OECD GOING DIGITAL TOOLKIT

POLICY NOTE

Open Science - Enabling Discovery in the Digital Age

OECD
BETTER POLICIES FOR BETTER LIVES

going digital

This Toolkit note was written by Alan Paic. It was reviewed by the Committee for Scientific and Technological Policy (CSTP), and it was declassified by the CSTP on 13 May 2021. The note was prepared for publication by the OECD Secretariat.

This Toolkit note is a contribution to the OECD Going Digital project, which aims to provide policy makers with the tools they need to help their economies and societies thrive in an increasingly digital and data-driven world.

For more information, visit *www.oecd.org/going-digital*.

#GoingDigital

Please cite this publication as:

Paic, A. (2021), "Open Science - Enabling Discovery in the Digital Age", *Going Digital Toolkit Note*, No. 13, *https://goingdigital.oecd.org/data/toolkitnotes/No13_ToolkitNote_OpenScience.pdf*.

Note to Delegations:

This document is also available on O.N.E. under the reference code:

DSTI/STP(2020)24.

Table of Contents

Figures

Boxes

Open Science - Enabling Discovery in the Digital Age

Data-driven innovation and data-intensive science hold immense promise to address grand societal challenges. Open science initiatives, which facilitate open access to publications, data, algorithms, software and workflows, play an essential role in accelerating needed scientific research and the innovation process itself. This Going Digital Toolkit note provides an overview of the open science movement, highlights achievements of open science including that in the context of the COVID-19 pandemic, identifies challenges to achieving all of the benefits that open science has to offer, and sheds light on the evolution of open science policies in a range of economies. The note also advocates a way forward that involves the seven pillars of the revised OECD Recommendation of the Council concerning Access to Research Data from Public Funding: (1) Data governance for trust; (2) Technical standards and practices; (3) Incentives and rewards; (4) Responsibility, ownership and stewardship; (5) Sustainable infrastructures; (6) Human capital; and (7) International co-operation for access to research data.

As society and economy are increasingly knowledge-based, data become a key resource. Data-driven innovation is transforming society through far-reaching effects on resource efficiency, productivity and competitiveness (OECD, 2015[1]). It also helps address many global challenges, such as climate, demographic changes, and pandemics. Data-intensive science is seen as the fourth paradigm, after empirical science, theoretical science and simulation (Hey, Tansley and Tolle, 2009[2]). Open science also creates spill-over effects and positive externalities, such as behavioural change, cultural and scientific exchange, and greater levels of trust induced by increased transparency.

The benefits from open science include: opportunity for new scientific discoveries, reproducibility of scientific results, facilitating cross-disciplinary co-operation, economic growth through better opportunities for innovation, increased resource efficiency, improved transparency and accountability regarding disbursement of public funds, better return on public investment, securing public support for research funding, and increasing public trust in research in general. Enhanced access also furthers other government/public missions, e.g. health, energy security and transportation. However, enhanced access also bears associated risks related to privacy, intellectual property, national security and the public interest, including the protection of rare and endangered species. These risks need to be adequately communicated and responsibly managed (OECD, 2020[3]).

Open science is disrupting the way science is done, catalysing the creative process and removing the barriers to the diffusion of knowledge. It increases openness through rapid, convenient and high-quality scientific communication between researchers and society at large, fostering better response to societal challenges, and providing business opportunities through the development of innovative products and services (Netherlands Ministry of Education, Culture and Science, 2018[4]). In global emergencies like the COVID-19 pandemic, open science policies can remove obstacles to the free flow of research data and ideas, and thus accelerate the pace of research critical to combating the disease (OECD, 2020[5]).

Due to the proliferation of data, science in general, and open science in particular are highly dependent on the rapid deployment of ever larger, ever faster data infrastructures, which are sustainable in the long term. Sustainability in this context means that high value data are available and re-usable over the long term, taking into account the evolution of hardware technologies, as well as software and data format obsolescence. Even though this note does not focus on data infrastructure per se, high-level recommendations regarding sustainable infrastructure are discussed as part of the dispositions of the recently adopted OECD Recommendation of the Council Concerning Access to Data from Public Funding (OECD, 2021[6]).

Open science spans a range of topics (Box 1), but this note focuses on the two major pillars: Open access to publications and open data.

Box 1. Definitions

Open science. There is no formal definition of open science. The OECD previously referred to it as representing efforts by researchers, governments, research funding agencies or the scientific community itself to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction as a means for accelerating research; these efforts are in the interest of enhancing transparency and collaboration, and fostering innovation. The three main aspects of open science are: open access, open research data, and open collaboration enabled through ICTs. Other aspects of open science – post-publication peer review, open research notebooks, open access to research materials, open source software, citizen science, and research crowdfunding – are also part of the architecture of an "open science system" (OECD, 2015[7]).

Based on a comprehensive analysis of 75 studies, 67 articles from reference journals and 8 official publications from Intergovernmental organisations' databases, Vicente-Saez and Martinez-Fuentes conclude that "open science is transparent and accessible knowledge that is shared and developed through collaborative networks" (Vicente-Saez and Martinez-Fuentes, 2018[8]). The term "knowledge" includes code, data, ideas, information, scientific outputs, scientific publications, and scientific results. The authors further include emerging trends on open science such as alternative reputation systems, open notebooks, open lab books, science blogs, collaborative bibliographies, citizen science, and open peer review.

Open access is understood as unrestricted online access to scientific articles, via a number of channels, such as institutional repositories, journal publishers' websites, researchers' webpages, etc. (OECD, 2015[7]).

Green open access refers to the practice of self-archiving the pre-print or the post-print of an article, generally by its author. The costs of green open access are generally covered by institutional funding or a percentage of research grants (OECD, 2015[7]).

Gold open access is access provided by a publisher. Under gold open access, generally the publishing costs and revenues are recovered through fees (OECD, 2015[7]).

Hybrid open access is open access provided by subscription-based journals where some articles are available in open access, provided that Article processing charges have been paid (OECD, 2015[7]).

Diamond open access is access provided by not-for-profit, non-commercial organisations, associations or networks publishing material that is made available

online in digital format, is free of charge for readers and authors and does not allow commercial and for-profit re-use (Fuchs and Sandoval, 2013[9]).

Open data refers to data access and sharing arrangements, where data can be accessed and shared, data that can be accessed and reused by anyone without technical or legal restrictions, free of charge (to the greatest extent possible), and used by anyone for any purpose subject, at most, to requirements that preserve integrity, provenance, attribution, and openness (OECD, 2015[7]). It is important to note that access to data is not a binary concept – rather, it can be staged to different degrees of openness, depending on the community of stakeholders involved. "As open as possible, as closed as necessary" is often used to illustrate the fact that while opening up data can help advance the Science, technology and innovation (STI) agenda, this needs to be balanced against issues of costs, privacy, security, intellectual property rights and preventing malevolent uses.

Sources: (OECD, 2015[7]); (Vicente-Saez and Martinez-Fuentes, 2018[8]); (Fuchs and Sandoval, 2013[9]).

Open science in practice

Publication of scientific results has been the norm since the very first scientific journals were created in the 17th century, enabling scientists to build on each other's ideas and make collective progress on the frontier of knowledge. However, access to such publications was not open and free, and considerable budgetary outlays were needed to provide scientists with this resource, which is an absolutely essential input to their work. Sharing the underlying data, algorithms, software and workflows, for the purpose of reproducibility, reuse for analysis and meta-analysis was anecdotal rather than systematic – in most cases the interested user needed to address a personal request to the author, who chose to grant access, or not. Provision of those additional resources was contingent on the initiative of an individual researcher contacting another researcher, requesting the additional information, subject to case-by-case decision-making.

The provision of open and free access to scientific publications, research data, and other digital resources needed by scientists can be justified by the public good argument – publicly funded research funded by taxpayers should be accessible to the public free of charge. In addition, free access to all aspects of scientific knowledge will create externalities for society, which in most cases surpass the cost of provision as well as the private value of those resources (Box 2).

Box 2. The value of open science

The European Commission commissioned a report which estimates the opportunity cost of not having Findable, Accessible, Interoperable and Re-usable (FAIR) data, using seven indicators to estimate the cost of not having FAIR research data: time spent, cost of storage, licence costs, research retraction, double funding, interdisciplinarity and potential economic growth. Using this method, they arrive at an estimation of EUR 10 billion annually (European Commission; PWC, 2019[8]).

An estimate of the impact of open access to Public Sector Information (PSI), commissioned by the European Commission, has reviewed several methods of estimating the economic value of PSI, as the value added of PSI with respect to the economy as a whole and private sector expenditure on PSI. Based on estimates of the spatial information industry, the narrow EU27 PSI industry has been estimated to EUR 18 billion, while the expanded economic impacts are estimated at EUR 70-140 billion, in 2008, representing 0.6-1.1% of EU27 gross domestic product (GDP) (Vickery, 2011[9]).

An Australian study estimated that data from research alone amounted to 0.15% to 0.4% of GDP in 2012, with potential upsides to 0.3% to 1% of GDP (Houghton, J., 2014[10]) while a study specifically on the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Data Access Portal, estimates benefits at 67 million AUD annually, nearly two orders of magnitude higher than the cost of dissemination (Sanderson, Reeson and Box, 2017[11]).

Sources: (European Commission; PWC, 2019[8]); (Vickery, 2011[9]); (Houghton, J., 2014[10]); (Sanderson, Reeson and Box, 2017[11]).

Fostering open science is therefore a desirable societal objective, even though in some cases it needs to be limited by measures to protect private, public, and community interests, including national security, intellectual property rights, privacy, personal data, the environment, key natural resources including water and minerals, and endangered species.

In a 2014 consultation by the European Commission, stakeholders identified open science as the dominant driver for the future of science, with high expectations of improved scientific integrity, better connection of science and society, and making science more responsive to societal challenges. Furthermore, a vast majority agreed that policy intervention is needed to mainstream open access to data and publications (European Commission, 2014[12]).

Open Access to publications

The origins of open access to publications can be traced to the 1960's and 1970's, with the creation of open access to bibliography with the *Educational Resources Information Center* and Medline, called *Pubmed* today. The

movement was amplified through the emergence of the Internet in the 1990's with freely accessible scientific articles in physics, mathematics and computer science on *ArXiv*, *BioMedCentral* and others (Machado, 2015[13]).

