuscripts; and 19 out of the remaining 20 were rejected because of insufficient raw data (Miyakawa, 2020). Thus, 97% of the 41 manuscripts did not present the raw data supporting their results when requested by an editor, suggesting a possibility that the raw data did not exist from the beginning, at least in some of these cases.

5. Key issues: barriers to open access

5.1 Assessments, incentives and metrics

The researchers who create new knowledge and seek to communicate it are at the heart of scholarly publishing. For most, the primary motivation is the urge to discover and to communicate their discoveries. The way they do so is strongly conditioned, however, by the criteria that are used to assess scientific excellence by research funding bodies and institutional hiring and promotion boards, particularly critical issues for young researchers who are involved in the intense competition for tenured academic posts. Given the profusion of demands for assessment, the temptation for those tasked with it is to reach for a routinely accumulated proxy metric for excellence rather than an expert evaluation of a candidate's work. The danger is that a proxy, rather than a direct measure, can have unintended consequences and can be 'gamed'.

Publication-related metrics such as the number of an individual's publications, the number of citations, their *h*-index and the impact factors of the journals they publish in are the most common bases for assessment. None of them are necessarily good indices of the value of an individual's scientific contribution compared with an impartial judgement of an expert in the relevant field. Their primary value is one of time-saving convenience.

The journal impact factor (JIF) is the one of greatest immediate relevance to this report because of its influence on the publishing habits of researchers and on the publishing market (section 3.1). However, the majority of citations received by a journal tend to come from a small number of articles (Nature Editorial, 2005; Garfield, 2007; Sauermann and Haeussler, 2017), such that a given article in a high-impact journal is as at least as likely to be weakly as much as strongly cited, making the JIF an uncertain indicator of an individual's scientific contribution. The JIF was roundly criticized by the 2013 San Francisco Declaration on Research Assessment (DORA) with the trenchant general injunction not to 'use journal-based metrics, such as journal impact factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist's contributions, or in hiring, promotion or funding decisions'. DORA stresses the need to improve research by using more robust means of assessment that focus on the primary values of insight, impact, reliability and reusability, rather than on questionable proxies. DORA is addressed to funders, institutions, publishers, creators of these metrics and to researchers, and by early 2020 had been endorsed by 1,954 organizations and 15,943 individuals worldwide. It argues that the practice of using impact factors as an index of scientific excellence creates biases and inaccuracies when appraising scientific research, and that they should not be used as substitutes (Alberts, 2013). It is the quality of scientific outputs that need to be recognized, not a flawed proxy of journal status.

Proxy metrics such as the JIF suffer from the consequences of 'Goodhart's law' (Goodhart, 1981; Fire and Guestrin, 2019) that 'when a measure becomes a target, it ceases to be a good measure', primarily because it can be, and is 'gamed'. Exactly that has happened (Tuchman, 2012; Caon, 2017; Chapman et al., 2019), indicating an urgent need for new approaches, but with the warning that if these are also based primarily on proxy measures, they too are likely to become inappropriate targets.

It is important to recognize however that even the numbers of citations in themselves are relatively crude metrics of scientific value. They are known to include strong biases; they are biased by self-citation (Van Noorden and Singh Chawla, 2019), by author nationality (Campbell, 1990) and do not necessarily score highly for solidity/plausibility, originality and societal value (Aksnes et al., 2019).

The JIF has a major impact on the publishing market, as discussed in section 3.1, by pressurizing researchers and their institutions to target publication in high-JIF journals. It is in the commercial interests of such journals to maintain the influence of the JIF, which has resulted in pressures to manipulate author lists and the numbers and nature of citations, to selectively publish in areas with the greatest numbers of researchers, and to submit superfluous citations that refer to the journal in which authors wish to publish (Chapman et al., 2019). These activities serve to reinforce the brand, and thus the market power of the journal, rather than reflecting the real value of published research (Brembs et al., 2013). Maintenance of the JIF also discriminates against those who choose to publish in journals that are specific to their research theme, and those scholars who create high-quality outputs but are unable to afford the high costs associated with high-JIF journals. Moreover, editors of high-JIF journals tend not only to choose articles on the basis of perceived scientific rigour as assessed by peer reviewers, but also on their view of the importance of the research. They thereby play a disproportionate role in determining the priorities of science and in generating pressure on scientists to 'over-cook' their results to ensure publication in these journals (Tijdink et al., 2016).

The over-use of a narrow evaluation metric such as the JIF also has a significant impact on researchers, universities and the research system. Given the hyper-competition for university posts, it places a relentless focus on individual achievement and delivery through the JIF route, it thins out research support through a university's interest in high-impact metrics, and places enormous time pressures on everyone, with a vicious circle of coercion on all to tick boxes and to conform. In some settings, where public funding for universities is directly linked to publication counts, researchers can be encouraged to sub-divide publications into many small units of publication.

This complex of issues is increasingly being addressed (e.g. Nationaal Platform Open Science, 2018; Bregman, 2020; FOLEC-CLACSO, 2020) and is the subject of an upcoming ISC study. One approach is to base evaluations on a narrative resumé that asks what has been contributed to the generation of knowledge, to the development of individuals, to the wider research community and to broader society (see Royal Society, Resumé for Researchers: <https://royalsociety.org/topics-policy/projects/research-culture/tools-for-support/resume-for-researchers/>). It is a more balanced approach than that offered by the JIF, though it lacks the easy convenience of a proxy metric. DORA has now been built on by the later Leiden manifesto (Hicks et al., 2015) and Jussieu Call for Open Science and Bibliodiversity (2017), that add weight to discussions of research assessment. Whereas DORA is STEM-focussed, the Leiden and Jussieu declarations are more sensitive to a broader range of disciplines, including the humanities and social sciences. The Leiden Declaration for example asserts the following:

- a) Quantitative evaluation should support qualitative, expert assessment.
- b) Measure performance by the research missions of the institution/group/researcher.
- c) Protect excellence in locally relevant research.
- d) Keep data collection and analytical processes open, transparent and simple.
- e) Allow those evaluated to verify data and analysis.
- f) Account for variation by field in publication and citation practices.
- g) Base assessment of individual researchers on a qualitative judgement of their portfolio.
- h) Avoid misplaced concreteness and false precision.
- i) Recognize the systemic effects of assessment and indicators.
- j) Scrutinize indicators regularly and update them.

It should be noted that principle (c) is particularly critical for priorities such as the those of the Sustainable Development Goals (SDGs), where location is an essential boundary condition for understanding.

5.2 Peer review issues

Peer review makes a fundamental contribution to the record of science. For academics, it is, or should be, part and parcel of their contribution to science, not a charitable activity that they undertake in their spare time. Peer review takes two forms: formal review for conventional journal articles and books at the behest of a publisher, and informal, unsolicited, expert post-publication or post-preprint scrutiny of a novel truth claim, which often uncovers major errors that have been missed by formal peer review. It plays a number of roles: in determining the importance, quality and novelty of a publication; in identifying unwarranted claims, demonstrably fallacious interpretations or lack of originality; and in suggesting improvements in comprehensibility or logic. Some reviews are rigorous and an asset to both author and journal editor, whether acceptance, rejection or revision are the outcomes. Some are trivial and of little value to author or editor, and some are improperly biased, whether recommending acceptance or rejection (Smith, 2006). Where methods of analysis or data treatment are complex, the task of rigorous review can be daunting.

Peer review has been a gatekeeper to the record of science, and has come to represent a quality threshold in the public eye, such that ‘unreviewed’ papers are greeted with suspicion. Although most scientists regard peer review as an essential part of the publication process, they know more than most that it is not an infallible process and that its failures can have serious consequences. In 1998 a paper submitted to the *Lancet* (Wakefield et al., 1998) claimed to show that the measles, mumps and rubella (MMR) vaccine predisposed children to autistic behaviour. Despite the small sample size ($n = 12$), the uncontrolled design and the speculative nature of the conclusions (DeStefano and Chen, 1999) the paper received wide publicity, and MMR vaccination rates began to drop because parents were concerned about the risk of autism after vaccination. Whilst retraction has deleted the paper from the scientific record, it has not removed it from humanity’s wider store of misinformation, and its echoes continue to undermine public health.

Whilst journal editors may have their own view of the importance of the work that they publish, the true test of significance comes from the longer-term, post-publication peer response of the scientific community, which is the ultimate bedrock on which the self-correcting property of science rests. Claims that are regarded as important tend to be subject to rigorous scrutiny and testing, resulting in invalidation, provisional acceptance or continuing debate. Many published claims are clearly not regarded as sufficiently important to justify such attention, and although some become significant at a later date, many are quietly forgotten, having failed in their quest for citation²⁰. A principle of Laplace, reworded by Carl Sagan, is relevant, that ‘extraordinary claims require extraordinary evidence’ (Gillispie et al., 1999), to which we might add that significant claims require significant scrutiny.

The very value of prepublication peer review has been strongly questioned (e.g. Smith, 2006; Sullivan, 2018; Heesen and Bright, 2020), with the comment by Smith (2006), a former chief executive of the BMJ Publishing Group, that it ‘is a flawed process, full of easily identified defects with little evidence that it works’ ... but that ‘it is likely to remain central to science and journals because there is no obvious alternative, and scientists and editors have a continuing belief in peer review. How odd that science should be rooted in belief’. It is a view that is strongly contested²¹, but given the significance of peer review for the continuing evolution of preprint servers and other innovative publishing systems, a systematic, evidence-based debate is needed if we are to understand how peer review, both pre- and post-publication, can best contribute to the rigour and creativity of the processes of scientific communication.

These issues have been highlighted during the COVID-19 pandemic, with implications for other potential emergencies. The long publication lag-time associated with prepublication-review journals was inconsistent with the urgent need for access to relevant research, whether that awaiting conventional publication, or research that was not yet ready for publication but which contained potentially valuable evidence. It led to large audiences, including policy-makers and the public, devouring preprint servers for access to the latest evidence. The biomedical servers medRxiv and bioRxiv were able to make preprints available within a day or two of submission, so that these two sites alone now (December 2020) host almost 7,000 COVID-relevant papers, which have been downloaded millions of times throughout the world (Barbour, 2020). However, preprints immediately relevant to crisis management have the potential to impede or derail management if they are flawed or misconceived, creating calls for rapid preprint review processes. The challenge to the international science community is to create a response that is rigorous, inclusive, responsive to diverse national capacities and needs, and of global scale. It has been suggested that the need for an action-ready

²⁰ Across all disciplines, and ignoring self-citation, 18% of papers remain uncited after 10 years (Lowe, 2018), although the rate appears to be diminishing as the habit of profuse citation becomes increasingly common.

²¹ See, for example, <https://www.publisso.de/en/advice/publishing-advice-faqs/peer-review/>

system might be explored through urgent dialogue with international scientific bodies, national academies and national research councils, convened by bodies such as the ISC and UNESCO (Rovenskaya et al., 2020). As the impacts of climate and environmental change become more severe, the demand for such rapid and open access to relevant research is likely to increase.

Whilst retractions of scientific journal articles are relatively rare (Brainard, 2018), a major challenge to peer review in the last decade has been a demonstration that the results of many papers in highly regarded journals have proved not to be reproducible (Miyakawa, 2020), a fact that had not been picked up by their peer reviewers. In many cases, this arose because the evidential data and associated metadata were not provided, and could not be accessed by reviewers. A more rigorous approach (see section 4.1) that requires data to be concurrently available for scrutiny under FAIR principles is essential. Even then, however, establishing whether the inferences drawn from large and complex datasets are statistically robust is an onerous task, even if the reviewer has the technical expertise to do so. Making use of innovative technologies, and potentially diversifying the role of editorial boards to include specialists to advise on data review could help create the additional resources needed to assess evidential data during the peer review process, though it would impact on costs (see section 5.5). There are methods of identifying fraud (Bolland et al., 2016), and of exploring the validity of data series that could be routinely applied to data-intensive work submitted for publication as essential parts of the armoury of publishers. Machine learning (ML) algorithms are also increasingly used in data analysis across a broad range of disciplines, suggesting the need for new norms about issues such as what should be disclosed in formal publications when using ML, how should algorithms be reported and should it be mandatory to make codes open, or at least available to peer reviewers? Disciplines, professional societies and the scientific unions have important roles to play here since disciplinary practice and culture will be important determinants of the way that these issues are resolved (Beam et al., 2020).

As the numbers of scholarly articles have increased, from 2.5 to 4.5 million per year in the last decade (Box 2), it has become increasingly difficult to obtain the services of appropriately expert reviewers (Gropp et al., 2017), in particular for multidisciplinary publications. Although the responsibility of scientists to undertake the non-trivial task of review is stressed at the head of this section, it is important to examine the extent to which incentives may be needed to maintain the review effort, and whether parts of the task could be automated. The idea of payment for reviews is regularly raised (e.g. <https://www.enago.com/academy/should-a-peer-reviewer-be-paid/>), but such an approach may be disruptive, particularly in undermining the business models of many open access publishers. It could create perverse incentives that work against trust in the review system, and perhaps undermine ‘reciprocity’ in peer review, whereby scholars review and expect to be reviewed. It is also a common experience that replacing personal responsibility to undertake voluntary but socially important tasks by payment can reduce rather than increase productivity (Carney, 2020). An alternative to providing material incentives is to reward reviewers in ways that recognize their contribution to science and scholarship. One such approach to giving credit for review work, and to increase fairness in review processes, is that of ‘open peer review’²², through which reviews could be seen as a form of publication that could be used in assessing researchers’ contributions (Rovenskaya et al., 2020). While there has been a steady growth in open peer review in past decades (Wolfram et al., 2020), the arguments for and against continue. Reviewers may not be willing to review papers from sources they feel could compromise their careers or relationships if their identity were known, whilst early-career researchers and others in subsidiary positions may feel inhibited from reviewing the work of those that have some authority over them (Smith, 1999).

²² There is no standard definition of open peer review, but it used here to mean peer review in which reviewers’ names and comments are published alongside the paper concerned.

There is increasing potential to relieve some of the pressures of formal peer review through the use of automated systems for aspects of editing and review, such as identifying image manipulation, validating data series, using ML training sets to clarify text and using networked distributed systems for new methods of open review. The experience of the *Journal of Open Source Software* (JOSS) which enables authors of research software to receive citation credit for their work is instructive in this regard (JOSS, 2019). JOSS has implemented an open review process via GitHub, and makes extensive use of automated tools to support the publication process, thereby saving costs. These systems are valuable in demonstrating the use of cost-saving automated systems in the publication management and review process for many smaller scientific publishers and international scientific organizations that operate on low margins.

5.3 Copyright and patents

Discussions of the copyright to published scientific work are almost invariably concerned with the rights and interests of authors and their institutions on the one hand and of publishers on the other (Tennant, 2019). Although the careers of researchers, the reputations of universities and the profits of publishers are important considerations for an efficient and effective publishing system, it is important to recognize that its ultimate purpose is to serve the public interest through the creation of new knowledge, which forms the rationale for the activities of most public and charitable funders of research. It is argued here that, with defined exceptions, the public interest is best served when the results of scientific inquiry are released into the public domain with, in the words of the Berlin Declaration, a 'free, irrevocable, worldwide, right of access to, and a licence to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship'.