Open access to publications results in a well-documented citation advantage, meaning that papers published under open and free access tend to have more impact on their peers, than their paid access equivalents[1]. While some of the observed effect may be due to the higher quality of such papers[2], it was shown in the domain of high energy physics that even top quality publications are less cited if they aren't released as open access preprints prior to publication (Piwowar et al., 2018[14]) (OECD, 2015[7]), (Boselli and Galindo-Rueda, 2016[15]).

An ever-increasing number of research organisations and funding agencies have adopted open science policies. For example, the number of Open Access policies worldwide has progressed more than eightfold since 2005, and currently concerns more than 1000 organisations worldwide, including 86 funding organisations (Figure 1). The overwhelming majority of open access policies (about 63%) are in Europe[3], 16% in North America, 7% in Asia, 6% in Latin America, 4% in Oceania and 3% in Africa. Close to half of those policies require open access, while another third request or recommend it. More than a third of the initiatives permit open access requirement to be waived. Also, a large number of initiatives allow for embargo periods before open access is enforced, as well as constraints imposed by publishers (ROARMAP, n.d.[16]).

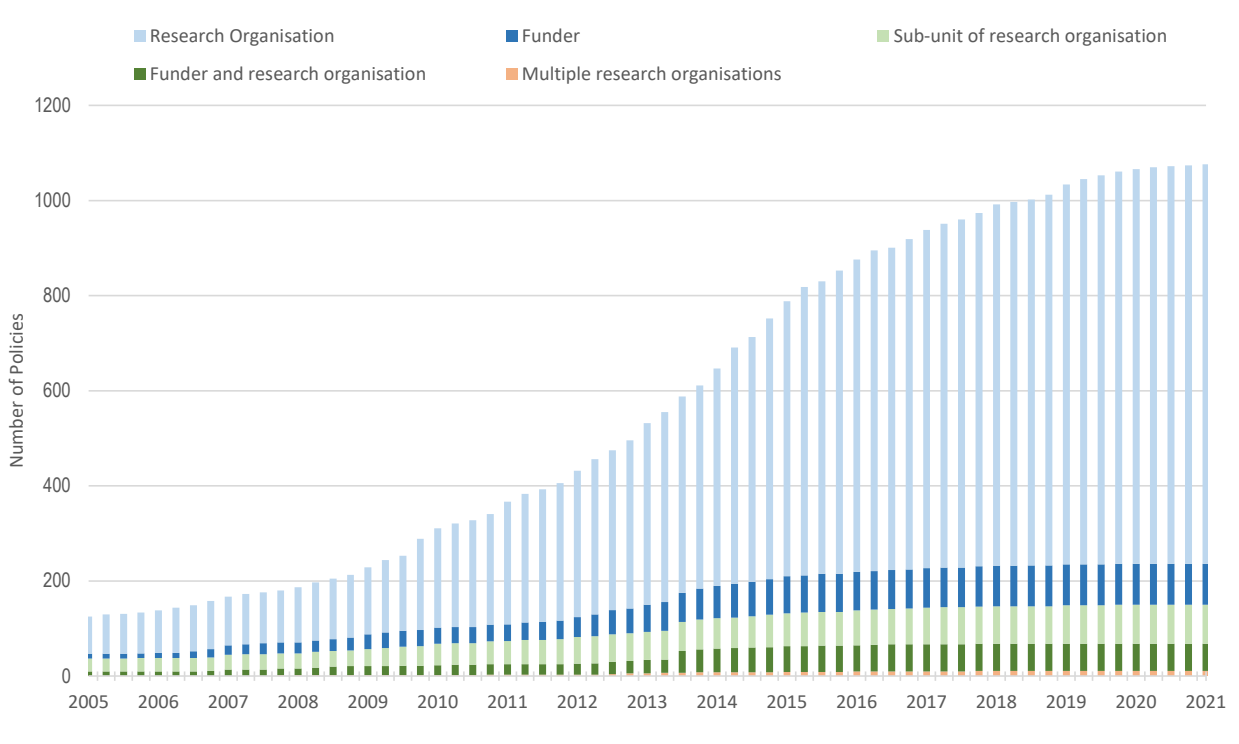
The various flavours of open access defined in Box 1 have strengths and weaknesses. While green open access is free to both the author and the reader, it most often relies on self-publishing, with authors depositing manuscripts on open access servers without a peer review process which provides some degree of quality check. Gold open access provides a rigorous peer review mechanism, but also requires authors to "pay to publish", imposing an article processing charge in order to cover those costs. Recently, an alternative to those two most widespread alternatives has been named "diamond open access". In this variant, the peer review process is being performed by hard-working volunteers, who should appreciate the possibility to review the very newest research in their field of expertise (Hoorn, 2014[17]).

[1] A minority of studies finds inconclusive results, or even an open access disadvantage.

[2] Some open access is so-called gold open access, where the author's institution has to pay for the paper to be published, therefore a selection is applied and only the worthiest publications are sponsored.

[3] Including the European Union, the United Kingdom, as well as Eastern Europe.

Figure 1. Adoption of open access policies



Note: Multiple research organisations designated institutions such as the Max Planck Society, the Academy of Sciences of the Czech Republic and others.
***Source*:** (ROARMAP, n.d.[16]).

In a recent large scale study on 300,000 articles (Piwowar et al., 2018[14]) find that at least 28% of all scholarly literature is open access (19 million papers), and the proportion is growing with time such that 45% of all papers published in 2015 (last year of the survey) were open access. It also finds that the dominant category of open access is not green, gold or hybrid open access, but articles made free to read on the publisher website, without an explicit open license. Open access also varies widely by discipline, with more than 80% of astronomy, astrophysics, embryology, tropical medicine and fertility papers published between 2009 and 2015 available in open access, while less than 10% of those in Pharmacy, Inorganic & Nuclear Chemistry, Criminology and Applied Chemistry are open access. Prevalence of green open access also varies widely, with more than 50% in Nuclear & Particle Physics, and less than 2% in Anaesthesiology and Tropical Medicine (Piwowar et al., 2018[14]).

A 2016 OECD survey of scientific authors showed that approximately 50-55% of documents are openly available 3-4 years after publication. Authors from emerging and developing countries tend to rely more on open access journals than their OECD counterparts for open access publishing (Boselli and Galindo-Rueda, 2016[15]).

Open data

The practice of sharing research data originated in disciplines where large international consortia collaborate on complex projects, such as in particle physics, geophysics, oceanography, molecular biology. A pioneering initiative in data sharing is the World Data Center (World Data System, n.d.[18])established in 1957 by the International Council for Science to serve the International Geophysical Year, a worldwide effort to study the earth, oceans, and atmosphere in a coordinated, synchronous way (Korsmo, 2010[19]). Another significant data sharing platform is Genbank, the genetic research database created in 1992 (Machado, 2015[13]). In the early days, data exchange proceeded through physical media, such as tape and disks, or through a myriad of private exchange networks, which eventually gave rise to the Internet. A large contribution to the emergence of the Internet was given by the European Laboratory for Particle Physics (CERN), where the concept of the World Wide Web was developed in 1990, as an answer to the ever-increasing needs of particle physicists to exchange large volumes of data (Berners-Lee et al., 1992[20]).

Open data can be defined simply as "data that can be accessed and reused by anyone without technical or legal restrictions" (Box 1). In the context of open science, the user usually bears no charges, even though some institutions reserve the right to charge for marginal costs of dissemination. In the Internet age, such cost is very often close to zero, except in cases where specific curation is needed, such as anonymization of personal data records.

Different business models for dissemination of research data include institutional subscription to research databases; open access in the "author pays" variant – authors or their employers pay for the cost of publishing in order to provide free access to the community; open-access archives and repositories, where organisations support institutional repositories and/or subject archives, and authors make their work freely available to anyone with Internet access; and a number of hybrid solutions, such as delayed open access and open choice (Houghton and Sheehan, 2009[21]); (OECD, 2017[22]).

Sharing research data by default has become the norm in many countries and disciplines, as long as no constraints are present due to privacy, national security, intellectual property or other legitimate public and private interests. More restricted access to data can be organised within the framework of safe environments, which rely on specific safe-software platforms, where only approved researchers can access the data within a specific environment, analyse them without extracting the actual sensitive data and then submit the results of their research for approval (Office of National Statistics UK, n.d.[23]).

A survey of 1 381 research data repositories worldwide and across all research disciplines shows that in 2015, 86% of the repositories provide open access to at least part of the data, 12% offer exclusively restricted access, and 2%

propose a combination of closed and restricted access. Out of the 86% repositories that are at least partly open, 50% are fully open, 32% have restricted parts of the datasets, 6% have embargoed datasets, and 6% have closed datasets (Kindling, 2017[24]).

The 2018 edition of the OECD International Survey of Scientific Authors (ISSA2) shows that on average, 67% of scientific production results in new data or code. Authors seem to be more likely to share their data than code. Code was archived on a repository or delivered to a journal as supportive material in about 20% of cases, whereas around 45% of authors shared their data using these means. Reuse is yet another barrier to be overcome, since even when shared, data is not always FAIR, i.e. not accompanied by relevant metadata nor compliant with relevant standards, and even fewer are the cases where an object identifier is assigned. Payment of a fee is required in about 12% of the cases. The main drivers for sharing of data identified were career objectives and peer expectations, rather than formal sharing requirements from funders. The most significant barriers identified were high dissemination costs as well as intellectual property issues (Bello and Galindo-Rueda, 2020[25]).

Open science policies

Policy settings in the last two decades have strived to reduce barriers to creating open access to scientific publications and data. This section discusses international initiatives, recommendations and guidelines that have been implemented through various national policies and strategies.

International initiatives, recommendations and guidelines

The 2002 Budapest Open Access Initiative is an early landmark, which positioned open access as an "unprecedented public good". It proposed some specific strategies, such as self-archiving by scholars and the creation of Open Access journals (Budapest Open Access Initiative, 2002[26]). The 2003 Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities contributed a definition of an Open Access contribution, and suggested active advocacy in favour of the Open Access paradigm (Berlin Declaration, 2003[27]).

A strong impetus to open access to research data was given by the OECD's Committee for Scientific and Technological Policy at its Ministerial Meeting in 2004, and the ensuing Recommendation of the Council concerning access to publicly funded research data, which established a pioneering set of principles (OECD, 2006[28]). These principles have contributed to foster a culture of openness and embed access and sharing in science policies (OECD, 2020[29]).

In 2012, the European Commission issued a Recommendation on access to and preservation of scientific information, calling for co-ordinated open access to scientific publications and data, preservation and reuse of scientific

information, and the development of ICT-based infrastructures among EU member states. This Recommendation was updated in 2018, extending the application of "open access that policies aim to provide researchers and the public at large with access to peer-reviewed scientific publications, research data and other research outputs free of charge in an open and non-discriminatory manner as early as possible in the dissemination process, and enable the use and reuse of scientific research results". It also underlines data-management planning is becoming a standard scientific practice (European Commission, 2018[30]).

Science Europe established a set of ten principles for the transition to Open Access in 2013. Those were updated in 2015 by a set of principles for Open Access publisher services (Science Europe, 2015[31])

In 2016, the European Commission published its Open Innovation, open science, Open to the World vision, incorporating its ambitious plans for a European Open Science Cloud (EOSC) (European Commission, 2016[32]). The EOSC aims to provide EU researchers with an environment offering free and open services for data storage, management, analysis and reuse across disciplines, achieved through connecting existing and emerging infrastructures, adding value and leveraging past infrastructure investment. The EOSC is expected to develop common specifications and tools to ensure data are FAIR and legally compliant with the General Data Protection Regulation and cybersecurity legislation of the European Union. It also foresees mechanisms for cost recovery on cross-border access (European Commission, 2018[30]).

The FAIR data principles were developed by a diverse set of stakeholders representing academia, scholarly publishers, industry and funding agencies (Wilkinson, M et al., 2016[33]). They are now becoming a mainstream reference for policy makers. The GO[4] FAIR In particular, the European Commission and EOSC Executive Board published six practical recommendations for implementation of FAIR practice (European Commission and EOSC Executive Board, 2020[34]).

The Research Data Alliance is an international forum that also enables the discussion and sharing of good practices. It was initiated in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards, and the Australian Government's Department of Innovation, with the goal of building the social and technical infrastructure to enable open sharing of data (RDA, 2014[35]). As of November 2020, the RDA had 63 organisational members and 11 000 individual members. RDA operates 99 working groups and interest groups that work on recommendations, which can be adopted as standards, as well as a number of supporting outputs. RDA recommendations address a broad range of issues related to interoperability,

[4] GO stands for 'Global Open'

data citation, data catalogues and work flows for research-data publishing (RDA, n.d.[36]).

The Committee on Data (CODATA) of the International Council for Science (ISC) was founded in 1966 to further the ISC vision of science as a global public good. CODATA promotes the FAIR principles and, more broadly, global collaboration to advance open science, and improve the availability and usability of data in all areas of research. The main priority areas of CODATA are: i) promoting principles, policies and practices for open data and open science; ii) advancing the frontiers of data science; and iii) building capacity for open science, by improving data skills and the functions of the national science systems necessary to support open data (CODATA, 2016[37]).

More recently, a group of national and European Commission funding agencies announced plans to impose Open Access publishing upon the recipients of their research funding under "Plan S" (Schlitz, 2018[38]). This is motivated by a slow-down in the overall progress of open access to publications shown in Figure 1. Both green and gold open access are supported (albeit with a cap on the article processing charges); hybrid is not allowed.

National policies

The 2017 European Commission (EC)-OECD science, technology and innovation (STI) policy survey (EC/OECD, 2018[39]) asked OECD member countries and partner economies to provide information about policy initiatives supporting open science and open access. The 181 policy initiatives cited cover the following areas: support for research infrastructure; national policies and strategies in favour of open access to data (often linked to broader open science strategies or open-government initiatives); creation of governance bodies to foster open access; network and collaborative initiatives aiming to facilitate open access to data (OECD, 2020[40]). Some of these are discussed in in the Annex.

Finland's open science and Research Initiative (ATT) is one of the most mature. It builds upon the previous Finnish Research Data Initiative, and aims to incorporate open science and research in the full research process. ATT covers strategy and policy development, services for researchers, skills development, promotion of interoperability, common infrastructures, enhanced reproducibility of research, and overall openness. It has demonstrated its impact as an accelerator of open science both in Finland and in the international context. It has been able to address a number of issues, such as digitalised services for the research field, creating reference architecture for open science, providing practical guidelines and support for researchers, and creating models and tools for open access and long-term preservation of metadata. In addition, the ATT initiative has provided benchmarking information allowing funding and

research organisations to determine their position regarding open science (Haapamaki, 2018[42]).

Other high-level initiatives – such as Norwegian, Korean and Slovenian strategies, the Netherlands' open science Policy and the UK Concordat – have yet to be implemented, and quantitative impact is therefore expected in the future. To date, they have had more qualitative achievements, such as raising community awareness and initiating dialogue about data sharing among key stakeholders (academia, policy makers, data repositories and the private sector). The UK Concordat concludes it has helped ensure that university leaders are aware of the needs for open research data, its benefits and the key challenge areas. It has ensured initial buy-in and commitment, and resulted in agreed principles for and high-level endorsement of Open Research Data (Bruce, 2018[43]).

In Korea, the debate on open access was linked to unleashing the power of big data and data-driven innovation during the Fourth Industrial Revolution (Shin, 2018[44]). In Argentina, consensus building around the portal has helped partly overcome the cultural barrier against data sharing (Luchilo, D'Onofrio and Tignino, 2018[45]). As pointed out by Slovenia (but valid in most cases), provisions for research data management and openness are not designed as standalone solutions, but are fully aligned with international recommendations. Also, the national approach to designing provisions is more efficient than the introduction of separate provisions by individual research-funding and research-performing organisations (Tramte, 2018[46]). Adoption of a common standard and a technical framework has also been reported by both Mexico (OpenAIRE) (CONACYT, 2018[47]) and Sweden (Generic Statistical Information Model [GSIM]). Sweden reports the GSIM acts as an effective common language of metadata and provides a framework for identifying what parts of the metadata and semantic descriptions need to be further curated, and which parts can be harvested and used "as is" (Eriksson and Nilsson, 2018[48]).

However, Colombia points out the GSIM common standard is a necessary support framework, but was not the strategy underpinning the construction of an initiative such as SiB Colombia (Escobar, Hernández and Agudelo, 2018[49]). As explicitly reported by Colombia (but implicitly stated in many other case studies), a set of principles and values is the basis for the trust necessary to build the initiative. The construction of an open-access ecosystem is a collective exercise: recognition and visibility are the first steps required to generate confidence. The discourse and the language must be coherent with the principles and values, and in turn with the actions. Additionally, it is important to build on the available infrastructure, rather than starting from a blank sheet. The European Open Research Data Pilot shows more quantitative impact, with a relatively high opt-in rate (68%) – demonstrating that researchers are willing to share their data in most cases. The researchers who opted out (32%) quoted

IP of data as the major reason. This finding would need additional analysis, to understand whether it relates to specific projects combining proprietary background knowledge with research financed by Horizon 2020 (European Commission, 2018[30]).

Other initiatives are still in the build-up stage and already have an impact, which can be further developed. For example, the Argentine Science and Technology Information portal launched in 2017 now contains 18 000 projects from the main funding sources, but still needs to include projects from other scientific organisations and universities (Luchilo, D'Onofrio and Tignino, 2018[45]). Mexico reports ten ongoing Open Research Data repositories, with 27 000 research datasets (CONACYT, 2018[47]). Canada's Open Government Portal allows users to search through over 81 000 open data and information from departments and agencies across the federal government (Treasury Board of Canada Secretariat, Open Government Team, 2018[50]). The Swedish infrastructure for register-based research (RUT) is operational, with a limited number of sensitive register datasets. However, the Swedish Research Council, Statistics Sweden and the National Board of Health and Welfare are currently discussing the need to devise technical solutions for data disclosure that can also link register data with big data, as well as the need to expand the dialogue to other actors in the e-infrastructure[5] ecosystem (Eriksson and Nilsson, 2018[48]).

Achievements of open science

Overall, significant progress has been achieved on open science in the past two decades, in a number of OECD member countries and partner economies. Overall awareness of the potential impact of open science has been enhanced, and specific policies have been implemented at national and institutional levels. At least 58 countries have dedicated national strategies and policies for open data and publications, and those that do not have such national policies often have a critical mass of bottom-up institutional policies.

At the level of academic institutions and data repositories, international co-operation is established in the forms of repository networks, for example OpenAire (see Annex), and increasingly comprehensive science clouds are being set up, including the European Science Cloud, the National Research Data Infrastructure in Germany, the Australian and African clouds, the National Institutes of Health Commons in the United States and Research Data Infrastructure for open science in Japan.

[5] ICT-based infrastructure

Box 3. Achievements in the pandemic context

In global emergencies like the coronavirus (COVID-19) pandemic, open science policies can remove obstacles to the free flow of research data and ideas, and thus accelerate the pace of research critical to combating the disease. The full genome of COVID-19 was published barely a month after the first patient was admitted into Wuhan hospital, as an open-access publication in The Lancet (Lu et al., 2020[52]). In the case of the SARS outbreak in 2002-03, this took five months. A large part of this delay is due to retention of information about the disease in the first months of the SARS epidemic.

In January 2020, 117 organisations – including journals, funding bodies, and centres for disease prevention – signed a statement titled "Sharing research data and findings relevant to the novel coronavirus outbreak", committing to provide immediate open access for peer-reviewed publications at least for the duration of the outbreak, to make research findings available via preprint servers, and to share results immediately with the World Health Organization (WHO) (OECD, 2020[51]).