Traditional scholarly publishing has often involved transfer of copyright from author, the originator, to publisher, the agent, which sets restrictions on how an author can share and reuse their work. This may inhibit interrogation of the record of science, such as in the inhibition of text and data mining, and also permits publishers to monetize the publications for which they have received copyright. Authors are typically the weaker negotiating partner in discussions with large publishers, as few non-specialists have a detailed understanding of copyright law, and copyright agreements tend to be established at the end of a publication process, by which time many authors are keen to expediate completion. Copyright law as it pertains to publishing is complex, poorly understood and frequently abused. These tensions are reflected in the popularity of sites such as ResearchGate and Sci-Hub for illicit file sharing by academics and the wider public (Lawson, 2017). It is perhaps unreasonable to expect researchers and librarians to become expert in copyright law, but research institutions and networks of researchers have a role to play in providing access to information on the different licences available and their implications. It has been argued above that the transfer of copyright fails the public interest test, with increasing calls for authors to recognize that unrestricted sharing helps to advance science faster than paywalled articles, that copyright transfer does a fundamental disservice to the research enterprise (Vessuri et al., 2013), and that it should be regarded as an unethical practice (Biasi and Moser, 2018). We should move beyond the proprietary commodification of scholarly knowledge towards an open communication system that is fit for purpose.

The tension between open access and paywalled copyright has been illustrated in the COVID-19 pandemic, when scientific authorities from twelve countries, including the US, Italy and South Korea, urged corporate publishers to make their papers relevant to COVID-19 openly and promptly available: '[we] urge publishers to voluntarily agree to make their COVID-19 and coronavirus-related publications, and the available data supporting them, immediately accessible' (Office of the Chief Science Advisor, Government of Canada, 2020). A petition with 2,000 signatures by 3 March 2020

stated: ‘Thousands of scientific studies about the coronavirus are locked behind subscription paywalls, blocking scientists from getting access to research needed to discover antiviral treatments and a vaccine to stop the virus’ (Napack, 2020). There was a valuable but time-limited response from publishers, though the papers in question remain under their original terms of copyright.

By contrast, open-access publishing typically permits authors to retain copyright whilst transferring non-exclusive rights to the publisher. A ‘Creative Commons’ (CC) licence (<https://creativecommons.org/licenses/>) is the most used form of licence where the intention is to grant permission for others to use an author’s creative work (Box 7). CC licences now comprise a vast and growing digital commons, with content that can be copied, distributed, edited, remixed and built upon within the bounds of copyright law. They were initially released in 2002, and are the most favoured means of observing the strictures on copyright in the Berlin Declaration. The preferred form of CC licence varies across different fields of study although CC BY is probably the most commonly used. In many individual disciplines of the humanities and social sciences, the integrity of argument often depends on careful and precisely phrased formulations (e.g. in philosophy or law), which are not protected by CC BY, and would be at risk. The integrity of orally-gathered material (some of which is of great ethical sensitivity, for example first-person accounts by people with mental health issues,

Box 7: Creative Commons licences favoured for open access

(from: <https://creativecommons.org>)

ATTRIBUTION: CC BY

This licence lets others distribute, remix, adapt and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

ATTRIBUTION-SHAREALIKE: CC BY-SA

This licence lets others remix, adapt and build upon your work even for commercial purposes, as long as they credit you and license their new creations under identical terms. This licence is often compared to ‘copyleft’ free and open-source software licenses. All new works based on yours will carry the same licence, so any derivatives will also allow commercial use. This is the licence used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.

ATTRIBUTION-NODERIVS: CC BY-ND

This licence lets others reuse the work for any purpose, including commercially; however, it cannot be shared with others in adapted form, and credit must be provided to you.

ATTRIBUTION-NONCOMMERCIAL: CC BY-NC

This licence lets others remix, adapt and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they do not have to license their derivative works on the same terms.

ATTRIBUTION-NONCOMMERCIAL-SHAREALIKE: CC BY-NC-SA

This licence lets others remix, adapt and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms.

ATTRIBUTION-NONCOMMERCIAL-NODERIVS: CC BY-NC-ND

This licence is the most restrictive of our six main licences, only allowing others to download your works and share them with others as long as they credit you, but they cannot change them in any way or use them commercially.

CC0

This licence enables scientists, educators, artists and other and other owners of copyright- or database-protected content to waive those interests in their works and thereby to place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

survivors of trauma or refugees) could be prejudiced by CC BY, risking changes to testimonies which would breach ethical guidelines. In such cases, the use of CC BY-ND licence would protect text from potential distortion by subsequent users. Authors choosing CC licences need to be aware of the implications of each, especially with regards to selecting 'NC' if they do not want their work to be subsequently used commercially.

With regards to the publication of software, GNU General Public Licenses (GNU-GPL) are amongst the most widespread of those that guarantee users the freedom to run, study, share and modify software. The GPL series are 'copyleft' (as opposed to copyright) licences that require any derivatives to be distributed under the same or equivalent licence terms as the original. GPL has been crucial to the success of Linux-based systems, giving the programmers who contribute to the kernel an assurance that their work will benefit all and remain free, rather than being exploited by software companies that are not be required to return any value to the community (Wheeler, 2006).

In the commercial environment, patenting is analogous to copyright in protecting the rights of originators. Recent decades have seen increased commercialization of publicly funded research results, with universities and research institutes being encouraged to collaborate with private companies on research projects in ways that have blurred the boundaries between scientific and commercial research. In this context, it is important to recognize that just as copyright only protects the expression of original ideas and not the research finding as such, patenting in a commercializing environment protects the rights of an originator whilst enabling the discovery to be publicized. Patents grant exclusive rights that prevent third parties from commercially exploiting, making, using, offering for sale, selling or importing the invention, for a limited period of time (generally 20 years). In return for this monopoly, the patent owner is required to disclose technical information about the invention in order for others to access it and to use it in further innovations. As a consequence, publicly accessible patent libraries can be a rich source of information about innovation, and should be more explored by academic researchers (Pereira et al., 2015). In addition, it is increasingly the case that commercial researchers publish their findings in scientific journals, sometimes for the purposes of 'defensive publication'²³. A further consequence of the rebalancing of priorities in universities and institutes towards commercialization is often to delay scientific publication until the patent application has been published, as to do otherwise would invalidate a patent application.

5.4 Indexing

Indexes of published scientific work are of great importance in signposting the existence of relevant scientific work to those who seek to find what is known in a given field of knowledge. The bodies that compile indexes of science publications such as Google Scholar, Web of Science (owned by Clarivate Analytics, formerly part of Thomson Reuters) and Scopus (owned by Elsevier) have, in essence, the indirect power to define scholarship and act as gatekeepers to research findings.

Such indexes tend to favour a circumscribed group of journals, often those from major publishers and mainly published in Europe and North America. They are dominated by English-language publications, with very little visibility of work in other languages. In addition, the research outputs of developing regions are poorly represented, stimulating the creation of a number of region-specific open-access publishing infrastructures that give visibility and access to the content of their high-quality journals (Vessuri et al., 2013). They include African Journals Online, Journals Online collections in several countries (Dunning et al., 2017) and SciELO and Redalyc in Latin America. Much would be gained if

²³ A defensive publication is an intellectual property strategy used to prevent another party from obtaining a patent on a product, apparatus or method.

these systems could be federated to create a more representative and inclusive view of the international scientific effort.

5.5 Costs and pricing

Pricing is fundamental to the future of publishing, to the realization of the principles set out in this report and in determining access by authors and readers. High prices, whether charged to readers or authors, can be insuperable barriers to access, particularly for those in poorly-funded institutions and in low- and middle-income countries, and fundamental barriers to open science. Prices should not be so great as to deny access to authors or readers.

It is important to differentiate between cost and price. Cost is dictated by the processes, materials and human resources needed in production. Price is determined by market structure, including the relationship between supply and demand and extent to which producers and consumers are informed about market conditions. In the case of scholarly journals, the market is built around incentives for publication, impact factors and production volumes that enable economies of scale. Markets are most efficient in allocating resources when all relevant information is made available to market players, and when assets are neither over- or under-valued. ‘Big deals’ (Box 2) are both barriers to market entry by new publishers (European Commission, 2006) and barriers to transparency, in ways that militate against market efficiency and have allowed major commercial publishers to charge high prices and generate excessive profits.

The cOAlition S consortium attempts to address this issue by calling for full transparency and monitoring of publication costs and fees as a basis for a more efficient market²⁴. A minimum level of compliance would involve publishers sharing information on the costs of services such as organization of peer review at the publisher level, but also to do so on a journal-by-journal basis (Wallace, 2020). However, some fear that Plan S will merely perpetuate and strengthen the positions of large for-profit publishers, who already dominate the market, thereby continuing to discriminate against authors in poorly-funded institutions or science systems, and limit competitiveness by preventing or discouraging innovation and the emergence of new players and new models, including non-commercial open access models (Aguado-López and Becerril-García, 2019).

Cost comparisons have proved difficult to make, although it is clear that for journals the shift from analogue to digital publishing has fundamentally changed costing structures. At their most basic, the core functions of publishers – organizing peer review processes, editorial oversight, making accept/reject decisions – have remained the same. It is the processes of dissemination and after-sales maintenance that have changed, to which we would add the duty to ensure that relevant data is concurrently available in FAIR format. Analogue processes of type-setting, printing, binding and physical distribution have been replaced by electronic formatting and online dissemination. Once printed and distributed, the analogue form is preserved on library shelves. The digital form however requires ongoing electronic storage, with significant maintenance and updating costs as technologies and formats change (Anderson, 2014). The cost comparison is therefore a complex one.

Publishers’ production costs also depend on the extent of ancillary activities designed to support the submission and publication process, including editing and marketing. In addition to a number of fixed costs (e.g. for web hosting or financial services), there are many factors that affect the publisher’s variable costs, such as whether editors receive payment, how many submissions the journal receives and how many submissions are rejected. Peer review is a major un-costed element of the publication

²⁴ We are sceptical that many commercial companies will reveal this data voluntarily, and if it is revealed, it is unlikely to be verifiable. Competition authorities, with judicial power to demand data, might be successful, but even then with difficulty.

process. The magnitude of the global peer review effort was estimated, in 2008, to have a monetary value of £1.9 billion (\$2.48 billion) (Hide, 2008). Given the estimated global value of the science journals market in 2020 of about \$10 billion per annum (Global Scientific and Technical Publishing, 2019–2023), free peer review represents a massive contribution by the scientific community to the publishing enterprise, including to the high profit margins of major commercial publishers.

In recent years, a number of publishers have exposed summaries of their costs, a helpful step towards transparency and to assessment of the value to science of different publishing models, and which it is hoped that the efforts of cOAlition S will clarify further. Several examples follow:

- The European Molecular Biology Organisation (EMBO) has published its cost breakdown, stating that it would cost €9,040 per published research article in 2019, and therefore that its current APCs of between €3,300 and €4,700 would not be sufficient to cover raw costs, which therefore must be subsidized from subscriptions. EMBO note that the cost of €9,040 per research article are also used to cover the costs of opinion pieces, editorials and news articles they publish, which usually do not generate income.
- The *Journal of Open Source Software* (JOSS; <https://joss.theoj.org/papers/published>) is an academic journal (ISSN 2475-9066) with a formal peer review process that publishes 300 papers per year, for which ‘camera ready’ submissions are required from authors. It is supported by a grant from the Alfred P. Sloan Foundation. The estimated annual operating and upgrading costs were the journal to receive no subsidy are \$31,400, which, for example would require an APC of about \$100 per paper. The initial set-up costs would have been about \$50,000. As is the norm for both not-for-profit and for-profit publishers, no fee is paid to its volunteer editors and peer reviewers. If editorial salaries were paid at a conventional rate, the required APC would be about \$1,300 (JOSS, 2019).
- The Multidisciplinary Digital Publishing Institute (MDPI), an open-access publisher, has published its costing model in order to comply with Plan S. MDPI charges APCs ranging within 1000–2300 Swiss Francs (~€950–2100) across its portfolio. They have in-house editors, but with rejection rates that are not excessively high, and a rapid peer review model that has drawn admirers.

The JOSS example highlights the benefits of building journal publishing or repositories in pre-existing hardware and software infrastructures that have capacities for perennial upgrading, such as those of universities. They represent an obvious opportunity to support digital publishing and repository functions in ways that are highly cost-efficient.

Grossmann and Brembs (2021) have presented a systematic, quantitative analysis to determine the actual costs of efficiently publishing a scholarly article using state-of-the-art technologies. They conclude that costs range from less than US\$200 per article in modern, large-scale publishing platforms using post-publication peer review, to about US\$1,000 per article in prestigious journals with rejection rates exceeding 90%. The latter contrasts dramatically with the prices actually charged for prestigious journals, which tend to be greater by an order of magnitude, and which continue to rise even in their APC-based open access journals.

A variety of models that do not depend on APCs are being developed across the world. For example, the Open Library of the Humanities, or PLoS’s Community Action Publishing (CAP) both rely on institutions acting altruistically to support publishing within their community. Other schemes such as ‘Subscribe to Open’, which provides open access to a year’s worth of content based on continued subscription, are intended to support a gradual shift to open access. Like flat-fee schemes, such as the

one in place at PeerJ, they rely on institutions and authors being willing to speculate that researchers will want to continue to read and publish in a certain journal over a long period of time.

It is clear that prices for high-standard publication can be very much lower, by an order of magnitude or more, than those charged by high-JIF commercial journals. If the processes of digital disruption and innovation that have characterized other business sectors are able to penetrate the current scientific publishing system at scale, there is a prospect for a substantial reduction in overall journal pricing, much to the benefit of science. One such route could be through more widespread development of overlay journals (see section 3.7); for example, in 2019, arXiv hosted over 1.6 million e-prints and received 155,866 submissions, with total costs of \$2,684,111 (arXiv, 2020), a cost of \$17.22 per submission. Scholastica provides an overlay hosting service for journals at \$99 per month plus \$10 per paper (Coles, 2019). Scholastica provides this service for *Discrete Analysis*, a new overlay journal for mathematical content in arXiv, and with no charges for readers or authors (Gowers, 2015).

6. A changing world of science and its implications for scientific publishing

The practices and priorities of science inevitably adapt to the needs of the times, to advances in scientific understanding and to developments in technology that permit new paths to discovery. The processes of scientific and scholarly publishing also need to adapt to these trends if they are to continue to serve the needs of science.