The Open COVID Pledge was launched in April 2020 by an international coalition of scientists, lawyers, and technology companies, and calls on authors to make all intellectual property (IP) under their control available, free of charge, and without encumbrances to help end the COVID-19 pandemic, and reduce the impact of the disease. Some notable signatories include Intel, Facebook, Amazon, IBM, Sandia National Laboratories, Hewlett Packard, Microsoft, Uber, Open Knowledge Foundation, the Massachusetts Institute of Technology, and AT&T. The signatories will offer a specific non-exclusive royalty-free Open COVID license to use IP for the purpose of diagnosing, preventing and treating COVID-19.

Sources: (Lu et al., 2020[52]). (OECD, 2020[51]).

An emerging trend is the establishment of specific governance set-ups, such as the National Chief Data Officer established in France to carry out his missions of co-ordinating government efforts to facilitate the provision, governance, production, circulation and reuse of government data, including research data (Ministère de l'Education Supérieure, de la Recherche et de l'Innovation, 2018[51]). Similarly, Canada established a Chief Science Advisor, who works to ensure that government science is fully available to the public and that government scientists can speak freely about their work (Treasury Board of Canada Secretariat, Open Government Team, 2018[50]).

The role of funding agencies in advancing the open science agenda cannot be overstated. In Europe, projects financed by the Horizon 2020 fund requires data management plans, as well as funders in a majority of European countries (17 out of 31, as of 2016) (European Commission, 2017[52]). At the same time, in the United States, major funders (National Institutes of Health and the National

Science Foundation) had similar requirements, albeit with no common standard for those requirements.

This has resulted in significant progress of Open Access as well as Open Data from publicly funded research. FAIR data is by now a well-established acronym for Findable, Accessible, Interoperable and Re-usable data, moving beyond implementation to adoption and even to becoming common practice.

The future of scholarly communication has started to move from planning to implementation and even adoption of more open practices. As pointed out above, Open Access is making progress: the proportion of Open Access publications is 45% in 2014. The cOAlition S2 consortium, and the associated Plan S will likely push this agenda even further.

Challenges

Overall, access to data currently lags behind access to publications: as of 2017 more than 92% of universities in Europe have open-access policies for publications in place or plan to have them in the near future, fewer than 28% had open access to data guidelines in place (Morais and Borrell-Damian, 2018[54]).

This is clearly not an infrastructure issue: over 83% of institutions either have their own repository or participate in a shared repository; 65% have their repository aggregated by the OpenAIRE portal/infrastructure, which aims to link the aggregated research publications to the accompanying research and project information, to enhance the reproducibility of scientific results. Institutional barriers to promoting research data management include internal factors (e.g. different "scientific cultures"), limited awareness of the benefits of research data and structural elements, such as the absence of policy guidelines at the national level, the lack of incentives to promote research data and increased costs (Morais and Borrell-Damian, 2018[54]).

Box 4. Challenges for open science in a pandemic context

While clinical, epidemiological and laboratory data about COVID-19 are widely available, including genomic sequencing of the pathogen, a number of challenges remain:

- All data is not sufficiently FAIR, resulting in difficulties in interpretation and reuse.
- Sources of data tend to be dispersed, and curation needs to be operated "on the fly".
- Providing access to personal health record sharing is a challenge under current data protection frameworks in most OECD countries.

- In some cases, it could be inferred that the transparency of the statistics may have guided governments to restrict testing in order to limit the number of "confirmed cases" and avoid the rapid rise of numbers.

Concerning open access to publications, challenges also remain:

- The current positive engagements by publishers (Box 3) are set to expire in months, and the sustainability in the long run is uncertain. It also concerns a small core of knowledge directly linked to COVID-19, and fails to open up the broader interdisciplinary knowledge base needed for full understanding of the virus. A recent study shows that less than one-third of the interdisciplinary publications referenced in COVID-19 are open-access (Larivière, 2020[57]).
- Preprints have been encouraged as a vehicle for rapid knowledge diffusion during the crisis, and this has largely proved positive. While preprint circulation allows for increased speed of diffusion, it presents risks of quality control. For example, a paper published on the BioRxiv server on 2 February erroneously asserted that the COVID-19 virus sequence might have been man-made. Luckily, the error was quickly spotted by fellow scientists and the paper was removed within hours.

Sources: (Larivière, 2020[57]); (OECD, 2020[51]).

A recent survey of *Research Assessment in the Transition to open science* revealed that assessment focuses mostly on published research outcomes and attracting funding, and, to a lesser degree, research impact and knowledge transfer. Consequently, evaluation is still largely driven by assessment of research publications, while open science and Access indicators are important or very important to only 28% of respondents, and even then, is mostly monitored at the institutional level, rather than driving incentive and reward structures for researchers (Saenen et al., 2019[55]).

Progress and motivation varies among different disciplines and institutions, among different actors and organisations, and among researchers at different stages of their career. There is also a lack of policy alignment across local, regional, national and international jurisdictions, such as across Member States, and no clear legal or regulatory framework, often associated with insufficient cost/benefit analysis of open science requirements (European Commission, 2020[56]).

Way forward for open science policies

The pioneering impetus of the OECD's Recommendation concerning Access to Research Data from Public Funding helped establish access and global sharing of research data as a major policy priority, with the ultimate goal of making the global science system more efficient and effective.

Box 5. Way forward in the context of the COVID-19 pandemic

- Develop data governance models that allow for open research data by default, while preserving individual privacy. In particular, enabling access to sensitive research data across borders on a more restricted basis in secure environments.
- Provide regulatory frameworks that would enable interoperability within the networks of large electronic health records providers, patient mediated exchanges, and peer-to-peer direct exchanges.
- Public actors, private actors, and civil society to develop and/or clarify a governance framework for the trusted reuse of privately-held research data toward the public interest.
- Secure adequate infrastructure (including data and software repositories, computational infrastructure, and digital collaboration platforms) to allow for recurrent occurrences of emergency situations.
- Ensure that adequate human capital and institutional capabilities are in place to manage, create, curate and reuse research data.
- Establish specific standards for COVID-19 data, in order to facilitate interoperability and reuse across countries and regions.
- Continue the broader discussion about progress towards open-access publishing, including initiatives such as Plan S, an international project that requires all scientific publications resulting from research funded by public grants be available in open access.

Source: (OECD, 2020[51]).

The scientific landscape has changed dramatically in the 15 years since the original OECD Recommendation was adopted in 2006. Data-driven innovation and data-intensive science are transforming society, and access to data has had far-reaching effects on the reproducibility of scientific results, diffusion of knowledge across society, cross-disciplinary co-operation, resource efficiency, productivity and competitiveness. Open science and open data have become mainstream trends, with at least 58 countries adopting dedicated national strategies and policies for open data and publications (EC/OECD, 2018[39]).

Such technological and policy advances required new policy guidance, which is why the OECD Recommendation on research data was recently updated. The revised Recommendation, released on January 20 (OECD, 2021[6]), reaffirms the relevance and importance of several key principles set out in 2006: openness, flexibility, transparency, legal conformity, protection of intellectual property, formal responsibility, professionalism, interoperability, quality, security, efficiency, accountability, and sustainability.

The revised Recommendation expands its scope to cover not only research data, but also related metadata (data about data, specifying their sources, methodology and limitations), as well as the bespoke algorithms, workflows, models and software (including code) that are essential for their interpretation.

Figure 2. OECD's Recommendation concerning Access to Research Data from Public Funding

Areas of policy guidance

***Source*:** (Paic, 2021[52]).

The Recommendation of the Council Concerning Access to Data from Public Funding paves the way forward, along seven priorities (Figure 2):

Data governance for trust: Make digital objects[6] from public funding openly accessible and reusable to the largest extent possible, while taking into account the need to restrict access for legitimate private, public, and community interests. They should be, to the greatest extent possible, accessible online in a timely fashion, findable and user-friendly, provided without discrimination as to the location or nationality of the user, and free of charge. In cases where access needs to be partially or totally restricted, it can be provided within safe and secure environments to certified users with clearance. Risks should be transparently managed and mitigated, especially related to access to sensitive categories of digital objects from public funding. These processes should occur in consultation with communities of stakeholders. Consent or comparable legal

[6] The term "digital objects" is used as shorthand for "research data and other research-relevant digital objects from public funding", and this covers data, metadata, algorithms, workflows, models and software (including code).

basis should be sought consistently, and roles and responsibilities of staff responsible for data access should be clear.

Technical standards and practices:. Foster compliance with technical standards and practices that make digital objects from public funding FAIR. Findability can be improved by assigning unique digital persistent identifiers and publishing descriptive metadata; accessibility will occur through development of infrastructure and services within and across scientific domains and disciplines; interoperability will be enhanced by requiring the use of semantic (including ontologies and scientific terminology), legal (rights of use), and technical (such as machine readability) standards; while re-use in the long term can be facilitated through the provision of high quality human-readable, machine-actionable, and open metadata and adequately maintained and supported bespoke algorithms, code, software, and workflows essential for re-use of data as free and open source. In addition, support is needed for the development, maintenance, adoption, dissemination, and implementation of technical standards that are open, freely accessible, and internationally agreed to the greatest possible extent.

Define responsibility, ownership and stewardship, which needs to be clearly delineated and allocated across the research data ecosystem, while also tailoring and implementing licensing to optimise scientific discovery and innovation and protect research data and digital object producers' rights. In particular, good practice for research data and software management needs to be adopted across the research system; access to digital objects resulting from public-private partnerships needs to be as open as possible while recognizing the legitimate interests of the private-sector partners. New uses of digital objects from public funding, such as for artificial intelligence and text- and data-mining techniques need to be made possible, and information about rights and licensing should be accessible in the metadata. The widest possible use of open licences should be encouraged, where these are appropriate.

Incentives and rewards. Effective models of reward and recognition need to provide incentives and remove disincentives for researchers and research support staff to provide access to digital objects from public funding. These should recognise and reward the provision of access as a recognised research output within researcher recruitment, advancement, and grant review processes, be based on robust and open indicators on the impact of access, including data and software citations, contributor taxonomies and dissemination of research outputs beyond publications and acknowledge data and code creators and maintainers as key contributors. Researchers may require a reasonable limited period of exclusive use of the digital objects they produce, for example to provide time for data analysis and preparation of final results and/or intellectual property claims.