6.1 Trends in science

There are major trends of demand, opportunity and challenge that are an important context for developments in scientific publishing:

- a) Global R & D funding has grown significantly in recent decades, doubling from less than \$1 trillion in 2005 to more than \$2 trillion in 2015, of which 75% is spent in just ten countries (Timmer, 2018). At the same time there has been increased support for and commitment to research in low- and middle-income countries, although much of this is attributed to China (Soete and Schneegans, 2015). The COVID-19 pandemic has, in many countries, illustrated the role of science as an essential component of national intellectual infrastructure for modern societies in confronting the many challenges that they face. A zoonotic disease like COVID-19 is just one of the global consequences that arise from the human assault on the systems of the planet, and of which global society is becoming increasingly aware. The local and the global inter-penetrate. Local events influence and are influenced by global events, such that there is growing awareness that the wellbeing of states and their citizens is dependent on solutions that only international collaboration can secure (Roosendaal and Geurts, 1997). We cannot escape the reality of an interconnected world. Viruses and climate change do not carry passports, global solutions require global collaboration, and scientific research and knowledge are essential in seeking them. In this context, the role of scientific publication systems is to maximize the efficiency (ubiquity and speed) with which scientific ideas, findings and information disseminate and are accessible globally.
- b) In the year 2000, Stephen Hawking, the cosmologist, predicted that ‘the next century (the 21st) will be the century of complexity’ (Gorban and Yablonsky, 2013). The era of ‘big data’, with rich and varied data streaming through acquisition and computation devices, coupled with artificial intelligence (AI) technologies, are making that prediction a reality. Complexity has become fundamental to modern science. It is inherent to many of the major challenges faced by science and humanity, from infectious disease, to the behaviour and functions of cities, to the deep structure of the universe, to national and global economies, to climate

change and global sustainability. Forward computational modelling is able to explore the dynamics of complex systems. The AI and ML technologies are able to identify their deep patterns and relationships in powerful and unprecedented ways, as illustrated by the recent giant leap in solving one of biology's grandest challenges, that of determining a protein's 3D shape from its amino-acid sequence, with profound implications for life science and medicine. The inductive power of AI and ML procedures in characterizing deep patterns in highly specific data-rich systems is a novel route to discovery, representing a major paradigm shift in science. It contrasts with the hitherto dominant paradigm that seeks more general, theory-enabled solutions. It involves an epistemic change in the way that the record of science is aggregated and what it means to 'publish' in the 21st century.

- c) The launch of the World Wide Web, 30 years ago, ushered in a new world of information, accessible to all who had devices to access it, both as readers and authors. The internet, the World Wide Web and social media have created a new global information and communication environment that has revolutionized social communication, and provided means of accessing and contributing to the distribution of information and knowledge in unprecedented ways. But they have also created a digital divide (International Telecommunications Union, 2019), with approximately 3.5 billion people globally lacking the means of access to what has become a fundamental means of human communication. Those that do have access have become accustomed to knowledge resources being 'free', ultimately funded from the information about themselves that they routinely give up as a marketable commodity to the technology platforms that host knowledge resources. These new social dynamics elude the traditional 'gatekeepers' of authorized wisdom. National media, newspapers and journals are no longer able to act as high-level filters for public information. The new technologies have the potential to enhance the reach of science, but they also create platforms for disinformation, enabling lobby groups to undermine scientific consensus on many critical issues such as climate change, vaccination, smoking and AIDS. In this setting, much of the record of science, though largely funded by citizens' taxes, lies on the wrong side of a paywall that they are less able, and less inclined, to climb. Effectively engaging with citizens and combating disinformation is one of the great challenges for modern science and its publication and dissemination systems.
- d) Much contemporary rhetoric speaks of a crisis of trust and rise of prejudice around the world. As the introduction to a 2019 Ipsos MORI report states, 'deference is dead and everywhere the elites and mainstream media are challenged by an angry populace'. However, the evidence then presented in the report and replicated elsewhere suggests otherwise (Ipsos MORI, 2019; Carter, 2020). Scientists are amongst the most highly trusted professionals in all countries that have been surveyed, and in many, trust is on the rise. Many societies have become politically more polarized, but this does not readily translate into trust/distrust of scientists or experts. The experts and elites are unjustifiably anxious. Trustworthiness is the essential issue: does the body do what it is supposed to do, do it well and is it transparent, with high ethical standards? These latter features are perceived to apply to public sector scientists more than those in the private sector, partly because of a perception of greater disinterestedness and of the norms to which public sector scientists are presumed to adhere, of transparency, replication, challenge and self-correction. Trust is a problem, and although it is a chronic rather than acute one, it is one where science must continue to justify high levels of trust, with trust in the means by which scientific understanding is published and disseminated being a key priority. The experience of science during the COVID-19 pandemic should add considerably to understanding how the trustworthiness of science and its communication can best be maintained.

These trends underline the fundamental need for scientific publishing to develop in ways that:

Facilitate global cooperation and ensure that the richness of diverse experiences and perspectives are drawn on in developing global solutions.

The current state of scientific publishing is one in which the Global North is more or less well-served, though at an unnecessarily high cost that discriminates against readers in the reader-pays publishing model and authors in the author-pays model. These disadvantages are exacerbated for low- and middle-income countries, mostly in the Global South. The consequences of these contrasts are that the highest profile journals, almost all of them published in Europe or North America, tend to carry the science of the Global North, costly to access in the South, whilst that of the Global South tends to circulate in southern regions only. Such publishing apartheid is a barrier to the flow of ideas and undermines the potential of the international science community to seek and create global solutions to global problems. This is not to underestimate the very positive benefit that locally circulated journals have in their regions.

Create ready access to the record of science and its data to enable deeper inquiry into the complexity that lies at the heart of many contemporary problems.

The donation of copyright to publishers of the results of publicly funded research is both a highly questionable practice and one that militates against the above objectives. It should cease, as advocated by the Berlin Declaration (2003). There is a strong case for national funders to intervene by forbidding this practice among those whose research they fund as a means of protecting a public asset. In addition, much of the archive of science lies in publishers' vaults, protected by high paywalls, and largely inaccessible to the techniques of text and data mining. The time-limited response of commercial publishers to demands for open access to these and other resources should be general, permanent and immediate.

Open access to the record of science to citizens, particular in areas of contemporary public concern.

The public desire for scientific information on contemporary matters has been exemplified during the COVID-19 pandemic (see section 5.3). The scientific community should consider creative ways in which efficient and effective bridges between highly technical scientific literature and the public can be created. It is instructive to recall that the mission statement of the journal *Nature*, published in its early years, stated the intentions: 'Firstly to place before the general public the grand results of scientific work and scientific discovery, and to urge the claims of science to a more general recognition in education and in daily life ...' and only 'secondly, to aid scientific men (sic) themselves, by giving early information of all advances made in any branch of natural knowledge throughout the world'. There are many creative attempts to refresh these principles in the era of the World Wide Web, social media and 'fake news' which need strong support from the scientific community.

6.2 Open science

The open science movement has arisen from the community's grassroots. Whilst openness has been at the core of scientific enquiry since the publication of the first scientific journals in the 17th century, the developments summarized in section 6.1 have stimulated new horizons of openness (Titz, 2016). The Budapest Declaration of 2002 on open-access publishing (section 3.3) could be regarded as its starting point, although shifts in the habits of scientific inquiry from one based on the hegemony of disciplinary science (mode 1), to more socially and thematically distributed systems of knowledge production, more responsive to societal needs and subject to multiple accountabilities (mode 2) (Gibbons et al., 1995; Nowotny et al., 2003) had been an ongoing development. In subsequent years, a number of influential reports made the case for open data, for opening the processes and

infrastructure of science²⁵, and for a more proactive, trans-disciplinary openness to society (Royal Society, 2012; Science International, 2015; National Academies of Sciences, Engineering and Medicine, 2018). The issue has now been taken up by UNESCO as the theme of a possible formal international recommendation (UNESCO, 2020), which it is hoped will be adopted by its 193 national members.

The benefits of a new era of open science are seen from a variety of perspectives (Bowman and Keene, 2018). Some advocate open science and greater sharing of data and information as means of increasing the efficiency of scientific inquiry²⁶. Some see benefits to interdisciplinary science in having open access to the record of science and to a wide variety of its data streams. Some see access to and integration of diverse, multi-dimensional data streams as means of analysing inherently complex problems. And some see open science as a democratizing process, in which openness is socially contextualized (Chan et al., 2019). The ISC takes a broad view (2020) that encompasses all these motivations, in defining open science:

Science that is open to scrutiny and challenge, and to the knowledge needs and interests of wider publics. Open science makes the record of science, its evolving stock of knowledge, ideas and possibilities accessible to all, irrespective of geography, gender, ethnicity or socio-economic circumstance. It makes the data and evidence of science accessible and reusable by all, subject to constraints of safety, security and privacy. It is open to engagement with other societal actors in the common pursuit of new knowledge, and to support humanity in achieving sustainable and equitable life on planet Earth.

This perspective provides the context and rationale for this report, for without fundamental and equitable improvements in access to the record of science, its data and evidence, by authors and readers, whether scientists, citizens or institutions, the vision inherent in this definition is likely to remain beyond our grasp. The extent to which modern publication mechanisms and processes enable or impede the development of open science is therefore a crucial issue.

The immense potential of open science to benefit humanity has been powerfully demonstrated during the COVID-19 pandemic by the response of scientists worldwide in sharing data, in creating working vaccines, in advising governments and informing populations. On 5 January 2020, a team led by Professor Zhang Yong-Zhen at Fudan University in Shanghai sequenced the genetic code of the virus that caused the initial outbreak in Wuhan, China. On 10 January, an Australian colleague of Zhang published the sequence on the website *virological.org*, which was the beginning of an unprecedented global scientific effort to treat and ultimately vaccinate against COVID-19 (Holmes E., Initial genome release of novel coronavirus 2020)²⁷. The Director of the US National Institutes of Health commented: 'I have never seen anything like this'... 'The phenomenal effort will change science – and scientists – for ever' (Sample, 2020). It is vital that science, governments and societies learn the lesson that we live in an interconnected world that needs an open, interconnected science if we are to address the many challenges that face global society, and that open science is properly incentivized (National Academies of Sciences, Engineering and Medicine, 2020b). The lessons for scientific publishing are clear: that the record of science must be openly and readily accessible, that preprints are of vital

²⁵ See <https://www.fosteropenscience.eu/foster-taxonomy/open-science-definition>, accessed 9 November 2020.

²⁶ The publicly funded virtual library of the NASA Astrophysics Data System (ADS) allows intelligent full-text searching across the entire corpus of astronomical literature. It is calculated (<https://arxiv.org/abs/astro-ph/0002104>) that 'the ADS increases the efficiency of astronomical research by 333 Full Time Equivalent (FTE) (2000 hour) research years per year, and that the value of the early development of the ADS for astronomy, compared with waiting for mature technologies to be adopted, is 2332 FTE research years'.

²⁷ Available from <http://virological.org/t/initial-genome-release-of-novel-coronavirus/319> (Accessed 3 February 2021).

importance during emergencies and that the development of agile and rapidly mobilized peer review systems is a priority.

6.3 A critique from the Global South

Not all are proponents of open science, or of the positive view of it expressed above. There is a developing critique from the Global South, particularly from Africa, that the assumptions and processes of open science and open access publishing as they have developed serve to reassert neo-colonial values in ways originally framed in the work of Franz Fanon (1961/2002). For many schooled in the confident setting of western science, 'the claim that open access may be a neo-colonial process seems incomprehensible' (Piron et al., 2017), after all, is not science universal? The latter argument must be carefully nuanced. The laws of physics may be universal, but social customs and characteristics of population health are not. Equally, there are epistemological diversities that reflect differing histories and values that lead to differing priorities and approaches. There is bias in the record of science as represented by indexes such as Web of Science and Scopus which, as noted in section 5.4, are dominated by outputs from the major commercial publishers, all located in Europe or North America, and largely representing science in these regions (Debat and Babini, 2020). The perspective that most scientific advances have been made in the Global North, and that northern priorities are global priorities, can lead to the exclusion of and contempt for knowledge and priorities in other regions (Nkoudou, 2016). Such a view implies that southern science needs to develop so that it looks more and more like that of the north. It is argued that 'these [northern] partners inevitably guide the problems and the methodological and epistemological choices of African researchers towards the only model they know and value, the one born at the centre of the world-system of science – without questioning whether this model is relevant to Africa and its challenges' (Piron et al., 2017). Such biases affect our understanding of the human and natural world, and make it more difficult for researchers from parts of Asia, Africa and Latin America to contribute effectively (Rad et al., 2018).

A global science community has become a greater reality in recent years, but it will not have come of age until it replaces a unipolar perspective with an inclusive universalism, open to a wider ecology of knowledge and capable of building an authentic global knowledge commons (Gruson-Daniel, 2015; Le Crosnier, 2015). It is hoped, in Africa for example, that the development of the African Open Science Platform (2018) and efforts to strengthen a currently weak publishing industry (Kigotho, 2021) will not only stimulate increased open content from within Africa, but also, crucially, provide the means of access to that content, in addition to bringing science closer to society, promoting fair and sustainable development (Nkoudou, 2016) and creating a more powerful African voice in global science.

There is a strong case for the redesign of open-science systems in the Global South so that they are well adapted to national needs (Onie, 2020). Across Asia, Africa and Latin America, young scientists are working to build the practice of open science, not merely by importing preformed systems from elsewhere, but in other ways that are well adapted to national and regional circumstances (Participants of African Open Science Platform Stakeholder Workshop, 2018). The importation of assessment systems based crudely on numbers of papers published in countries where research cultures are being shaped inadvertently encourages bad practice, with researchers resorting to predatory journals (section 3.10) to boost their publication count and to falsifying peer reviews (Onie, 2020). Similarly, open-data practices require well-developed repositories and protocols for their function and use, so that their lack in some low- and middle-income countries has led to data sharing merely resulting in data being taken either by local or international researchers without the originator's permission or with the originator acting as a research assistant rather than a respected collaborator.

6.4 Other contrary voices

In addition to the above criticisms, there is also pragmatic scepticism to open science based on possible unintended consequences: that it could threaten the competitive, closed, sometimes solitary conduct that has fostered many major scientific achievements (Bahlai et al., 2019); that it might undermine the career potential of many young scientists for whom publication in a high-impact journal offers upward mobility (ibid); that it is inimical to commercial exploitation (Krishna, 2020); or that open access could lead to a dramatic increase in APC prices (Grove, 2021).

It is also important to note other, more principled arguments against open science, which tend to be conservative or radical (Lancaster, 2016). The conservative critique defends the right of the individual against the collective. This argument was trenchantly stated in an editorial published in the *New England Journal of Medicine* (Longo and Drazen, 2016) which described the ‘emergence of a new class of research parasites’, and also commented that some of these parasites might seek to examine whether the original study was correct, a response that implicitly but directly conflicts with a fundamental principle of scientific rigour (section 2a). A similar sentiment was expressed some years ago by a Microsoft executive, who referred to open-source computer programmes as a ‘cancer’, although the company has since joined the movement to liberate the world’s data through its ‘Open Data Campaign’.

The radical critique (Tyfield, 2013) argues that the release of vast troves of data, papers or research results, although potentially beneficial to science as an enterprise, simply exacerbates the trend towards the increasing marketization and corporatization of science that disproportionately benefit large corporations. It is argued (Tyfield, 2013) that open science opens the door to capture of publicly funded research value by commercial platforms (section 3.1), yet more ‘metrics’ of productivity to ‘incentivize’ scholars to work harder and a focus on the system-wide progress of science, ignoring costs and benefits to individuals, whether scientists or non-scientists. It is a perspective that should be taken seriously and debated within the science community, in particular in relation to issues of governance (section 7.5).

7. Exploiting digital potential

The widespread replacement of analogue by digital technologies in the last decades of the 20th century created a ‘digital revolution’ that has had profound impacts on society and on science, but with still much latent potential to be realized. Digital technologies together constitute a ‘general purpose technology’ that is the power underlying the so-called fourth industrial revolution because of its applicability to most human, societal and economic purposes²⁸ and its downward pressure on costs. It continually transforms itself in boosting productivity across all sectors and industries. It is globally pervasive, unleashing an unprecedented new era of innovation that has profound implications for science, industry and society. In an era when digital technologies are disrupting most industries, scientific publishing remains remarkably unscathed. Although thousands of publishers currently operate in the sector (including a new wave of open access publishers), just a few companies account for the majority of all articles published, most still relying on subscriptions (Box 3). How has a model created in the 17th century managed to persist until the present day? Although major commercial publishers can draw on the resources required for technological innovations that support their existing publishing model, innovations in modes of publication, for example in preprints, repositories and in open-access publishing have come from relatively recent entrants to the field.

²⁸ An example of the digital impact is illustrated by the sequencing of the human genome. This was first announced in 2003. It had taken 10 years and cost US\$4 billion. Today, as a digitally-controlled process, it can take less than a day and cost less than US\$1000.

7.1 The digital impact on the research cycle

All disciplines, whether or not data-intensive, from the humanities to the natural sciences, operate in a digital world where all the elements of the research process are connected or connectable. The art historian, the linguist, the engineer, the health economist, the physicist and the biologist, whether employed in a well-funded research institution or working independently, can and do digitally file their notes, discussions, observations, data and versions of text destined for conventional publication, in ways that permit them to be linked together as parts of a research workflow. For convenience of discussion, we represent the research process as an idealized cycle as shown in Figure 4, in which its stages are as follows:

- Formulating approaches for problems that need solutions, or for intuitive hypotheses that can be tested by observation or experiment.
- Developing a work plan and seeking funding if necessary.
- Undertaking a research programme to collect evidence to test or develop a hypothesis.
- Formulating a thesis.
- Submitting the thesis for peer review.
- Publishing the thesis including supporting data, possibly as binary outputs.
- Open critique by peers.
- Reformulating a further stage of research.

In the era of paper and print technologies, the staging points in the cycle tended to be relatively discrete and well-defined, with, in many cases, publication being a self-contained end point. As we have argued above, observations and data are first-class outputs of scientific inquiry. Where they provide evidence for a published truth claim and cannot be published within the concept article, they should be made concurrently available and consistent with FAIR principles, as part of a binary duo (Fig. 4). There are fewer excuses for failing to do so in the digital era than there were in the era of print and paper.

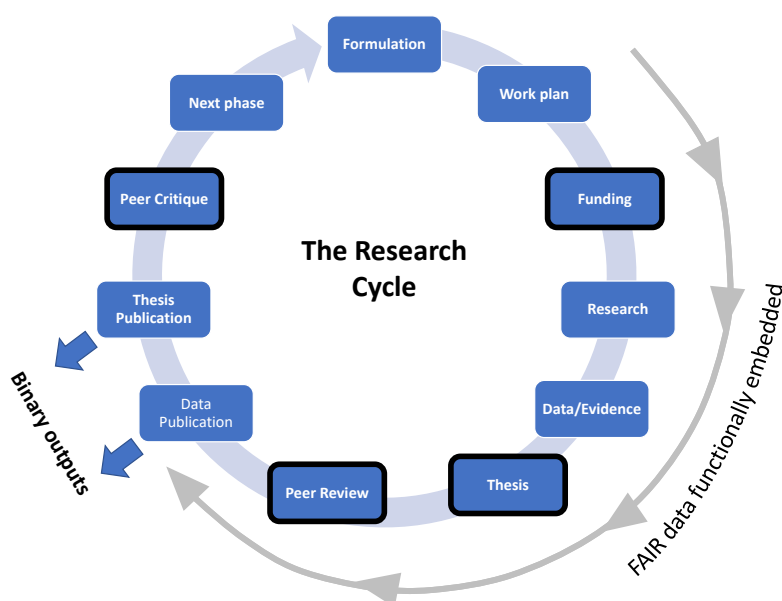


Figure 4. A simplified, schematic representation of the ‘research cycle’. The points where the cycle can be broken by failure are shown by black-rimmed boxes. Blue arrows represent formally published binary outputs, of concept paper and data publication. The grey ‘arrowed’ arc indicates the importance of routine data management using

FAIR principles that should ideally be embedded in the workflow as a readily useable functionality. The cycle may be more complex and more discontinuous. It may, for example, involve interactions with other research actors, and it may be short-circuited by creative insights, or by urgent demands for new knowledge in the face of emergency. Funding may not be required and several phases of publication may occur.

However, the concepts of openness and accessibility need not end there. In the paper and print era, documents relating to other parts of the cycle, including sharing and exchange of data, methods, software, preprints, discussion papers, funding proposals and details of collaborations, all tended to be lost. The advent of pervasive digital technologies has changed that. All elements of the cycle are connectable, with the possibility of digital interoperability across the cycle. The ease of digital production, its flexibility and connectivity compared with its print antecedent, creates the potential for access to much richer strands of work, thought and creativity. It is a potential that is increasingly exploited, reflecting the injunction of Principle VII in section 2.

7.2. Linked digital infrastructures for the research cycle

There is much value in elements of the research cycle that has hitherto been difficult to realize, for even if kept in archives, the metadata required to understand and use these elements are rarely available. Much if not all science is now done in a digital environment where research elements can be represented digitally and are thus potentially accessible and ‘publishable’. They include registration of research plans (e.g. for clinical trials), research materials, open databases for research protocols, repositories for datasets and software programmes, identifier registries for researchers, projects and research organizations, and peer reviews. All these elements are digitally resolvable, linkable, computable and interoperable. Together they form a linked digital infrastructure for the research process shown in Figure 4.

If researchers are to exploit these potentials, it is important to make management of workflows and outputs as easy as possible, with management processes embedded as normal parts of the routine of research. The Jupyter Notebook is an example of such a new kind of scholarship. It is an open-source web application that enables creation and sharing of documents that contain live code, equations, visualizations and narrative text. It can be used for data cleaning and transformation, numerical simulation, statistical modelling, data visualization, ML and much more (<https://jupyter.org>). It could accompany a conventional article, or be self-standing, as a means of releasing more of the value embedded in the research cycle.

Linked digital infrastructures may not only support the research process and produce publishable outputs, but also provide information about the research process that can help researchers, universities and funders to manage and assess research, as shown in Figure 5. They may include data about such matters as funding, productivity by discipline and institution. They are data **about** science, in contrast to the data **of** science discussed in section 4. If these were accessible, they would allow others to access, connect, analyse and reorganize ‘research information’ to better verify, understand, analyse, use and apply it. High levels of digital interoperability now make it possible to link patterns of publication by discipline, by geography, by citation and through time. Many of these elements of the digital infrastructure are routinely collected and collated by publishers through the submissions made to them and the articles that they publish (Fig. 5). These data are of value to researchers as a means of tracking publication in their discipline, to universities in managing their research effort and their researchers, and to national bodies in assessing patterns of production and productivity.

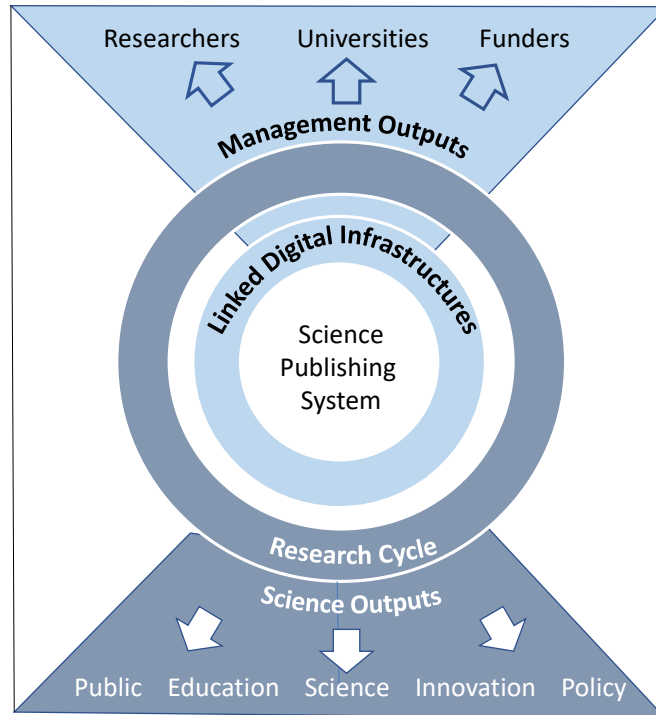


Figure 5. A schematic diagram of the digital mode of science publishing, showing the relationship between the research cycle (Fig. 4) and linked digital infrastructures. The research cycle generates scientific knowledge. The infrastructures generate strategic knowledge about science of value to those concerned with its management.

In principle this system is one in which formal publication in journal, monograph or book, no longer needs to be separately costed, but is simply one of the outputs of the research process, albeit, together with data, the most important. Its costs are simply the costs of doing research in the digital era. Provided that appropriate open-source software is made available to manage these processes, there is no reason why the unfunded researcher, or the researcher without institutional support, should not be able undertake these tasks.

7.3 Monetizing the research cycle

Section 7.1 and Figure 4 locate publishing activities as part of a spectrum of connectable research rather than as separate and discrete processes. Some of the major commercial publishers (Springer-Nature, Elsevier and Wiley) recognize this reality and are now extending their business models beyond simple content provision in journals, monographs and textbooks to the other infrastructures of the research cycle (Posada and Chen, 2018). Some (e.g. Elsevier, Pearson and Cengage) increasingly see themselves as data companies, and some IT companies are moving into the research data field. Such companies increasingly provide not only research support tools such as bibliographies and research activity syntheses, but also research assessment systems, productivity tools, online learning analytics and management systems that are derived from the data acquired from their publishing activities, and which are critical to the work of research institutions and those that fund them (Aspesi et al., 2019). This is not a replacement for the already lucrative journal and book publishing business, but an addition to it. These developing business models seek to monetize the whole of the research cycle, its management and assessment.

Data analytics are seen by some companies as the lucrative new frontier. Recent negotiations between Dutch universities and Elsevier have revealed the trade-off the publisher is prepared to make in

exchange for an extensive pilot programme on metadata (de Knecht, 2019). Elsevier's strategy is based on two key priorities (Aspesi et al., 2019):

- a) To protect its core journal business, and minimize the impact of the open access movement on that business, the company works to increase market share, for whilst its high-impact journals are important expressions of the brand, its open access, APC-based business model is built around bulk publishing. Although high-impact journals are the most sought after, the bundled deals that are frequently struck do not discriminate between these and journals of lesser standing.
- b) To expand the services that relate to all parts of the workflow of the research cycle and to sell those services to all the stakeholders. The more data Elsevier collects from citation and readership and its various databases (Scopus, Science Direct, Mendeley, SSRN and Bepress), the better it is able to enhance its competitive position through the analysis of research and publication patterns, the quality and reach of collaborative networks, and the identification of researchers likely to become future leaders in their fields (who might then be offered editorial board positions), ahead of other publishers. There are three important targets for these services that could result in excessive domination of the market:
 - **Researchers.** Elsevier wants to sell discovery services that researchers will use to learn about new fields and form hypotheses, sell them tools for grant writing to fund their experiments, sell them tools for data collection/laboratory notebooks, sell them tools for writing manuscripts, sell them preprint repository services for papers which will be preferentially submitted to Elsevier journals with the data deposited in Elsevier-held repositories, and sell them tools for analysing the performance of these published papers and datasets, which will be discussed on Elsevier's social networks of researchers. These all-in-one services will create a lock-in for researchers based on what their university purchases, with switching costs and financial penalties for using any non-Elsevier-owned product.
 - **Universities, funders, governments.** The target is to sell to these bodies services that enable them to assess the productivity of specific research areas or groups and to provide metrics to assess careers. In a 2015 investor presentation, Elsevier explicitly indicated its intent to increasingly serve university administrations, funding bodies and governments with tools aimed at estimating and improving the productivity of research and optimizing funding decisions.
 - **Businesses and investors.** Elsevier, like other companies to which researchers have relinquished copyright as a condition of publication, legally owns a treasure trove of insight, and increasingly of data. NASDAQ believes that about 30% of Elsevier's market capitalization is derived from the appropriation of academic research capital (Aspesi et al., 2019). Partnering with venture capitalists to exploit this treasure trove could prove to be the most lucrative of its options.

The strategies and processes described above are well advanced, such that it is timely for all stakeholders to consider whether to accommodate to them or to plan and work for a different future. The fundamental issues for the scientific community are not only whether or not the pervasive presence and influence of a commercial provider in the heart of the research system will bring a net benefit to Principles I–VII in section 2.1, but also whether or not the monopolistic approach outlined in (b) above is in the interests of the scientific enterprise as a global public good.

7.4 The emergence of commercial digital platforms for science

The increasing availability of data about the processes that occur in a given setting or market has given rise to a new form of organization, the digital platform. It enables a data-driven world rather than a process-driven world, in which the platform handles transactions between different players. Such a platform ideally integrates data harvesting processes, data-ingestion, ML systems to perform rules-based tasks, analytical engines and, increasingly, AI engines or tools that allow platforms to talk to other software. Such integrated digital platforms are able to produce information in a form that players on the platform find more valuable and more immediate than conventional modes of transaction. Whilst early-stage platforms were often ‘two-sided’, connecting buyers and sellers in simple transactions, there are now ‘multi-sided’ platforms that bring together consumers, service providers and other stakeholders to facilitate value exchange between them as part of a larger ecosystem (Gatti, 2020). Not only do the parties contribute to and benefit from the platform, they also generate greater utility through their participation in leveraging data generated from the platform. Examples of successful digital platforms include the following: **social media platforms** such as Facebook, Twitter, Instagram and LinkedIn; **knowledge platforms** such as Google, Yahoo! and Bing; **media sharing platforms** such as YouTube, Spotify and Vimeo; and **service-oriented platforms** such as Uber and Airbnb. The platform concept has had a dominant influence in the sectors in which it operates, not because of ownership – Airbnb owns no residences and Uber no taxis – but because their control of the data that links physical services to customers is so valuable that competition tends to be squeezed out. Their processes invariably raise significant questions about privacy, data rights and data preservation.

The trend amongst major scientific and scholarly publishers to transform themselves into technology or data companies, as discussed in section 7.3, is enabled by the data that they draw from their publishing and indexing activities (for example, the value of most scientific papers to their publishers lies more in their reference lists than it does in their scientific content). The end point of this transformation could be a new set of multi-sided relationships governed by shared standards and resetting of roles and responsibilities in **science/data/knowledge platforms** with the power to monopolize provision of the services listed in section 7.3b.