Sustainable infastructures need to be developed and maintained to support the findability, accessibility, interoperability, and reusability of digital objects from public funding free of charge at the point of use. In particular, governments should develop strategies to ensure sustainable infrastructures, including data and software repositories and services, which would prioritise digital objects from public funding for short-, medium-, or long-term preservation, improve interoperability among global research infrastructures to leverage national investments and innovation, and to encourage interdisciplinarity, safeguard digital objects determined to be of high-value in the long term, including through provision for maintenance and support in cases the specific infrastructure ceases to exist; and ensure an appropriate match between funding instruments, review criteria regarding the need for long term preservation of digital objects , and the expected longevity of infrastructure. Private investment in research data infrastructures needs to be encouraged, with investment in the skills needed to manage and use them, while taking measures to facilitate their openness, reliability and integrity, and to protect the public interest over the long term by avoiding vendor lock in and ensuring data portability.

Human capital is necessary to realise the potential benefits of enhancing access to research-relevant digital objects from public funding. Strategies are needed to develop skills necessary for data-driven research and innovation, both at a basic competence level for all researchers and students as well as a cadre of dedicated data managers, stewards, and research software engineers, and cohorts of researchers with advanced (e.g., PhD-level) data-intensive research and data science skills. Likewise, it is important to improve understanding, by relevant policymakers and research management staff, of approaches for effective management of digital objects and ensure data literacy skills by citizens to enable them to make effective use of research data, and develop appropriate learning and training programs and resources to support these objectives. Data scientists and research software engineers need to be attracted and retained across the breadth of scientific disciplines, notably through attractive career paths, including facilitation of transfers between scientific disciplines, as well as careers combined with private sector experience; and recognition and reward of data management skills and software development skills as high value added to publicly funded research and innovation..

International co-operation for access to research data is needed in order to enable free exchange of ideas and enhance scientific discovery, notably where making use of datasets across borders bilaterally or multilaterally can help the advancement of science and contribute to solving global societal challenges. For this, we need to develop common definitions, data and security standards, and certification processes to enhance access, including access to sensitive data, across national borders. With specific regard to sensitive data, including

personal data, we need to explore the interoperability of legal and ethical frameworks to enhance data access across borders while protecting legitimate private, public, or community interests; and work towards developing internationally compatible procedures for: (i) determining the degree of sensitivity of data; (ii) establishing criteria and protocols for the certification of institutions and researchers gaining access to such data; and (iii) establishing technological standards and approaches for secure remote access to such datasets. In addition, we should seek to achieve synergies in data infrastructure through relevant and appropriate funding, governance and collaboration schemes for international data infrastructure, and co-operate to build a digitally-skilled research workforce through exchange programmes.

Annex. A selection of Open Science Initiatives

Open Access to Publications

Sponsoring Consortium for Open Access Publishing in Particle Physics (SCOAP3), International

Responsible entity: CERN

Description: SCOAP3 is a global partnership of over 3 000 libraries, funding agencies and research centres of 44 countries and three intergovernmental organisations. It focuses on high-energy physics and grants Open Access to key journals without authors incurring any fees. It reimburses publishers for costs arising out of Open Access whilst the publishers lower their subscription fees. As all publishers reduce subscription fees, the resulting funds are pooled in a common fund out of which publication costs are financed. The SCOAP3 journals are accessible for any researcher to publish their results and copyrights will remain with the authors. Further, the repository is placed on top of the CERN Invenio digital library open source platform. Since 2014, over 38 000 articles have been funded by SCOAP3. For its third phase of 2020-2022, it expects a record number of articles (23 000) with a total contract value of more than 29 million Euros.

Read more: *https://scoap3.org.*

INSPIRE-HEP, International

Responsible entity: CERN, DESY (German Electron Synchrotron), Fermilab, SLAC (Stanford National Accelerator Laboratory)

Description: Launched in 2010, INSPIRE-HEP is an Open Access digital library in the field of high-energy physics. It succeeds the 'Stanford Physics Information Retrieval System (SPIRES)' database, the leading literature platform for high-energy physics since the 1970s. It is a collaborative project of four laboratories, namely CERN, DESY, Fermilab and SLAC. It provides scientific papers as well as citation metrics, extracts from papers or internal experiments and tools to facilitate metadata improvements such as for purposes of author disambiguation. Besides literature, it provides databases on key researchers involved in high-energy physics as well as critical institutions, conferences, job vacancies and experiments. It encompasses more than 500,000 high-energy physics related articles, including journal papers, preprints, e-prints, technical reports, conference papers and theses, comprehensively indexed by the SLAC and DESY libraries since 1974 (Princeton, 2020).

Read more: *https://inspirehep.net.*

DataCite consortium, International

Responsible entity: DataCite

Description: Founded in 2009 by research organisations, universities and libraries of six countries, DataCite is a not-for-profit organisation committed to enhancing the quality of data citation. It follows the FAIR principle and strives for better data citation which is needed to improve and facilitate access to digital research data, develop standards, enhance acceptance around validity of digital data as well as to foster data archiving which addresses verification and repurposing of data and results. As such, it serves as a data registry of Digital Object Identifiers (DOIs) and assigns them to datasets. Further, it thereby it simplifies data management and provides a platform which grants access to all metadata of submitted articles and datasets.

Read more: *datacite.org.*

Persée portal, France

Responsible entity: French Ministry of National Education, École Normale Supérieure de Lyon, French National Centre for Scientific Research, University of Lyon

Description: Developed in 2005, the Persée portal is a digital library which offers open access to academic journals, notably in French, making it one of the largest francophone human and social sciences portals. It offers free access to more than 700,000 documents from 300 collections covering a wide range of disciplines within the humanities and social sciences. Its objective lies in the sharing of knowledge, to promote scientific heritage for the benefit of research and innovation by offering free access, open source and standards, interoperability and reuse of data.

Read more: *www.persee.fr.*

Japan Science Technology Information Aggregator, Electronic (J-STAGE), Japan

Responsible entity: Japan Science and Technology Agency (JST)

Description: The J-STAGE platform offers electronic collections of Japanese academic journals, proceedings and reports from different Japanese scientific associations. Its principles are free access (over 90% of articles), reuse and redistribution. Its objectives lie in enhancing and supporting the submission of manuscripts, peer-reviewing, page lay-outs and dissemination of electronic journals. It promotes communication and internationalisation of Japanese science and technology publications. As of November, 2020 it covers 25 subject areas, has published more than 2 000 journals. J-STAGE also entails the

Journal@rchive initiative of 2002 which aims at electronically archiving less recent and historically significant publications in print.

Read more: *https://www.jstage.jst.go.jp/browse/-char/en.*

Open Access policy for public sector research funds and foundations, Denmark

Responsible entity: Independent Research Fund Denmark, Innovation Fund Denmark and Danish National Research Foundation

Description: In order to strengthen the impact of research, the Open Access policy's objectives lies in open and free access to publicly financed research results for all. It offers access to research findings in publications in journals and conference proceedings. Originally developed in 2012, Denmark's National Strategy for Open Access of 2018 resulted in stricter rules and requirements for Open Access. Partial or full grant holders must ensure free access to published articles via Open Access and with the permission of the magazine. It is required that they publish a parallel digital version of the final, peer-reviewed scientific article (Ministry of Higher Education and Science, 2019).

Read more: *https://ufm.dk/en/research-and-innovation/cooperation-between-research-and-innovation/open-access/Publications/councils-and-foundations.*

Plan S, European Union

Responsible entity: cOAlition S, supported by European Commission and European Research Council (ERC)

Description: Established as a joint effort by twelve European countries in 2018, Plan S is an initiative engaged in the promotion of open access publishing. It sets out an action plan involving free and open access to scientific publications which were financed with public grants in respective open access journals and platforms as of 2021. It outlines 10 principles required to accomplish this objective: authors are meant to obtain copyrights with an open license, preferably the Creative Commons Attribution license (CC BY) in line with the Berlin Declaration. Secondly, funders are to specify criteria on the basis of which high quality open access journals, platforms and repositories operate. Further, the funders intend to establish new open access journals of high quality where appropriate as well as necessary open access infrastructure. Open access fees are to be provided by funders and research institutes as opposed to the individual researcher whist ensuring transparency of fees. In addition and also for transparency purposes, policy-makers, universities, research organisations etc. are encouraged to streamline decision-making. Moreover, monitoring and compliance management will fall under the responsibility of the funders. Finally, the funders are required to provide

evaluation of research outputs solely based on the intrinsic value of research, irrespective of publication channel or publisher.

Read more: *www.coalition-s.org/why-plan-s/.*

Open Access to Data

Zenodo, international

Responsible entity: CERN, European Commission

Description: Developed under the OpenAIRE program by the European Commission in 2013, Zenodo is a digital platforms which collects research outputs and databases and provides open and free access for anyone to use. Co-funded by the European Commission and under operational leadership of CERN, which is an OpenAIRE partner and primary actor in open source, access and data, Zenodo enables the distribution and sharing of reports, data sets and research software. It is an all-encompassing tool which can be used by researchers worldwide, welcoming outputs from every discipline. In March 2020, Zenodo launched a Coronavirus Research Community - COVID-19 accepting data from all scientific disciplines and sub-disciplines. Additionally, it welcomes all formats of information such as presentations, audio/video, posters, publications, datasets etc. (CERN, 2020).

Read more: *https://zenodo.org.*

European Open Science Cloud, EU

Responsible entity: joint effort by the European Commission, the European Bioinformatics Institute of the European Molecular Biology Laboratory(*EMBL-EBI*), the *Elixir* infrastructure and the *COMPARE project*, as well as the EU Member States.

Description: The European open science Cloud (EOSC) aims to provide a platform to host and process research data in line with supporting the EU's global lead in science. Based on current and evolving data frameworks it offers European researchers a seamless and secure-access infrastructure to store, manage and process data from different sources. Accordingly, it underpins the open science, Open Innovation and Open to the world policies by delivering the following cornerstones: improving global data findability and accessibility (FAIR data); fostering skills recognition of researchers (careers); addressing concerns revolving around data subject privacy, access and intellectual property rights; facilitating replicability of findings and reduce wastage of data, e.g. clinical trial data (research integrity); helping elucidate funding models for the generation and preservation of data; as well as limiting rent-seeking and fostering innovative research services (new business models) (European Commission, 2018).