Antitrust regulators around the world are grappling with how to maintain control over digital industries to prevent monopolization by a very few platforms and to ensure socially beneficial innovation and outcomes (Gatti, 2020). There is widespread agreement amongst regulators and commentators that traditional antitrust and legal tools for assessing (and, where necessary, mitigating) the economic impact of actions by dominant firms are ill-suited to digital industries. Moreover, as faith in unfettered free markets has waned in recent years, and concerns grow about inequality, sustainability and the scale and influence of Big Tech, legislators are moving to modernize antitrust legislation and to challenge monopoly power²⁹. Many goods in digital markets are free and highly valued by users (e.g. Google Search) but may still have such damaging consequences on an industry as a whole that are sufficient to outweigh those benefits. Assessing the long-term consequences of an acquisition is extremely difficult in digital markets, and the length of time needed to impose legislative sanctions means that by the time an action is found to have been illegal or anticompetitive, the whole industry has moved on and it is too late to prevent the damaging repercussions.

²⁹ As we write, a US Congressional Committee is assessing the actions of four of the largest digital companies: Alphabet (Google’s parent), Amazon, Apple and Facebook. Numerous reports have been commissioned on this issue by national regulators: <https://research.chicagobooth.edu/stigler/events/single-events/antitrust-competition-conference/world-reports-digital-markets>.

A recent report by the Digital Competition Expert Panel for the UK Treasury recommends that regulators should:

Take a forward-looking approach that creates and enforces a clear set of rules to limit anticompetitive actions by the most significant digital platforms while also reducing structural barriers that currently hinder effective competition. These rules should be based on generally agreed principles and developed into more specific codes of conduct with the participation of a wide range of stakeholders. Active efforts should also make it easier for consumers to move their data across digital services, to build systems around open standards, and to make data available for competitors, offering benefits to consumers and also facilitating the entry of new businesses. Implemented effectively, this approach would be more flexible, predictable and timely than the current system. (Furman et al., 2019)

However, relying on national antitrust authorities to mitigate the potentially damaging activities of large international publishing entities is not likely to be a realistic alternative to community action (Gatti, 2020).

7.5 Governance of digital infrastructures: perspectives from the science community

Scientific publishing may now be entering a phase of platform development where the investment and technological capacities of major corporations are able to propel changes so rapidly that they outpace and by-pass regulatory, legal and governance systems designed to protect the public interest and those that seek to deliver it. It is a moment when the scientific community and its institutions, primarily the universities and the funders of research, need to take stock and consider whether these long-term interests are under threat, and – if necessary – to act collectively to protect them.

There are important roles for the private sector in providing services to public sector research, but there are also fundamental issues of equity and governance at stake. In terms of equity, the high current costs of access, either by authors or readers, may not deter take up in rich countries and well-funded institutions in the Global North, but it has a chilling effect in low- and middle-income countries in Global South in particular, resulting in a profound north–south divide. It is in addressing this issue and its consequence that a global response from the scientific community is required. In terms of governance, an urgent debate is required about the respective roles of commercial companies and the institutions of the scientific community. Where should the responsibility lie for the governance of:

- business management metrics for the research cycle;
- assessments of staff and student performance in universities;
- data necessary to underpin the strategies of institutions and national research systems;
- the findings that arise from publicly funded research represented by the vast troves of data contained in the record of science?

Should the governance of these systems that are so essential to the future of science be in the hands of private companies, or should it be located within the scientific community? The Scholarly Publishing and Academic Resources Coalition (SPARC) has made its view clear: ‘The need for academic institutions to act to retain control of infrastructure, data and data analytics is here to stay. It is critical for academic leaders to acknowledge that data and its uses play central roles in the operations and the future of their institutions, and take control of how it is managed as a strategic asset’ (Aspesi et al., 2019).

The above issues were raised in the discussion paper circulated to ISC members in July 2020. There was a very strong response that their governance should be in the hands of the scientific community and its institutions rather than in the hands of companies whose primary responsibilities are to their

shareholders. There is of course a role for commercial entities but it should be in providing a service to a system governed by the scientific community rather than one that they themselves govern.

Avoiding the emergence of monopoly (sections 7.3 and 7.4) is a crucial issue. National regulations and legalities tend to have a limited impact in restraining monopolistic activity by multinational companies, but the international scientific and scholarly community has the collective potential to apply significant pressure on the publishing activities of even the largest publishers. Creating a body that could identify and highlight issues of concern and anticompetitive activities by publishers, and facilitate coordinated responses by institutions globally when they negotiate contracts with publishers, could be a powerful force. Such a **digital markets unit** would provide a useful reference point and mechanism for building consensus amongst institutions and countries that are negotiating large and often complex multifaceted contracts with commercial publishing platforms (Gatti, 2020). The cost of creating and operating such a unit is likely to be small in comparison to the benefits it could have for the scientific and scholarly community.

8. Summary assessment: how well are science and scholarship served?

The principles set out in section 2 are presented as an essential basis for scientific and scholarly publishing if it is to be effective in supporting science as a global public good. The analyses in sections 3–7 explore the current state and trajectory of publishing in relation to these principles and the processes that determine them. We now assess the extent to which the present system of publishing satisfies the principles, whilst section 9 goes on to propose priorities for action designed to correct the direction of travel.

Principle I: There should be universal open access to the record of science, both for authors and readers, with no barriers to participation, in particular those based on ability to pay, institutional privilege, language or geography.

Many of the major business models discriminate on the basis of ability to pay by either authors or readers or both. They inhibit access to the record of science by scientists, inhibit public access and fracture the global scientific community by erecting expensive APC-based paywalls around their journals, thereby excluding authors from less well-funded institutions and from low- and middle-income countries.

Principle II: Scientific publications should carry open licences that permit reuse and text and data mining.

Too many publishers maintain the requirement for authors to transfer copyright to them, and too many authors are prepared to do so as the price of publishing in their journal of choice, rather than adopting an open licence. As a consequence, a large part of the record of science cannot be readily accessed for further research, mined for the knowledge it contains or accessed rapidly in response to emergencies. Transferring copyright for publicly funded research to publishers represents privatization of a public asset that should be regarded as illegitimate.

Principle III: Rigorous and ongoing peer review must continue to play a key role in creating and maintaining the public record of science.

The peer review system, essential as it is to scientific publication, is buckling under an increasingly heavy burden of demands. New ways are needed to manage the process and to adapt to the diversification of modes of scientific publishing. The system of peer review must also adapt to the need for a more agile response to crisis when there are urgent demands for rapid access to emerging knowledge. The scientific community and its institutions, including universities and funders, needs to work with publishers in seeking solutions.

Principle IV: The data and observations on which a published truth claim is based should be concurrently accessible to scrutiny and supported by necessary metadata.

Data and observations should also be ‘published’ and recognized and incentivized as first-class outputs of research. The failure to do so has played a significant part in the replication crisis of the last decade. Development of norms, protocols and business models for open-data deposition is an important priority. Publishers should work with the science community to make concurrent deposition of relevant data a prerequisite for publication in ways that are consistent with Principles I and II.

Principle V: The record of science should be maintained in such a way as to ensure open access by future generations.

Maintaining the record of science for future generations has become both more important in view of the global challenges faced by humanity, and more difficult in view of the increasing diversity of modes in which information and ideas are created and communicated. This has been recognized through initiatives such as LOCKSS (<https://www.lockss.org>), CLOCKSS (<https://clockss.org>) and Portico (<https://www.portico.org>), but needs to be expanded into a globally coordinated system of libraries as a more inclusive enterprise. It is an issue that UNESCO should take up as part of its current work on open science. The further development, federation and interoperation of digital repositories, governed in the global public interest, are important priorities for the long-term record of science.

Principle VI: Publication traditions of different disciplines should be respected, while at the same time recognizing the importance of inter-relating their contributions in the shared enterprise of knowledge.

The priorities and processes of publishers have tended to be driven by the demands from the large markets represented by the STEM disciplines. It is important to recognize that one size does not fit all, but that systems should be sufficiently flexible to accommodate a range of disciplinary needs whilst also tackling the increasing need for collaboration and interoperability across the disciplines of science in addressing major global problems.

Principle VII: Publication systems should be designed so as to continually adapt to new opportunities for beneficial change rather than embedding inflexible systems that inhibit change.

The digital revolution has created new opportunities for more efficient and effective ways of serving the principles set out above, but much of the current system is still based on operational models inherited from the era of print and paper, such that sustainable innovations that satisfy priorities for access, inclusion, agility and price are difficult to sustain.

Although support for the above principles is strong within the scientific community (section 2.2), scientists and their institutions have not exercised their market power, as consumers of publishers’ services, to effect change in ways that are consistent with these principles. Section 3.1 discusses the convoluted structure of the market that makes it inefficient in delivering on these priorities. Section 7.4 draws attention to a current trend that could lock researchers and their institutions into monopolistic research systems controlled by publishers. The response of the community (section 7.5) has been that this would not be a desirable outcome and that governance of these processes must be in the hands of the scientific community and its institutions if they are to serve the global public good.

9. Priorities for action

The current system of scientific and scholarly publishing is a ‘mixed economy’ of for-profit and not-for-profit operations, variously involving private sector commercial bodies, publicly funded systems, and institutionally based, learned society and independent operations. We expect this mix to be

maintained whilst advocating that there should be a shared view of purpose operating to serve the global public good by adhering to the Principles I–VII in section 2. On this basis, the system falls far short of what is needed if it is optimally to serve the needs of science and the global public good. There are a large number of problematic issues that are sometimes a cause and sometimes a consequence of these failures, which are discussed in detail in the foregoing text. These details are however integral to larger issues of structure, purpose and process where clarity is needed about the necessary direction of travel for Principles I–VII to be realized. They are market structures, digital opportunities and governance.

9.1 Market structures

A few major commercial publishers hold dominant positions in the journals market based on a combination of high-impact journals and a high volume of diverse journals of varying quality. Journal branding, largely a consequence of the choices made by researchers, creates luxury goods that sell for high prices, out of reach of all but those from well-funded institutions or science systems, and where a large proportion of scientific output is placed behind high paywalls to the detriment of the taxpayers who paid for much of the knowledge in the first place. The high cost of access to journals also fractures the international scientific community along economic fault lines. Researchers in high-income countries tend to see publishing through the lens of their high-impact journals³⁰. High APCs systematically exclude many researchers in low- and middle-income countries from publishing in international open-access journals, for example in Africa (Nabyonga-Orem et al., 2020), whilst many Latin American states collaborate in a publicly funded, scholar-led system which is economically efficient but regionally contained. The consequence is that the global circulation of ideas and of new knowledge is less efficient than it needs to be, when solving global challenges depends upon global engagement. It is a system where the science and scientific perspectives of the low- and middle-income countries are largely excluded from interaction with the perspectives of the Global North, with the unfortunate consequences described in section 6.3. It is not a question of extending the North American–European publishing model to the Global South, but of developing more effective global interchange. Although ‘there is enough money within the global system to cover the current rate of publication’ (Eve, 2020), and thereby to facilitate access to both readers and authors, it is not distributed in ways that match need or productivity.

It is an imperative for the scientific community that the market should evolve into one that is structured to optimize value to science. Ideally it should be a diverse ‘mixed economy’ adaptable to a variety of needs and circumstances, involving both public and private investors, and in which publishers (as suppliers) would compete for the business of authors (as customers) based on three criteria: values (our VII principles), service and price. Contracting terms would ensure that the community retained adequate control irrespective of the service provider. Such a structure would strengthen competition and customer-focussed innovation, avoid monopolies and promote inclusion. By contrast, in the internal ISC discussion on market efficiency, one member commented: ‘aren’t profit margins an indicator of efficiency (making more money while charging the same market rates)? By advocating against profit margins, one reduces incentives for improvement and efficiency. Why make your processes or services better if you’re not going to make any more money from doing so?’ The answer is that in an efficient market, companies with large profit margins risk being undercut by companies able to deliver the same or better services at a lower price. If they are not at risk, what is wrong with the market? Excessive profit margins are a sign of a dysfunctional market, where new

³⁰ Economics is an example of the distortion of a field of study by the use of high-impact journals as a proxy for scientific achievement. Publication in the ‘top five’ economics journals, all published in or concerned primarily with USA priorities, is commonly used as an arbiter of advancement in the field, even outwith the USA. The consequence is to bias the concerns of many major non-US economics departments towards US issues (Heckman and Moktan, 2019).

entrants are inhibited from competing effectively and where there are no barriers to monopoly control. It is a major issue that is exemplified in the operations of digital platforms such as Google and Facebook, and which sections 7.3 and 7.4 suggests is the direction of travel of major commercial publishers.

The brand based on impact factors is a major barrier to development of a more efficient market. If brand awareness were to collapse, so would excessive pricing. This can be done, and has happened, for example in mathematics (section 1.3). The self-interest of individuals in publishing in a journal with a 'high impact factor' because of their use in assessments of merit, should be balanced against the harms this creates in maintaining excessively high prices that exclude access by the many, reflecting the classical tension between freedom and responsibility (Mill, 1859). In moving towards a more rational market, it is vital that universities and institutes adopt the precepts of important international declarations (e.g. those of DORA, Leiden, Jussieu and Helsinki) that argue against using journal impact factors in assessing contributions to science. It is however eight years since DORA was published, and notwithstanding the numbers of signatories, its impact has been less than hoped. Although journals carry the impact factor, it is the papers that they contain that have scientific impact, such that we should focus less on the journal and more on the paper, with the possibility, indeed the desirability, that the very existence of the journal might be challenged.

9.2 Digital opportunities

Digital technologies and the emergence of the World Wide Web have brought down the cost of distributing research outputs to an additional reader to something very close to zero. In this situation, arbitrarily denying access to the findings of research that the public has paid for is difficult to justify, particularly if there are alternative, feasible ways of paying 'first copy' costs (Gatti, 2020).

The journal is a survivor from the print and paper age. In an era of digital publishing, papers do not need to be bound into a physical journal, nor do they need to be associated together in discrete, digital-only volumes, although there are some fields, particularly in the humanities, where hard-copy journals are still in demand, and of course digital products do not satisfy the ongoing demand for books. In the digital-only case, the essential requirements of science and scholarship are simply for universally-accessible indexing systems so that papers can be found, high standards of agile peer review (see section 5.2), repository systems with good functionality and open access copyright.

A move away from a version of record embedded in a date-stamped journal towards a self-standing, date-stamped digital object (article, data or other output) that has been subject to appropriate peer review and with necessary persistent identifiers could reduce costs and overcome many of the current financial barriers that inhibit the international flow of ideas. However, the journal system also acts as a rough guide to the location of scholarship in specific fields that many of us find valuable, although the great increase in the number of journals in recent decades has eroded this function. The discoverability of 'journal-independent' papers would be powerfully enabled by an efficient, formalized, universal keyword system that utilizes the power of AI to undertake thematic searches to facilitate discovery (the 'Findable' in FAIR). Field or domain specific repositories such as arXiv operate discovery systems in this way, but in the era of complexity and the SDGs, a more comprehensive approach is required. Such an approach could combine a precision and comprehensiveness far superior to those of a conventional web browser if it were able to utilize text and content mining as a universal norm in accessing the record of science, including its data and metadata³¹. Such a

³¹ As analyses of complex, multidisciplinary systems increase in frequency and diversity, it becomes increasingly difficult to find a journal that speaks both to the focus of such a study as well as its disparate component contributions. A more precise, more comprehensive discovery routine that targets individual papers would serve

development could dramatically increase the reach of the many diverse online publishing and preprint systems that have recently developed (section 5.5) and would be a powerful enabler of an inclusive open science.

9.3 Governance

Section 7.4 draws attention to the increasing inroads of major commercial publishers in areas that have hitherto been beyond their competence, in research infrastructure, assessment, management and strategy. These developments pose serious questions for the science community about whether this is a positive and helpful development, or a negative one, with these functions and, in effect monopoly control, left to companies whose primary responsibility is to their shareholders rather than to science. Our survey of the scientific community (section 2.2), as represented by ISC membership, drew a strongly negative reaction to these developments, with the comment that governance of such matters should lie primarily within the science community and its institutions. This is an issue of some urgency, as the pace of development can be so fast as to change circumstances beyond the point of no return, emphasizing the importance of early community deliberation. Establishing a 'digital markets' unit' for scientific publishing (section 7.5) could be a powerful step in early mitigation of potentially damaging trends.

9.4 Context and action for change

From small beginnings, the last two decades have seen the modern movement for open science and open access to the record of science progressively penetrate into all parts of the scientific enterprise and its stakeholders, though it has not been without critics (sections 6.3 and 6.4). Numerous open science initiatives have been created across the disciplines, in universities, amongst national and international funders, amongst publishers, in the creation of open science advocacy and policy bodies, and have led to UNESCO developing a global open science recommendation that is proposed for universal adoption and promotion through the actions of UNESCO's 193 member states (UNESCO, 2020).

The unprecedented collaborative response of the international science community to the COVID-19 pandemic, with the development of vaccines from a standing start to public vaccination in less than a year, has been a massive illustration of the power of open science in mobilizing the energies and creativity of the scientific community, and across the academic/industrial interface, in responding to a global emergency. It is vital that the potentials that it has demonstrated are not allowed to decay through a relapse to business as usual, but that the scientific community captures those potentials and embeds them in its normal modes of work. Open access to the record of science and the large-scale exploitation of new forms of early dissemination of novel, relevant results was an indispensable part of the response to the pandemic, as were its global inclusivity and the open sharing of results and ideas. These processes must be retained as parts of a new norm both within and beyond those scientific communities that were centrally involved in the pandemic response.

From this perspective of open access to the record of science, the ISC will now work with its members, national academies, international scientific unions and associations, and other regional and national science bodies, to seek tractable solutions to the major problems of scientific and scholarly publishing identified by this report: a highly inefficient market system, economic sustainability, global inequities and monopolistic trends that work against innovation and towards private governance of public assets. The ISC will also seek to engage and collaborate with national and international funders, with

potential readers more efficiently than a search of journal titles or conventional indexes. It would also dispel the need for ever more specialized journals.

universities, with open science bodies, publishers and individual scientists to create a powerful and broadly-based coalition for change to ensure that the processes of efficient dissemination and use of scientific work are central parts of a revitalized open science.

References

1. ISC (International Science Council). 2019. *Advancing Science as a Global Public Good – Action Plan 2019–2021*. Paris, ISC. DOI: 10.24948/2019.09
2. ISC (International Science Council). 2018. *High-level Strategy*. Paris, ISC. DOI: 10.24948/02
3. Stiglitz, J. E. 1999. Knowledge as a global public good. I. Kaul, I. Grunberg, M. Stern (eds.), *Global Public Goods*. Oxford University Press, New York, <http://pinguet.free.fr/stiglitz1999.pdf> (Accessed 15 July 2020).
4. Roosendaal, H. E. and Geurts, P. A. T. M. 1997. Forces and functions in scientific communication: an analysis of their interplay. <https://research.utwente.nl/en/publications/forces-and-functions-in-scientific-communication-an-analysis-of-t> (Accessed 15 July 2020).
5. STM (International Association of Scientific, Technical and Medical Publishers). 2018. *The STM Report: An Overview of Scientific and Scholarly Publishing*. Fifth Edition. The International Association of Scientific, Technical and Medical Publishers.
6. Raff, W. R. 2012. Publishing in the biomedical sciences: if it's broken, fix it! *Biol. Open*, Vol. 1, pp. 1055–7. <https://doi.org/10.1242/bio.20122477> (Accessed 20 January 2021).
7. Cohen, W. M. and Levinthal, D. A. 1990. Absorptive capacity: A new perspective on learning and innovation. *Adm. Sci. Q.*, Vol. 35, No. 1, pp. 128–52. <https://doi.org/10.2307/2393553>
8. Cohen, W.M., Nelson, R.R. and Walsh, J.P. 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not). *NBER Working Paper Series*, Working Paper No. 7552. <http://www.nber.org/papers/w7552>
9. <https://doi.org/10.3386/w7552>
10. Murray-Rust, P., Molloy, J. C. and Cabell, D. 2014. Open Content Mining. Moore, S. A. (ed.), *Issues in Open Research Data*. London, Ubiquity Press, pp. 11–30.
11. Baker, M. 2016. 1,500 scientists lift the lid on reproducibility. *Nature*, <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970> (Accessed 15 July 2020).
12. Begley, C. and Ellis, L. 2012. Raise standards for preclinical cancer research. *Nature*, Vol. 483, No. 7391, pp. 531–3. <https://doi.org/10.1038/483531a>
13. Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science*, Vol. 349, No. 6251. DOI: 10.1126/science.aac4716
14. Nosek, B. A. 2015. Promoting an open research culture: author guidelines for journals could help to promote transparency, openness, and reproducibility. *Science*, Vol. 348, No. 6242, pp. 1422–5. doi:10.1126/science.aab2374.
15. Vinopal, J. and McCormick, M. 2013. Supporting digital scholarship in research libraries: scalability and sustainability. *J. Libr. Adm.*, Vol. 53, No. 1, pp. 27–42. doi:10.1080/01930826.2013.756689.
16. Fyfe, A., Coate, K., Curry, S., Lawson, S., Moxham, N. and Røstvik, C. M. 2017. Untangling academic publishing: a history of the relationship between commercial interests, academic prestige and the circulation of research. *Zenodo*, <http://doi.org/10.5281/zenodo.546100>.
17. Baldwin, C. 2004. What do societies do with their publishing surplus? The Association of Learned and Professional Society Publishers. http://www.alpsp.org/Ebusiness/Libraries/Publication_Downloads/NFPsurpluses.sflb.aspx?download=true (Accessed 20 January 2021)

18. de Knecht, S. 2019. Leaked document on Elsevier negotiations sparks controversy. Science Guide. <https://www.scienceguide.nl/2019/11/leaked-document-on-elsevier-negotiations-sparks-controversy/> (Accessed 23 July 2020).
19. Van Noorden, R. 2020. Nature journals announce first open-access agreement. *Nature*, <https://www.nature.com/articles/d41586-020-02959-1> (Accessed 12 January 2021).
20. Shu, F., Mongeon, P., Haustein, S., Siler, K., Alperin, J. P., and Larivière, V. 2018. Is it such a big deal? On the cost of journal use in the digital era. *Coll. Res. Libr.*, Vol. 79, No. 6, p. 785. doi:<https://doi.org/10.5860/crl.79.6.785>. (Accessed 12 January 2021).
21. Nabe, J. and Fowler, D. C. 2012. Leaving the big deal: consequences and next steps. *The Ser. Libr.*, Vol. 62, No. 1–4, pp. 59–72. doi:10.1080/0361526X.2012.652524
22. Schonfeld, R. 2019. Is the value of the big deal in decline? The Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2019/03/07/value-big-deal-leakage/> (Accessed 12 January 2021).
23. Larivière, V., Haustein, S. and Mongeon, P. 2015. The oligopoly of academic publishers in the digital era. *PLoS ONE*, Vol. 10, No. 6. <https://doi.org/10.1371/journal.pone.0127502>
24. Terry, R. 2005. Funding the way to open access. *PLoS Biol.*, Vol. 3, No. 3. <https://doi.org/10.1371/journal.pbio.0030097>
25. Budapest Open Access Initiative Declaration. 2002. Available at: <https://www.budapestopenaccessinitiative.org/read> (Accessed 17 February 2021).
26. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. 2003. Available at: <https://openaccess.mpg.de/Berlin-Declaration> (Accessed 17 February 2021). Brainard, J. 2021. Open access takes flight. *Science*, Vol. 371, No. 6524, pp. 16–20. DOI: 10.1126/science.371.6524.16
27. Burchardt, J. 2014. Researchers outside APC-financed open access: implications for scholars without a paying institution. *SAGE Open*, <https://doi.org/10.1177/2158244014551714>
28. Tennant, J. 2019. ‘Transformative’ open publishing deals are only entrenching commercial power. <https://world.edu/transformational-open-access-publishing-deals-are-only-entrenching-commercial-power/> (Accessed 23 July 2020).
29. Schönfelder, N. 2020. Article processing charges: mirroring the citation impact or legacy of the subscription-based model? *Quantitative Science Studies*, Vol. 1, No. 1, pp. 6–27. https://doi.org/10.1162/qss_a_00015
30. Harvard Faculty of Arts and Sciences Open Access Policy. 2008. Harvard Library Office of Scholarly Communication. <https://osc.hul.harvard.edu/policies/fas/> (Accessed 16 July 2020).
31. Piwowar, H., Priem, J. and Orr, R. 2019. The future of OA: a large-scale analysis projecting Open Access publication and readership. *bioRxiv*, <https://doi.org/10.1101/795310>
32. Science Europe. 2018. Science without Publication Paywalls: ‘cOAlition S’ for the Realisation of Full and Immediate Open Access. Published 4 September 2018 at: <https://www.scienceeurope.org/news/science-without-publication-paywalls-coalition-s-for-the-realisation-of-full-and-immediate-open-access/> (Accessed 20 January 2021).
33. Houghton, J. and Vickery, G. 2005. Digital Broadband Content: Scientific Publishing. Working Paper. Paris, Organisation for Economic Co-operation and Development, Working Party on the Information Economy.
34. CLACSO. 2019. CLACSO Institutional Presentation, CLACSO. https://www.clacso.org/wp-content/uploads/2019/05/Institutional_presentation.pdf (Accessed 17 February 2021)
35. Debat, H. and Babini, D. 2019. Plan S in Latin America: A precautionary note. *PeerJ Preprints*, Vol. 7, e27834v2. <https://doi.org/10.7287/peerj.preprints.27834v2>

36. cOAlition S. 2019. Plan S: Principles and Implementation. <https://www.coalition-s.org/addendum-to-the-coalition-s-guidance-on-the-implementation-of-plan-s/principles-and-implementation/> (Accessed 20 January 2021)
37. Holmwood, J. 2018. The expansion of open access is being driven by commercialisation, where private benefit is adopting the mantle of public value. LSE Impact Blog, <https://blogs.lse.ac.uk/impactofsocialsciences/2018/10/02/the-expansion-of-open-access-is-being-driven-by-commercialisation-where-private-benefit-is-adopting-the-mantle-of-public-value/>
38. Frass, W. 2015. 2014. Taylor & Francis and Association of Learned and Professional Society Publishers Society Survey: Challenges Facing Learned Societies. *Editors' Bulletin*, Vol 10, No. 2, pp. 23–32. DOI: 10.1080/17521742.2014.998009
39. Wise, A. and Estelle, L. 2019. Society Publishers Accelerating Open access and Plan S (SPA-OPS) project. Welcome Trust. Collection. <https://doi.org/10.6084/m9.figshare.c.4561397.v3>
40. Coalition Publica. <https://www.coalition-publi.ca/>. (Accessed 17 July 2020).
41. Rodrigues, E. 2020. Revitalising the role of universities in scholarly communication. European University Association Expert Voices. <https://www.eua.eu/resources/expert-voices/150-revitalizing-the-role-of-universities-in-scholarly-communication.html>. (Accessed 17 July 2020).
42. Scholastica. 2019. How journals are using overlay publishing models to facilitate equitable OA. Published on the Scholastica blog, 25 October 2019 <https://blog.scholasticahq.com/post/journals-using-overlay-publishing-models-equitable-oa/>. (Accessed 20 January 2021)
43. Mellor, D. 2020. Conflict between open access and open science: APCs are a key part of the problem, preprints are a key part of the solution. Center for Open Science (COS). <https://www.cos.io/blog/conflict-between-open-access-and-open-science-apcs-are-key-part-problem-preprints-are-key-part-solution> (Accessed 20 January 2021)
44. Ferwerda, E., Pinter, F. and Stern, N. 2017. A landscape study on open access and monographs: policies, funding and publishing in eight European countries. Zenodo, <https://doi.org/10.5281/zenodo.815932>
45. Frederick, J. K. and Wolff-Eisenberg, C. 2020. Ithaka S+R US Library Survey 2019. Ithaka S+R. <https://doi.org/10.18665/sr.312977>.
46. McKenzie, L. 2017. Biologists debate how to license preprints. *Nature*. doi:10.1038/nature.2017.22161
47. Schneider, L. 2019. Preprinters of the world unite. For better science. <https://forbetterscience.com/2019/01/28/preprinters-of-the-world-unite/>. (Accessed 19 January 2021)
48. Babini, D. 2020. Toward a global open-access scholarly communications system: A developing region perspective. Eve, M. A. and Gray, J. (eds.) *Reassembling Scholarly Communications Histories, Infrastructures, and Global Politics of Open Access*. Massachusetts Institute of Technology Press.
49. Alperin, J. P. 2015. The Public Impact of Latin America's Approach to Open Access. PhD diss., Stanford University. <https://purl.stanford.edu/jr256tk1194>.
50. Open AIRE guidelines. 2015. <https://guidelines.openaire.eu/en/latest/>
51. Universities UK Open Access and Monographs Group. 2019. *Open Access and Monographs: Evidence Review*. <https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Pages/open-access-monographs-evidence-review.aspx> (Accessed 17 July 2020).
52. Adema, J. and Schmidt, B. 2010. From service providers to content producers: New opportunities for libraries in collaborative open access book publishing. *New Rev. Acad. Librariansh.*, Vol. 16, pp. 28–43.