Through the European COVID-19 Data Platform in the context of EOSC, researchers can store and share data sets, in a GDPR-compliant manner, including of DNA sequences, protein structure, (pre-) clinical trial and epidemiological data. As of Sep. 2020, more than 73 000 users have participated in sharing and uploading of data which is typically very difficult to find, such as raw and assembled viral sequences to combat the coronavirus pandemic. 2 million requests have been made since its launch on April 20th (EOSC Portal, 2020).

Read more: *www.eosc-portal.eu.*

OpenAIRE, EU

Responsible entity: European Commission

Description: The Open Access Infrastructure for Research in Europe (OpenAIRE) was launched in 2009 to promote and adopt the Open Access Policy with the objective to accelerate research and strengthen innovation in the EU. It comprises a pan-European network, National Open Access Desks (NOADs) consisting of local open science experts across the EU working within their member state and internationally to support researchers in adopting open science practices. The OpenAIRE platform makes research accessible, re-usable and facilitates project reporting as it links research outputs to funding information. OpenAIRE is funded under the EU Horizon 2020 program for Research and Innovation and collaborates closely with other EU initiatives such as the European open science Cloud. The OpenAIRE Covid-19 Open Research Gateway serves as a single access point for the discovery and coordination of research outcomes (publications, data, software) on COVID-19. It will gather disease-related data around the virus as well as for future research on long-term social implications.

Read more: *www.openaire.eu.*

Data61, Australia

Responsible entity: Commonwealth Science and Industrial Research Organisation (CSIRO)

Description: CSIRO's Data61 is a leading data innovation network in Australia which draws from employees and researchers of 70 countries and 30 university partners. Established in 2016, its mission is to explore the impact of the digital-era's disruption in the various sectors of its researchers by adopting a 'Digital+Domain' focus which aims to provide new interdisciplinary solutions to structural change arising of the data-driven economy. It specifically targets large-scale projects with global outreach in different fields including cybersecurity, digital regulation, shared data and data privacy, personalised medicine and wellness, food provenance, biosecurity, the response to the

COVID-19 pandemic, smart cities initiatives as well as storable renewable energies. Its ultimate objective is to support Australia's structural transformation from an industrial, resource-based economy towards a sustainable knowledge economy. It intends to reach this achievement with strategic partnerships with governments, students and the private sector to provide cutting-edge research for global challenges.

Read more: *https://data61.csiro.au.*

Leveraging AI to tackle public health challenges, Canada

Responsible entity: Canadian Institutes of Health Research (CIHR)

Description: With the purpose of enhancing COVID-19 research and faster knowledge absorption, the Canadian Institutes of Health Research (CIHR) has used its investments, expertise and connections in various fields including knowledge mobilisation, open science, patient-oriented research and high-priority evidence reviews. One resource provided by CIHR for COVID-19 research is the 'Linkage Tool for the COVID-19 Rapid Response Funding' which constitutes a database for researchers, individuals and organisations to share data related to the coronavirus. In addition, CIHR provides funding for research outputs and other initiatives. As such, the 'Strategy for patient-oriented research' (SPOR) helps to connect researchers, health administration and policy makers nationwide to address specific evidence needs and to avoid duplication of work concerning the virus. 300 distinct COVID-19 related activities have been reported, for instance guidelines on preventing infections in people aged 60 and above to advise the World Health Organisation on the protection of residents in long-term care facilities.

Read more: *https://cihr-irsc.gc.ca/e/52025.html.*

Connecting the Czech Republic, Czech Republic

Responsible entity: Czech Ministry of Industry and Trade, CzechInvest in cooperation with the Ministries of Health, Defence and Transport, the Office of the Government of the Czech Republic, the CzechTrade Agency, the Confederation of Industry of the Czech Republic Digital, a number of open data experts, universities, and communities.

Description: Amid the COVID-19 pandemic, the 'Connecting Czech Republic' (SPOJUJEME ČESKO) platform has enabled online communication of employees, to collect data on combating the coronavirus, the Buy Safely service improving food safety as well as funds collection for ventilators. The platform is based on collaboration of the state, scientists, universities, communities, start-ups, NGOs, large companies and SMEs. This continuously updated information system helps address the pandemic by giving rise to various initiatives as a direct outcome of this fostered cooperation.

Read more: *https://spojujemecesko.cz.*

GISAID Initiative, Germany

Responsible entity: Federal Republic of Germany

Description: Founded in 2008, The GISAID initiative, formerly Global Initiative on Sharing All Influenza Data initiative, is a global science initiative endorsing open-access with focus on genomic data of influenza viruses including the COVID-19 virus. It offers free and publicly available virus data whilst securing legal certainty around the use of such data. Virus sequence data as well as related clinical and epidemiological data are shared in a timely manner, even prior to publication. It proactively supports novel research tools for the analysis of influenza viruses by facilitating the integration of different tools to examine GISAID data. As a global platform it relies on public-private partnerships with Brazil, Singapore, the USA, Sanofi and Seqirus. The importance of the GISAID initiative to combat global health crises has been acknowledged by the G20 health ministers in 2017 as part of the Berlin declaration.

Read more: *https://www.gisaid.org.*

Korean Bioinformation Centre (KOBIC)'s COVID-19 Research Information Portal, Korea

Responsible entity: Korean Bioinformation Centre (KOBIC)

Description: As the national research centre in the fields of genomics, proteomics, systems biology and personalised medicine, KOBIC is in charge of coordinating data and bioinformation from various laboratories and institutions across South Korea. It provides a central data portal to facilitate knowledge sharing between research groups. Its COVID-19 Research Information Portal offers information regarding COVID-19 but also other epidemics such as SARS, MERS etc., notably through dielectric data, biomechanics, the virus' protein structures and genome mutation information. It thus provides a database highly needed amid fast-paced and time-sensitive COVID-19 research. The objective lies in supporting scientific publications which focus on diagnostic tests, vaccines and treatments.

Read more: *www.kobic.re.kr/covid19/.*

SURF, the Netherlands

Responsible entity: SURF Foundation

Description: SURF is a co-operative association of Dutch educational and research institutes aiming to maintain a network for collaborative innovation by ensuring free access to the most recent ICT facilities for cutting-edge research and talent development. The member institutions engage in

cooperative purchase or development of digital services and solutions to further knowledge sharing in the research and scientific community of The Netherlands. To fully benefit from increasing amounts of data which become available, the strategy relies on unlimited open access, data security and well-integrated data, computing and network facilities. Further, research institutions are encouraged to strengthen international partnerships and promote mutual use of infrastructure and services. For these purposes SURF is proactively involved in identifying the research community's needs and to engage in policy development in this regard, e.g. through cooperation with international organisations.

Read more: *www.surf.nl/en.*

Horizon 2020 Open Research Data (ORD) Pilot and Data Management Plan (DMP), EU

Responsible entity: The European Commission, The Commissioner for Research, Science and Innovation, The Directorate General for Research and Innovation

Description: Aimed at the beneficiaries of H2020 grants, this initiative seeks to promote open access and re-use of research data coming from Horizon 2020 projects as well as to implement data management as part of the research process. Further, it strives to consider the importance of protection of intellectual property and personal data while advocating for open access. Targeted at beneficiaries coming from the scientific community, civil society and the corporate sector, it covers all thematic areas of Horizon 2020 research and innovation projects. Based on the 'FAIR' principle, researchers are therefore encouraged to ensure their data is findable, accessible, interoperable and re-usable as well as appropriate management of such data during and after the project. These objectives are contributing to the European Commission's pursuit of a systemic change towards Open Science which builds upon a networked and collaborative approach to research and open access. Open access in the framework of research and development refers to peer-reviewed scientific publications and its underlying data. More specifically, open access to research data is expected to improve transparency, accountability, the quality of scientific research through re-use of data and thereby innovation, to the benefit of the private sector at large. The ORD pilot serves as a regulatory guideline to make open data the default option, requiring H2020 participants to submit a DMP in which data management, collection and use are outlined. Between 2014 and 2016 around two thirds of beneficiaries adhered to the rules while one third chose to opt out, most notably for IPR purposes.

Read more: *https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm.*

Finnish Open Science and Research Initiative (ATT), Finland

Responsible entity: Ministry of Education and Culture, Academy of Finland and Tekes, CSC – IT Center for Science

Description: The key objective of the Finnish Open Science and Research Initiative (ATT) is to spur a national shift towards open science by making science and research more reliable, fostering open science principles and through enhancement of the social impact of science and scholarship. This initiative was running from 2014 – 2017 and targeted researchers, research administrators, policy makers and funding institutions among others. Subsequently, the framework was transposed to the Federation of Finnish Learned Societies which is in charge of coordinating the national open science action plan. The initiative was based on three complementary aspects: research publications, research data and research methods. Further, its aim was to integrate open science and research into the whole research process in order to determine the visibility and impact of science and research in the innovation sphere and society more broadly. By fostering transparency and a collaborative approach to research it followed the objective to render Finnish research more competitive and innovative. As a consequence, these efforts facilitate knowledge generation by research institutes and universities and accessibility to society. Highlighted components include data governance to create trust, IPR, findability, machine readability of data as well a reward systems for data authors, albeit not in effect yet. In addition, the initiative helped competences of public authorities to manage, create and re-use data.

Read more: *https://avointiede.fi/en.*

Mexican Open Science Policy: Open Institutional Repositories Program (CONACYT), Mexico

Responsible entity: National Council of Science and Technology

Description: Initiated in 2015, the Open Institutional Repositories Program CONACYT is a national policy under the Mexican Law for Science and Technology developed in reference to the Budapest Open Access Initiative and the OECD Principles and Guidelines to Open Science among others. The program focuses primarily on public sector institutions of the scientific community and their research data, even though private sector institutions are also targeted, this far there has been no request to participate. The objective is to collect, preserve and secure Open Access to publicly funded science, technology and innovation data as well as to attain respective databases. As part of this initiative the aim was to develop more than 10 institutional data repositories which are connected to the national repository since 2019. Further, it enhanced institutional capacities for research data management and more than 20,000 data sets are openly accessible. A key achievement has been the establishment

of a single national technical framework (OpenAIRE) for open research data management, while more than 10 data repositories with more than 27 000 data sets are in progress.