- <https://www.tandfonline.com/doi/full/10.1080/13614533.2010.509542> (Accessed 17 July 2020).
53. Eve, M. P. 2020. The open access monograph conundrum can be solved. *Times Higher Education*, 3 November 2020. <https://www.timeshighereducation.com/blog/open-access-monograph-conundrum-can-be-solved> (Accessed 20 January 2021).
 54. White, K. 2019. Publications output: U.S. trends and international comparisons. Science and Engineering Indicators. National Science Foundation (NSF). <https://nces.nsf.gov/pubs/nsb20206/executive-summary>
 55. UNESCO. 2015. *UNESCO Science Report: Towards 2030*. Paris, UNESCO.
 56. Mandke, K. 2019. Publish or perish: How is this still a thing? Published on the Behavioural and Social Sciences community from Nature Research on 10 October 2019: <https://socialsciences.nature.com/posts/52940-is-it-publish-or-perish-for-phd-students> (Accessed 12 January 2021).
 57. Grudniewicz, A., Moher, D. and Cobey, K. D. 2019. Predatory journals: no definition, no defence, *Nature*, Vol. 576, pp. 210–2. doi: 10.1038/d41586-019-03759-y (Accessed 20 January 2021).
 58. Siler, K. 2020. There is no black and white definition of predatory publishing. LSE Impact of Social Sciences blog 13 May 2020. <https://blogs.lse.ac.uk/impactofsocialsciences/2020/05/13/there-is-no-black-and-white-definition-of-predatory-publishing/> (Accessed 23 July 2020).
 59. Xia, J., Harmon, J. L., Connolly, K. G., Donnelly, R., Anderson, M. R. and Howard, H. A. 2014. Who publishes in "predatory" journals? *J. Assoc. Inf. Sci. Technol.*, Vol. 66, No. 7, pp. 1406–17. doi:10.1002/asi.23265. hdl:1805/9740. S2CID 40929915
 60. Shen, C. and Björk, B. C. 2015. 'Predatory' open access: a longitudinal study of article volumes and market characteristics. *BMC Med.*, Vol 13, No. 230. <https://doi.org/10.1186/s12916-015-0469-2>
 61. Seethapathy, G. S., Santhosh Kumar, J. U. and Hareesha, A. S. 2015. India's scientific publication in predatory journals: need for regulating quality of Indian science and education. *Curr. Sci.*, Vol. 111, p. 11.
 62. Singh Chawla, D. 2020. Predatory-journal papers have little scientific impact. *Nature*, 13 January 2020. <https://www.nature.com/articles/d41586-020-00031-6>
 63. Brainard, J. 2020. Articles in 'predatory' journals receive few or no citations. *Science*, Vol. 367, No. 6474, p. 129 doi:10.1126/science.aba8116
 64. Callaghan, S., Donegan, S., Pepler, S., Thorley, M., Cunningham, N., Kirsch, P., Ault, L., Bell, P., Bowie, R., Leadbetter, A., Lowry, R., Moncoiffé, G., Harrison, K., Smith-Haddon, B., Wetherby, A. and Wright, D. 2012. Making data a first class scientific output: data citation and publication by NERC's Environmental Data Centres. *Int. J. Digit. Curation*, Vol. 7, No. 1, pp. 107–13. <https://doi.org/10.2218/ijdc.v7i1.218>
 65. Miyakawa, T. 2020. No raw data, no science: another possible source of the reproducibility crisis. *Mol. Brain*, Vol. 13, p. 24. <https://doi.org/10.1186/s13041-020-0552-2>
 66. National Academies of Sciences, Engineering, and Medicine. 2019. *Reproducibility and Replicability in Science*. Washington, DC, The National Academies Press. <https://doi.org/10.17226/25303>
 67. Darwin, C. R. 1871. *The Descent of Man, and Selection in Relation to Sex*. London, John Murray, Vol. 1. 1st edition, p. 385.
 68. Feynman, R. P. 1974. Cargo Cult Science. Caltech Commencement Address, available at: <http://calteches.library.caltech.edu/51/2/CargoCult.htm> (Accessed 15 February 2021).
 69. Nuzzo, R. 2014. Scientific method: Statistical errors. *Nature*, Vol. 506, pp. 150–152 (13 February 2014) doi:10.1038/506150a

70. Nosek, B. A., Ebersole, C. R., DeHaven, A. C. and Mellor, D. T. 2018. The preregistration revolution. *Proc. Natl Acad. Sci.*, Vol. 115, No. 11, pp. 2600–6.
71. Cai, S., Gallina, B., Nyström, D. and Seceleanu, C. 2019. Data aggregation processes: a survey, a taxonomy, and design guidelines. *Computing*, Vol. 101, pp. 1397–429. <https://doi.org/10.1007/s00607-018-0679-5>
72. Royal Society and British Academy. 2017. *Data Management and Use: Governance in the 21st Century*. London, The British Academy. <https://royalsociety.org/-/media/policy/projects/data-governance/data-management-governance.pdf>
73. Royal Society, 2012. *Science as an Open Enterprise*. London, The Royal Society Science Policy Centre. <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/>
74. Science International. 2015. *Open Data in a Big Data World*. Paris, International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP).
75. National Academies of Sciences, Engineering, and Medicine. 2018. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, DC, The National Academies Press.
76. Hackert, M. L., Van Meervelt, L., Helliwell, J. R. and McMahan, B. 2016. Open data in a big data world: a position paper for crystallography. International Union of Crystallography. <https://www.iucr.org/iucr/open-data> (Accessed 21 July 2020).
77. National Academies of Sciences, Engineering, and Medicine. 2020a. *Life-Cycle Decisions for Biomedical Data: The Challenge of Forecasting Costs*. Washington, DC, The National Academies Press. <https://doi.org/10.17226/25639>
78. Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data*, Vol. 3, 160018. <https://www.nature.com/articles/sdata201618> (Accessed 21 July 2020).
79. Dunning, A., de Smaele, M. and Böhmer, J. 2017. Are the FAIR data principles fair? *Int. J. Data Curation*, Vol. 12, No. 2. <http://www.ijdc.net/article/view/567> (Accessed 21 July 2020).
80. Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., Parsons, M., Robinson, R. and Wyborn, L. 2019. Make scientific data FAIR: All disciplines should follow the geosciences and demand best practice for publishing and sharing data. *Nature*, Vol. 57, pp. 27–9. <https://www.nature.com/articles/d41586-019-01720-7> (Accessed 21 July 2020).
81. PLoS ONE. 2019. Data Availability. <https://journals.plos.org/plosone/s/data-availability> (Accessed 23 July 2020).
82. Garfield, E. 2007. The evolution of the Science Citation Index. *Int. Microbiol.*, Vol. 10, No. 1, pp. 65–9, DOI: 10.2436/20.1501.01.10
83. Sauermann, H. and Haeussler, C. 2017. Authorship and contribution disclosures. *Sci. Adv.*, Vol. 3, No. 11, e1700404. doi:10.1126/sciadv.1700404
84. Not-so-deep impact. 2005. *Nature*, Vol. 435, pp. 1003–4. doi:10.1038/4351003b
85. Alberts, B. 2013. Impact factor distortions. *Science*, Vol. 340, No. 6134, p.787 doi:10.1126/science.1240319
86. Goodhart, C. 1981. Problems of monetary management: the U.K. experience. Courakis, A. S. (ed.). *Inflation, Depression, and Economic Policy in the West*, pp. 111–46.
87. Fire, M. and Guestrin, C. 2019. Over-optimization of academic publishing metrics: observing Goodhart’s Law in action, *GigaScience*, Vol. 8, No. 6., giz053.
88. Tuchman, G. 2012. Commodifying the Academic Self. *Inside Higher Ed*, February 6. Accessed May 30, 2016. <https://www.insidehighered.com/views/2012/02/06/essay-gaming-citation-index-measures> (Accessed 20 January 2021).

89. Caon, M. 2017. Gaming the impact factor: where who cites what, whom and when. *Australas. Phys. Eng. Sci. Med.*, Vol. 40, No. 2, pp. 273–6. DOI 10.1007/s13246-017-0547-1
90. Chapman, C. A., Júlio César Bicca-Marques, J. C., Calvignac-Spencer, S., Fan, P., Fashing, P. J., Gogarten, J., Guo, S., Hemingway, C. A., Leendertz, F., Li, B., Matsuda, I., Hou, R., Serio-Silva, J. C. and Stenseth, N. C. 2019 Games academics play and their consequences: how authorship, h-index and journal impact factors are shaping the future of academia. *Proc. R. Soc. B.*, Vol. 286, No. 1916. <http://dx.doi.org/10.1098/rspb.2019.2047>
91. Van Noorden, R. and Singh Chawla, D. 2019. Hundreds of extreme self-citing scientists revealed in new database. *Nature*. <https://www.nature.com/articles/d41586-019-02479-7>
92. Campbell, F. M. 1990. National bias: a comparison of citation practices by health professionals. *Bull. Med. Libr. Assoc.*, Vol. 78, No.4, pp. 376–82.
93. Aksnes D. W., Langfeldt, L. and Wouters, P. 2019. Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, doi:10.1177/2158244019829575
94. Brembs, B., Button, K. and Munafò, M. 2013. Deep impact: unintended consequences of journal rank. *Front. Hum. Neurosci.*, Vol. 7, p. 291. <https://doi.org/10.3389/fnhum.2013.00291>
95. Tjldink, J., Vinkers, C. and Otte, W. 2016. Are scientific findings exaggerated? Study finds steady increase of superlatives in PubMed abstracts. LSE Impact of Social Sciences blog. <https://blogs.lse.ac.uk/impactofsocialsciences/2016/01/26/are-scientific-findings-exaggerated/>. (Accessed 15 July 2020).
96. Nationaal Platform Open Science. 2018. Memo: Researcher Recognition and Rewarding. https://www.openscience.nl/files/openscience/2019-02/notitie-erkennen-en-waarderen-van-onderzoekers_en-gb.pdf (Accessed 20 January 2021).
97. Bregman, R. 2020. *Humankind*. London, Bloomsbury Publishing.
98. FOLEC-CLACSO (Latin American Forum for Research Assessment – Latin American Council of Social Sciences). 2020. *Towards a Transformation of Scientific Research Assessment in Latin America and the Caribbean*, available at: <https://www.clacso.org/wp-content/uploads/2020/05/FOLEC-DIAGNOSTICO-INGLES.pdf> (Accessed 22 January 2021).
99. Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. 2015. Bibliometrics: the Leiden Manifesto for research metrics. *Nature*, Vol. 520, pp. 429–31. doi:10.1038/520429a
100. Jussieu Call for Open Science and Biodiversity. 2017. Signed on 10 October 2017, available at: <https://jussieucall.org/jussieu-call/> (Accessed 20 January 2021).
101. Wakefield, A. J., Murch, S. H., Anthony, A., Linnell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., Valentine, A., Davies, S. E. and Walker-Smith, J. A. 1998. [RETRACTED] Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *Lancet*, Vol. 351, No. 9103, pp. 637–41. [https://doi.org/10.1016/S0140-6736\(97\)11096-0](https://doi.org/10.1016/S0140-6736(97)11096-0)
102. DeStefano F. and Chen R. T. 1999. Negative association between MMR and autism. *Lancet* Vol. 353, pp. 1987–8). <https://www.rti.org/publication/negative-association-between-mmr-and-aut,ism>
103. Lowe, D. 2018. Not so many uncited papers, actually. *Science Translational Medicine*. <https://blogs.sciencemag.org/pipeline/archives/2018/02/20/not-so-many-uncited-papers-actually>
104. Gillispie, C. C., Gratton-Guinness, I. and Fox, R. 1999. *Pierre Simon Laplace, A Life in Exact Science*. Princeton, NJ, Princeton University Press.

105. Smith, R. 2006. Peer review: a flawed process at the heart of science and journals. *J. R. Soc. Med.*, Vol. 99, No. 4, pp. 178–82.
106. Sullivan, B. 2018. Is it time for pre-publication peer review to die? PLoS SciComm, published 28 August 2018 at: <https://scicomm.plos.org/2018/08/28/is-it-time-for-pre-publication-peer-review-to-die/>
107. Heesen, R. and Bright, L. K. 2020. Is peer review a good idea? *Br. J. Philos. Sci.* (in print) <https://doi.org/10.1093/bjps/axz029> (Accessed 20 January 2021).
108. Barbour, V. 2020. Science publishing has opened up during the coronavirus pandemic. It won't be easy to keep it that way. *The Conversation*, 27 July 2020: <https://theconversation.com/science-publishing-has-opened-up-during-the-coronavirus-pandemic-it-wont-be-easy-to-keep-it-that-way-142984> (Accessed 9 November 2020).
109. Rovenskaya, E., Kaplan, D. and Sizov, S. 2020. *Bouncing Forward Sustainably: Pathways to a post-COVID World: Strengthening Science Systems. 2nd Consultation Report*. International Institute for Applied Systems Analysis and the International Science Council. <http://pure.iiasa.ac.at/id/eprint/16673/1/Second%20Consultation%20Report%20-%20Strengthening%20Science%20Systems%201.pdf> (Accessed 9 November 2020).
110. Vlasschaert, C., Topf, J. M. and Hiremath, S. 2020. Proliferation of papers and preprints during the coronavirus disease 2019 pandemic: progress or problems with peer review? *Adv. Chronic Kidney Dis.*, Vol. 27, No. 5, pp. 418–26. <https://doi.org/10.1053/j.ackd.2020.08.003> (Accessed 20 January 2021).
111. Vosoughi, S., Roy, D. and Aral, S. 2018. The spread of true and false news online, *Science*, 9 Mar 2018: pp. 1146–1151. DOI: 10.1126/science.aap9559 (Accessed 20 January 2021).
112. Brainard, J. 2018, Rethinking retractions. *Science*, Vol. 362, No. 6413, pp. 390–3. DOI: 10.1126/science.362.6413.390 (Accessed 20 January 2021).
113. Bolland, M. J., Avenell, A., Gamble, G. D. and Grey, A. 2016. Systematic review and statistical analysis of the integrity of 33 randomized controlled trials. *Neurology*, Vol. 87, No. 23, pp. 2391–402. <https://doi.org/10.1212/WNL.0000000000003387> (Accessed 21 July 2020).
114. Beam, A. L., Manrai, A. K. and Ghassemi, M. 2020. Challenges to the reproducibility of machine learning models in health care. *JAMA*, Vol. 323, No. 4, pp. 305–6. doi:10.1001/jama.2019.20866_ (Accessed 21 July 2020).
115. Gropp, E., Glisson, S., Gallo, S. and Thompson, L. 2017. Peer review: a system under stress. *Bioscience*, Vol. 67, No. 5, pp. 407–10.
116. Wolfram, D., Wang, P., Hembree, A. and Park, H. 2020. Open peer review: promoting transparency in open science. *Scientometrics*, 26 May 2020. <https://openresearch.community/documents/open-peer-review-promoting-transparency-in-open-science-dietmar-wolfram-et-al-scientometrics-26-may-2020> (Accessed 20 January 2021).
117. From Moral to Market Sentiments – Mark Carney: How We Get What We Value. 2020. BBC Radio 4. 4 December, available at: <https://www.bbc.co.uk/programmes/m000py8t> (Accessed 17 February 2021).
118. Smith, R. 1999. Pros and cons of open peer review. *Nat. Neurosci.*, Vol. 2, pp. 197–8. <https://doi.org/10.1038/6295>.
119. Lawson, S. 2017. Access, ethics and piracy. *Insights*, Vol. 30, pp. 25–30.
120. Vessuri, H., Guédon, J.-C. and Cetto, A. M. 2013. Excellence or quality? Impact of the current competition regime on science and scientific publishing in Latin America and its implications for development. *Current Sociology*. <http://eprints.rclis.org/23682/>