Read more: *www.conacyt.gob.mx.*

National Plan Open Science (NPOS), the Netherlands

Responsible entity: Ministry of Education, Culture and Science (OCW), co-signed by several research associations

Description: Under the motto 'as open as possible, as closed as necessary' the Dutch government has enacted the open science and open access norm in academic research. In 2017, the NPOS was signed and focuses on the following four main aspects: *100% open access publishing* ensures that all publicly financed scientific publications will be accessible to all, worldwide, for consultation and data re-use purposes as of 2020. Through *Optimal (re-) use of data* the initiative seeks to induce exchange of research data among scientists based on universal data standards. *Recognising researchers. Corresponding evaluation and valuation systems* aims at shifting the valuation focus away from the number of publications in highly reputable journals towards adoption of a wider set of criteria including quality of education, valorisation, leadership and good data stewardship. *Encouraging and supporting Open Science* calls for a coordinated approach to proactively support open science and the National Plan Open Science is meant as a lead initiative for other countries in this regard. Further, under this initiative the 'roadmap to open access 2018-2020' foresees several key measures for further progress on open access. As such, reading contracts of universities with leading publishers are to be made openly accessible without further costs, applicable to all disciplines. In order to increase their bargaining position for this matter, universities need to look for alternative publication platforms other than the established publishing journals. Additionally, the Dutch government strives for a European lobby on open access to foster international cooperative efforts to exert pressure on publishers. Ongoing search for alternative monitoring of open access publications and archiving so as to retain some of the copyrights as opposed to transferring them to publishers are identified as further points to fully achieve open access targets.

Read more: *www.openscience.nl/en.*

AMED support for R&D on the coronavirus disease, Japan

Responsible entity: The Japan Agency for Medical Research and Development (AMED)

Description: The Japan Agency for Medical Research and Development (AMED) promotes integrated research, from basic research to practical application. In

view of Japan's ageing society, AMED has the objective to generate the world's most advanced medical technologies and services in order to ensure health and longevity for Japanese lives. It therefore emphasises medicine, drugs and medical devices as strategic industries for policy-making. Accordingly, to promote industrialisation AMED fosters research institutions' capacity to acquire intellectual property rights, enhance cooperation among companies in Japan and joint research with international organisations involved in medical R&D.

In January, 2020 AMED signed a joint statement committing them to adopt data sharing policies and urging researchers to share their data. The statement called for immediate and free access to relevant research findings in academic publications, also via open platforms before peer-review. It also incorporated measures to be taken by AMED, including development of a research infrastructure for drug discovery related to infectious diseases, diagnostic kits, therapeutic methods and vaccine development (AMED, 2020).

Read more: *www.amed.go.jp/en/news/topics/covid-19.html.*

Strategy to Promote Sharing and Use of Research Data for Innovative Growth, Korea

Responsible entity: Ministry of Science and ICT (MSIT), Korea Institute of Science and Technology Information (KISTI), in close collaboration with other public research institutes

Description: As an early adopter of e-government initiatives since the early 2000s, South Korea is a leader in open government among OECD member countries. Starting in 2017, the *Strategy* led by MSIT fosters the sharing and usage of research data of publicly funded research projects with the objective to promote innovation at national level. This *Strategy* encompasses policy measures such as investment plans and legislative amendments to increase awareness of the value of research data as well as to induce scientific and socio-economic efforts for new findings out of shared data. The initiative is targeted at researchers, businesses (in particular SMEs), civil society and the public. Its key policy aspects include the support for a research data centre for each field of research, the establishment of a national platform to connect these individual data centres, education and training for data scientists, a legal basis for sharing and use of data as well as data management plans, and to promote innovative capacities arising out of these shared research data. As part of this action plan, the *Strategy* makes use of five sub-strategies, namely the issuance of basic principles for data sharing, the promotion of community- and capacity-building processes, a digital infrastructure with additional support to reduce the researcher's burden, the prioritisation of data-sharing strategies in data-intensive fields first and pilot projects to show-case the potential value of

data-sharing (i.e. projects on genome data, future chemicals and artificial intelligence data).

Read more: *www.msit.go.kr/web/msipContents/contentsView.do?cateId=_policycom2&artId=1383763.*

Infrastructures for Register-based Research – a government commission to the Swedish Research Council, Sweden

Responsible entity: Ministry of Education and Research, The Swedish Research Council

Description: In 2012, the Swedish Research Council was assigned this commission to improve access to register data for research. Public services in Sweden have been accumulating data as part of data registers, e.g. to optimise health care quality, which have come to be useful databases for interdisciplinary researchers to draw from. The initiative, originally envisioned for four years which has since been made permanent, aims at improving the quality of registers, databases, sample collections and research datasets. In particular by enhancing metadata and semantics it is expected to support dissemination of knowledge, re-use of data as well as to significantly reduce time needed to obtain access to register data. Furthermore, it foresees active support in metadata management and the curation of such metadata through education and training initiatives. The gathered data is collected based on the FAIR principles and comes from higher education institutions and businesses conducting research with register data. The initiative has acted as a frontrunner among Nordic countries and induced collaborations with NordForsk funded projects.

Read more: *www.vr.se/english/mandates/research-infrastructure/register-based-research.html.*

Concordat on Open Research Data, United Kingdom

Responsible entity: UK Research and Innovation, Universities UK, Wellcome Trust

Description: Developed in 2015, this policy seeks to achieve that the collected and generated data of the UK research community's members is openly accessible so that it may be (re-)used by others while complying with legal, ethical, disciplinary and regulatory standards and remaining respectful of costs. Large potential benefits are expected from the re-use of research data and results, notably on economic growth, enhanced resource efficiency, safeguarding public support for publicly financed research as well as increasing public trust in research. It considers legitimate reasons for confidentiality such as IPR, privacy and security concerns, therefore the intention of the concordat

was to develop 10 principles and standards which are suitable for all parties involved. These principles state that transparency is critical and data collection/ curation should follow the FAIR principles as well as acknowledge the importance of data citation and credit recognition. However, it considers that appropriate skills acquisition needs to be fostered as data collection and generation require a different skillset than the curation and archiving of such. As a consequence, it calls upon all stakeholders to ensure that skills and capacity building through training opportunities and the funding of such is imperative. The policy also aims to assign clear roles to each stakeholder to improve accountability and respects researchers' autonomy in that it places particular emphasis on situation-specific circumstances and rejects a 'one size fits all' approach. While not compulsory, the policy outlines expectations as regards good practice of data management and dispersion. Most importantly, it aims to support and affect cultural change towards open data, open access and open science although it is acknowledged that this implementation takes time. In addition, the concordat has the objective to secure the UK' leadership role in research as it becomes increasingly international and for this purpose, seeks to alleviate barriers to cross-border co-operation.

Read more: *www.ukri.org/about-us/policies-standards-and-data/good-research-resource-hub/open-research/.*

Federal Data Strategy to provide Open Access to research, United States

Responsible entity: U.S. government

Description: In order to increase access to publicly funded research data and results the U.S. government gathers all U.S. government open data resources in a single-point of access at its website data.gov. Through its Federal Data Strategy it seeks to provide a coordinated effort to enhance open access of data with the ultimate goal of fostering scientific progress, innovation and international collaboration in science and technology. The strategy is based on four principles including 'Enterprise Data Governance' which specifies priorities for government data management as a strategic asset and assigns roles for data privacy, security and confidentiality protection as well as for monitoring of compliance with standards. 'Access, Use and Augmentation' refers to increased data access by ensuring its timely availability in more effective formats, maximising the amount of publicly accessible data which is non-sensitive and making use of technologies to provide open access while respecting privacy and security concerns. The 'Decision-making & Accountability' principle foresees that public authorities publish management expenditures and performance indicators to improve accountability while also informing policy-making through research data on the effectiveness of the government's decisions. Lastly, 'Commercialisation, Innovation and Public Use' envisions to make

federal data accessible to external stakeholders for commercial purposes so as to increase innovation and market efficiency. In addition to the Federal Strategy, the White House Office of Science and Technology Policy (OSTP) has adopted several directives which for instance instruct more than 20 research-funding government agencies to provide strategies in order to maximise public access to peer-reviewed publications and to increase findability and re-usability of research results. In 2017, these agencies accounted for more than 110 billion USD in research funding and have adopted various strategies, although they typically require adherence to data management plans with specific evaluation plans. To monitor and improve coordinative action in this regard, a Sub-Committee on Open Science has been formed in 2018 to ensure preservation, findability, accessibility and usability of publicly funded research data and results. As of February 2019, more than 300 000 datasets were listed on data.gov.

Read more: *data.gov.*

References

Bello and Galindo-Rueda (2020), "Charting the digital transformation of science", *OECD Science, Technology and Industry Working Papers*, https://doi.org/10.1787/18151965. [27]

Berlin Declaration (2003), *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*, https://openaccess.mpg.de/67605/berlin_declaration_engl.pdf. [29]

Berners-Lee, T. et al. (1992), "World-Wide Web: The Information Universe", *Electronic Networking*, pp. 52-28, https://www.legalanthology.ch/t/berners-lee-et-al_world-wide-web_1992.pdf. [22]

Boselli, B. and F. Galindo-Rueda (2016), "Drivers and Implications of Scientific Open Access Publishing: Findings from a Pilot OECD International Survey of Scientific Authors", *OECD Science, Technology and Industry Policy Papers*, No. 33, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jlr2z70k0bx-en. [17]

Bruce, R. (2018), "The concordat on open research data", https://community.oecd.org/servlet/JiveServlet/downloadBody/149486-102-1-264203/OECD-open-access-to-data-UK-Case-Study-v01_02%20(002).pdf. [44]

Budapest Open Access Initiative (2002), *Budapest Open Access Initiative*, https://www.budapestopenaccessinitiative.org/read. [28]

CODATA (2016), *CODATA's Mission*, http://www.codata.org/about-codata/our-mission. [39]

CONACYT (2018), *Mexican Open Science Policy Case Study: Open Institutional Repositories Program*, https://community.oecd.org/community/cstp/enhanced-data-access. [48]

EC/OECD (2018), *Science, Technology and Innovation Policy - STIP Compass*, https://stip.oecd.org/stip.html. [41]