121. Biasi, B. and Moser, P. 2018. *Effects of Copyrights on Science—Evidence from the US Book Republication Program*. Cambridge, MA, USA, National Bureau of Economic Research.
122. Office of the Chief Science Advisor, Government of Canada, Call for Open Access to COVID-19 Publications, published 13 March 2020
https://www.ic.gc.ca/eic/site/063.nsf/eng/h_98016.html
123. Napack, B. Unlock Coronavirus research for scientists, petition published February 2020 on change.org <https://www.change.org/p/brian-napack-unlock-coronavirus-research-for-scientists-89a7ce07-6a46-4ed1-8a73-9745c2496df5>
124. Wheeler, D. 2006. GPL, BSD, and NetBSD - why the GPL rocketed Linux to success. Posted 1 Sep 2006 at: <https://dwheeler.com/blog/2006/09/01/>
125. Pereira, C. G., da Silva, R. R. and Geciane, P. 2015. The scientific information provided through patents and its limited use in scientific research at universities. *Braz. J. Sci. Technol.*, Vol. 2, No. 1. DOI: [10.1186/s40552-015-0007-y](https://doi.org/10.1186/s40552-015-0007-y)
126. European Commission, Directorate-General for Research, Information and Communication Unit. 2006. *Study on the Economic and Technical Evolution of the Scientific Publication Markets in Europe*.
https://ec.europa.eu/research/openscience/pdf/openaccess/librarians_2006_scientific_pub_study.pdf
127. Wallace, N. 2020. Open-access science funders announce price transparency rules for publishers. *Science*, <https://www.sciencemag.org/news/2020/05/open-access-science-funders-announce-price-transparency-rules-publishers>. (Accessed 17 July 2020).
128. Aguado-López, E. and Becerril-García, A. 2019. Latin America’s longstanding open access ecosystem could be undermined by proposals from the Global North. *LSE Impact of Social Sciences* blog. <https://blogs.lse.ac.uk/latamcaribbean/2019/11/06/latin-americas-longstanding-open-access-ecosystem-could-be-undermined-by-proposals-from-the-global-north/>. (Accessed 17 July 2020).
129. Anderson, K. 2014. Confounded complexity — Pondering the endless upgrade paths of digital publishing. Posted 20 November 2014 at:
<https://scholarlykitchen.sspnet.org/2014/11/20/confounded-complexity-pondering-the-endless-upgrade-paths-of-digital-publishing/> (Accessed 19 January 2021)
130. Hide, B. 2008. How much does it cost, and who pays? The global costs of scholarly communication and the UK contribution. *Serials*, Vol. 21, No. 3, pp. 194–200.
131. Global Scientific & Technical Publishing 2019–2023,
<https://www.researchandmarkets.com/reports/4912342/global-scientific-and-technical-publishing-2019>
132. JOSS (Journal of Open Source Software). 2019. Cost models for running an online open journal. *JOSS Blog*. <https://blog.joss.theoj.org/2019/06/cost-models-for-running-an-online-open-journal>
133. Grossmann, A. and Brembs, B. 2021. Current market rates for scholarly publishing services [version 1; peer review: awaiting peer review]. *F1000Research*, 10:20
<https://doi.org/10.12688/f1000research.27468.1>
134. arXiv Annual Update, January 2020, published 6 January 2020 at:
https://arxiv.org/about/reports/2020_update
135. Coles, P. 2019. The Cost of the Open Journal of Astrophysics. Posted 1 February on *In the Dark*: <https://telescope.wordpress.com/2019/02/01/the-cost-of-the-open-journal-of-astrophysics/>
136. Gowers, T. 2015. *Discrete Analysis* — an arXiv overlay journal. Posted on Gowers’s Weblog September 10 2015 <https://gowers.wordpress.com/2015/09/10/discrete-analysis-an-arxiv-overlay-journal/>

137. Timmer, J. 2018. The global state of science. Posted 19 January 2018 on Ars Technica: <https://arstechnica.com/science/2018/01/the-global-state-of-science/>
138. Soete, L. and Schneegans, S. 2015 Presentation of the 2015 UNESCO Science Report Towards 2030. https://en.unesco.org/sites/default/files/usr15_presentation_luc_soete.pdf
139. Gorban, A. N. and Yablonsky, G. S. 2013. *Grasping Complexity*. <https://arxiv.org/abs/1303.3855v1> (Accessed 22 July 2020).
140. International Telecommunications Union. 2019. *Measuring Digital Development: Facts and Figures 2019*. <https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf> (Accessed 23 July 2020).
141. Ipsos MORI. 2019. Trust: The Truth. <https://www.ipsos.com/ipsos-mori/en-uk/ipsos-thinks-trust-truth> (Accessed 20 January 2021).
142. Carter, J. 2020. The American public still trusts scientists, says a new Pew survey, posted 29 September 2020, <https://www.scientificamerican.com/article/the-american-public-still-trusts-scientists-says-a-new-pew-survey/> (Accessed 14 January 2021).
143. Titz, S. 2016. The Long March to Open Science. Swiss National Science Foundation-Swiss Academies. <http://www.snf.ch/en/researchinFocus/newsroom/Pages/news-160902-horizons-the-long-march-to-open-science.aspx> (Accessed 22 July 2020).
144. Gibbons, M., Trow, M., Scott, P., Schwartzman, S., Nowotny, H. and Limoges, C. 1995. The new production of knowledge: The dynamics of science and research in contemporary societies. *Contemp. Sociol.*, Vol. 24, No. 6. DOI: 10.2307/2076669
145. Nowotny, H., Scott, P. and Gibbons, M. 2003. Introduction: Mode 2 revisited: the new production of knowledge. *Minerva*, Vol. 41, pp. 179–94. <https://link.springer.com/article/10.1023/A:1025505528250> (Accessed 22 July 2020).
146. UNESCO. 2020. First draft of the UNESCO Recommendation on Open Science (SC-PCB-SPP/2020/OS/R1) available at: <https://unesdoc.unesco.org/ark:/48223/pf0000374837>
147. Bowman, N. D. and Keene, J. R. 2018. A layered framework for considering open science practices. *Commun. Res. Rep.*, Vol. 35, No. 4, pp. 363–72
148. Chan, L., Okune, A., Hillyer, B., Albornoz, D. and Posada, A. (Eds). 2019. *Contextualizing openness: situating open science*. Ottawa, The University of Ottawa Press. <https://www.idrc.ca/sites/default/files/openebooks/Contextualizing-Openness/9781552506110.html#ch1> (Accessed 22 July 2020).
149. International Science Council. 2020. Open Science for the 21st Century, Draft ISC Working Paper: https://council.science/wp-content/uploads/2020/06/International-Science-Council_Open-Science-for-the-21st-Century_Working-Paper-2020_compressed.pdf
150. Sample, I. 2020. The great project: how Covid changed science for ever. *The Guardian*, 15 December 2020: <https://www.theguardian.com/world/2020/dec/15/the-great-project-how-covid-changed-science-for-ever>
151. National Academies of Sciences, Engineering, and Medicine. 2020b. *Advancing Open Science Practices: Stakeholder Perspectives on Incentives and Disincentives: Proceedings of a Workshop—in Brief*. Washington, DC, The National Academies Press. <https://doi.org/10.17226/25725>
152. Fanon, F. 1961/2002. *Les damnés de la terre* [The Wretched of the Earth]. Paris, Editions la Découverte. (In French).
153. Piron, F., Diouf, A., Madiba, M., Nkoudou, T., Ouangré, Z., Tessy, D., Achaffert, H., Pierre, A. and Lire, Z. 2017. Le libre accès vu d’Afrique francophone subsaharienne [Open Access viewed from francophone sub-Saharan Africa]. *Revue Française des Sciences de l’Information et de la Communication*, No. 11 (in French). <https://doi.org/10.4000/rfsic.3292>

154. Debat, H. and Babini, D. 2020. Plan S in Latin America: a precautionary note. *Schol. Res. Commun.*, Vol. 11, No. 1. <https://src-online.ca/index.php/src/article/view/347> (Accessed 23 July 2020).
155. Mboa Nkoudou, T. H. 2016. Le Web et la production scientifique africaine : visibilité réelle ou inhibée ? [The web and the production of African science: real or inhibited visibility?] Project SOHA <http://www.projetsoha.org/?p=1357> (Accessed 23 July 2020), (in French).
156. Rad, M. S., Martingano, A. J. and Ginges, J. 2018. Toward a psychology of *Homo sapiens*: making psychological science more representative of the human population. *Proc. Natl Acad. Sci. USA*, Vol. 115, pp. 11401–5.
157. Gruson-Daniel, C. 2015. *D'une Open Science mainstream à une science ouverte engagée : enjeux de la recherche dans les pays du Sud* [From mainstream open science to engaged, open science: issues for research in the Global South], Présenté à Colloque du CIRAD, Montpellier. <https://journals.openedition.org/rfsic/3292?lang=fr> (in French)
158. African Open Science Platform. 2018. The Future of Science and Science of the Future: Vision and Strategy for the African Open Science Platform. <https://doi.org/10.5281/zenodo.2222418>
159. Kigotho, W. 2021. African academics may perish even when they have published. *University World News*, Africa Edition, 14 January 2021: <https://www.universityworldnews.com/post.php?story=20210110220840860> (Accessed 20 January 2021).
160. Le Crosnier, H. 2015. *Une introduction aux communs de la connaissance: Recueil d'articles* [An introduction to knowledge commons: selected articles]. Caen, C & F Éditions (in French).
161. Onie, S. 2020. Redesign open science for Asia, Africa and Latin America. *Nature*, Vol. 587, No. 5, pp. 35–7.
162. Participants of African Open Science Platform Stakeholder Workshop, September 2018, Participants of African Open Science Platform Strategy Workshop, March 2018, Advisory Council, African Open Science Platform Project, Technical Advisory Board, African Open Science Platform, Boulton, G., Hodson, S., Serageldin, I., Qhobela, M., Mokhele, K., Dakora, F., Veldsman, S. and Wafula, J. 2018. *The Future of Science and Science of the Future: Vision and Strategy for the African Open Science Platform (v02)*. Zenodo. <http://doi.org/10.5281/zenodo.2222418>
163. Bahlai, C., Bartlett, L. J., Burgio, K. R., Fournier, A. M. V., Keiser, C. N., Poisot, T. and Stack Whitney, K. 2019. Open science isn't always open to all scientists. *Am. Sci.*, Vol. 107, No. 2, p. 78. DOI: 10.1511/2019.107.2.78
164. Krishna, V. V. 2020. Open science and its enemies: challenges for a sustainable science–society social contract. *J. Open Innov. Technol. Mark. Complex.*, Vol. 6, No. 3, p. 61.
165. Grove, J. 2021. Plan S launches, but will it drive down publisher fees? Times Higher Education, January 13, 2021. <https://www.timeshighereducation.com/news/plans-launches-will-it-drive-down-publisher-fees> (Accessed 20 January 2021).
166. Lancaster, A. 2016. Open Science and its Discontents. Ronin Institute: Reinventing Academia. <http://ronininstitute.org/open-science-and-its-discontents/1383/> (Accessed 16 July 2020).
167. Longo, D. L. and Drazen, J. M. 2016. Data sharing. *New Engl. J. Med.*, Vol. 374, pp. 276–7.
168. Tyfield, D. 2013. Transition to Science 2.0: “Remoralizing” the Economy of Science. *Economic Aspects of Science*, Vol. 7, No. 1. <https://spontaneousgenerations.library.utoronto.ca/index.php/SpontaneousGenerations/article/view/19664>. (Accessed 16 July 2020).

169. Posada, A. and Chen, G. 2018. *Inequality in Knowledge Production: The Integration of Academic Infrastructure by big publishers*. Toronto, Canada, ELPUB.
<https://hal.archives-ouvertes.fr/hal-01816707> (Accessed 17 July 2020).
170. Aspesi, C., Allen, N., Crow, R., Daugherty, S., Joseph, H., McArthur, J. and Shockey, N. 2019. *SPARC Landscape Analysis: The Changing Academic Publishing Industry – Implications for Academic Institutions*. 10.31229/osf.io/58yhb.
171. Gatti, R. 2020. Business Models and Market Structure within the Scholarly Communications Sector. International Science Council Occasional Paper. DOI: 10.24948/2020.04
172. Furman, J., Coyle, D., Fletcher, A., McAuley, D. and Marsden, P. 2019. *Unlocking digital competition*. Report of the Digital Competition Expert Panel.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785547/unlocking_digital_competition_furman_review_web.pdf (Accessed 9 November 2020).
173. Nabyonga-Orem, J., Asamani, J. A., Nyirenda, T. and Abimbola, S. 2020. Article processing charges are stalling the progress of African researchers: a call for urgent reforms. *BMJ Global Health*, Vol. 5, e003650.
174. Heckman, J. J. and Moktan, S. 2019. Publishing and promotion in economics: the tyranny of the top five. National Bureau of Economic Research, Working Paper 25093.
<http://www.nber.org/papers/w25093>
175. Mill, J. S. 1859. *On Liberty*. www.gutenberg.org

Appendix: Report production and review

Working Group members

- Dr Bianca Amaro, President, La Referencia; and Coordinator of the Brazilian Open Science Program, Brazilian Institute of Information in Science and Technology
- Dr Dominique Babini, Open Science Advisor, Latin American Council of Social Sciences (CLACSO)
- Professor Michael Barber, Emeritus Professor of Mathematics and former Vice Chancellor, Flinders University
- Professor Geoffrey Boulton (Chair and principal author), Regius Professor of Geology Emeritus, University of Edinburgh; Member of the International Science Council Governing Board
- Professor Robin Crew, Senior Research Fellow (Entomology), Centre for the Advancement of Scholarship, University of Pretoria; Past-President, Academy of Science of South Africa
- Professor Luke Drury, Emeritus Professor of Astrophysics, Dublin Institute for Advanced Studies, and former President, Royal Irish Academy
- Professor Martin Eve, Professor of Literature, Technology and Publishing, Birkbeck, University of London
- Professor Sari Hanafi, Professor of Sociology, American University of Beirut; President, International Sociological Association
- Mary Lee Kennedy, Executive Director, Association of Research Libraries, USA
- Professor Nathalie Lemarchand, Professor of Geography, University Paris 8; Vice-President, International Geographical Union
- Professor Anna Mauranen, Professor of English Philology, University of Helsinki; President, Finnish Academy of Science and Letters
- Dr Ravi Murugesan, Associate, International Network for Advancing Science and Policy (INASP)
- Dr Joseph Mwelwa, JointMindsConsult, Botswana
- Dr François Robida, French Geological Survey (BRGM) and International Union of Geological Science Commission for the Management and Application of Geoscience Data
- Mr Peter Strickland, Executive Managing Editor, International Union of Crystallography
- Professor Zhang Xiaolin, Chinese Academy of Science, China

Special Adviser on the Economics of Publishing

Dr Rupert Gatti, Faculty of Economics, University of Cambridge

International Science Council Staff

Lead: Lizzie Sayer, Senior Communications Officer

Support: Zhenya Tsoy, Digital Communications

Review Committee

- Professor Pearl Dykstra (Chair), Director of Research of the Department of Public Administration and Sociology at the Erasmus University Rotterdam
- Ms Hazami Habib, Chief Executive Officer, Academy of Sciences Malaysia
- Dr Molapo Qhobela, Chief Executive Officer, National Research Foundation, South Africa

- Dr Peter Suber, Director of the Harvard Office for Scholarly Communication, Director of the Harvard Open Access Project (Berkman Klein Center), and Senior Researcher, Berkman Klein Center
- Professor Sally Wyatt, Professor of Digital Cultures, Technology & Society Studies, Maastricht University

----ENDS----



**International
Science Council**

The global voice for science