Eriksson, M. and M. Nilsson (2018), *OECD Case study report RUT*, https://community.oecd.org/community/cstp/enhanced-data-access. [49]

Escobar, D., A. Hernández and M. Agudelo (2018), *SiB Colombia: "Enhanced Access to Public Data for Science, Technology and Innovation"*, https://community.oecd.org/community/cstp/enhanced-data-access. [50]

European Commission (2020), *Progress on Open Science: Towards a Shared Research Knowledge System*, http://dx.doi.org/Doi:10.2777/00139. [59]

European Commission (2018), *Case Study of Policy Initiative for Open Access to Research Data: Horizon 2020 Open Research Data (ORD) Pilot and Data Management Plan*, https://community.oecd.org/community/cstp/enhanced-data-access. [32]

European Commission (2017), "Access to and preservation of scientific information in Europe - Report on the implementation of Commission Recommendation C(2012) 4890 final", https://op.europa.eu/en/publication-detail/-/publication/676f8a3b-62f6-11e8-ab9c-01aa75ed71a1/language-en. [55]

European Commission (2016), *Open innovation, open science, open to the world – A vision for Europe*, EU Publications, https://publications.europa.eu/en/publication-detail/-/publication/3213b335-1cbc-11e6-ba9a-01aa75ed71a1. [34]

European Commission (2014), "Validation of the results of the public consultation on science 2.0: Science in transition", http://ec.europa.eu/research/consultations/science-2.0/science_2_0_final_report.pdf. [14]

European Commission and EOSC Executive Board (2020), *Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group*, EU Publications, http://dx.doi.org/doi:10.2777/986252. [36]

European Commission; PWC (2019), *Cost-benefit analysis for FAIR research data*, https://op.europa.eu/en/publication-detail/-/publication/d375368c-1a0a-11e9-8d04-01aa75ed71a1/language-en. [10]

Fuchs, C. and M. Sandoval (2013), "The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious", *tripleC*, pp. 428-443, https://openaccess.city.ac.uk/id/eprint/3078/1/. [9]

Haapamaki (2018), *Report of the Finnish ATT Initiative*, https://community.oecd.org/docs/DOC-149009. [43]

Hey, Tansley and Tolle (2009), *The Fourth Paradigm: Data-intensive scientific discovery*, https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/. [2]

Hoorn, E. (2014), "Diamond open access and open peer review: An analysis of the role of copyright and librarians in the support of a shift towards open access in [19]

the legal domain.", *Open Journal Systems*, Vol. 20/1, https://www.rug.nl/research/portal/files/12356585/Diamond_open_access_and_open_peer_review.htm.

Houghton, J. (2014), *Open research data report to the Australian National Data Service (ANDS)*, https://www.ands.org.au/__data/assets/pdf_file/0019/393022/open-research-data-report.pdf. [12]

Houghton, J. and P. Sheehan (2009), "OECD Science, Technology and Industry Policy Papers", *Economic Analysis and Policy*, Vol. 39, https://doi.org/10.1016/S0313-5926(09)50048-3. [23]

Kindling, M. (2017), "The landscape of research data repositories in 2015: A re3data analysis", *D-Lib Magazine*, Vol. 23/3/4, http://dx.doi.org/10.1045/march2017-kindling. [26]

Korsmo, F. (2010), "THE ORIGINS AND PRINCIPLES OF THE WORLD DATA CENTER", *Data Science Journal*, Vol. 8, https://www.jstage.jst.go.jp/article/dsj/8/0/8_SS_IGY-011/_pdf. [21]

Larivière, V. (2020), *The Coronavirus (COVID-19) outbreak highlights serious deficiencies in scholarly communication*, https://blogs.lse.ac.uk/impactofsocialsciences/2020/03/05/the-coronavirus-covid-19-outbreak-highlights-serious-deficiencies-in-scholarly-communication/. [57]

Luchilo, L., M. D'Onofrio and M. Tignino (2018), *Case study: The Argentine Science and Technology Information Portal*, https://community.oecd.org/community/cstp/enhanced-data-access. [46]

Lu, R. et al. (2020), "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding", *The Lancet*, Vol. 395, pp. 565-74, https://www.thelancet.com/article/S0140-6736(20)30251-8/fulltext. [52]

Machado, J. (2015), *Open Data and Open Science*, Brazilian Institute for Information in Science and Technology. [15]

Ministère de l'Education Supérieure, de la Recherche et de l'Innovation (2018), *Ouverture des données publiques et de recherche. De l'administrateur général des données (national Chief Data Officer).) au plan national pour la science ouverte.*, https://community.oecd.org/community/cstp/enhanced-data-access. [54]

Morais, R. and L. Borrell-Damian (2018), *2016-2017 European Universtites Association Open Access Survey results*, European University Association, [56]

http://www.eua.be/Libraries/publications-homepage-list/open-access-2016-2017-eua-survey-results (accessed on 19 June 2018).

Netherlands Ministry of Education, Culture and Science (2018), "The Netherlands – National Plan for Open Science", https://community.oecd.org/servlet/JiveServlet/downloadBody/149103-102-1-263392/The%20Netherlands%20-%20case%20study%20for%20open%20access%20to%20data%20for%20STI%20-%20for%20OECD.....pdf. [4]

OECD (2021), *OECD Recommendation of the Council Concerning Access to Data from Public Funding*, OECD Publishing, Paris, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347. [6]

OECD (2020), *Enhanced Access to Publicly Funded Data for Science, Technology and Innovation*, OECD Publishing, Paris, https://dx.doi.org/10.1787/947717bc-en. [3]

OECD (2020), *Enhanced Access to Publicly Funded Data for Science, Technology and Innovation*, OECD Publishing, Paris, https://dx.doi.org/10.1787/947717bc-en. [31]

OECD (2020), *Enhanced Access to Publicly Funded Data for Science, Technology and Innovation*, OECD Publishing, Paris, https://dx.doi.org/10.1787/947717bc-en. [42]

OECD (2020), *Recommendation of the Council on Consumer Product Safety*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0459. [53]

OECD (2020), *Why open science is critical to combatting COVID-19*, http://www.oecd.org/coronavirus/policy-responses/why-open-science-is-critical-to-combatting-covid-19-cd6ab2f9/. [5]

OECD (2018), *Community site: Enhanced Access to Publicly Funded Data for Science, Technology and Innovation*, https://community.oecd.org/community/cstp/enhanced-data-access. [61]

OECD (2017), "Business models for sustainable research data repositories", *OECD Science, Technology and Industry Policy Papers*, Vol. 47, https://doi.org/10.1787/23074957. [24]

OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, https://doi.org/10.1787/9789264229358-en [1]

OECD (2015), "Making Open Science a Reality", *OECD Science, Technology and Industry Policy Papers*, No. 25, OECD Publishing, Paris, https://dx.doi.org/10.1787/5jrs2f963zs1-en. [7]

OECD (2006), *Recommendation of the Council concerning Access to Research Data from Public*, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347. [30]

Office of National Statistics UK (n.d.), *Secure Research Service*, https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#the-five-safes. [25]

Paic, A. (2021), *OECD Innovation Blog - Making data for science as open as possible to address global challenges*, https://oecd-innovation-blog.com/2021/01/20/oecd-recommendation-access-research-data-public-funding-update-covid-19/. [60]

Piwowar et al. (2018), "The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles", *Peerj*, https://peerj.com/articles/4375/?utm_source=TrendMD&utm_campaign=PeerJ_TrendMD_0&utm_medium=TrendMD. [16]

RDA (2014), *RDA governance document*, http://dx.doi.org/10.15497/RDA00001. [37]

RDA (n.d.), *All recommendations and outputs*, https://www.rd-alliance.org/recommendations-and-outputs/all-recommendations-and-outputs. [38]

ROARMAP (n.d.), *Registry of Open Access Repository Mandates and Policies (ROARMAP)*, http://roarmap.eprints.org/. [18]

Saenen, B. et al. (2019), *Research Assessment in the Transition to Open Science - 2019 EUA Open Science and Access Survey Results*, https://eua.eu/downloads/publications/research%20assessment%20in%20the%20transition%20to%20open%20science.pdf. [58]

Sanderson, T., A. Reeson and P. Box (2017), *Understanding and unlocking the value of public research*, https://www.researchgate.net/profile/Todd_Sanderson/publication/313756769_Understanding_and_unlocking_the_value_of_public_research_data_OzNome_social_architecture_report/links/58a4e30f92851cf0e391e90c/Understanding-and-unlocking-the-value-of-public-resear. [13]

Schlitz, M. (2018), "Science Without Publication Paywalls: cOAlition S for the Realisation of Full and Immediate Open Access", *PLOS Medicine*, https://doi.org/10.1371/journal.pmed.1002663. [40]

Science Europe (2015), *Principles on Open Access to Research Publications*, https://www.scienceeurope.org/media/4kxhtct2/se_poa_pos_statement_web_final_20150617.pdf. [33]

Shin, E. (2018), "Korean Case Report on Enhanced Access to Public Data for Science, Technology & Innovation", *Case study contributed to CSTP project on enhanced access to data*, https://community.oecd.org/community/cstp/enhanced-data-access. [45]

Tramte, P. (2018), *Case study on research data management and openness in Slovenia*, https://community.oecd.org/community/cstp/enhanced-data-access. [47]

Treasury Board of Canada Secretariat, Open Government Team (2018), *Case study: Canada's Open Government Portal*, https://community.oecd.org/community/cstp/enhanced-data-access. [51]

Vicente-Saez, R. and C. Martinez-Fuentes (2018), "Open Science now: A systematic literature review for an integrated definition", *Journal of Business Research*, pp. 428-436, https://doi.org/10.1016/j.jbusres.2017.12.043. [8]

Vickery, G. (2011), "Review of recent studies on PSI re-use and related market developments.", *Information Economics*, https://www.researchgate.net/profile/Graham_Vickery2/publication/265573154_Review_of_Recent_Studies_on_PSI_Re-Use_and_Related_Market_Development/links/5f5a828e4585154dbbc86872/Review-of-Recent-Studies-on-PSI-Re-Use-and-Related-Market-Development.pdf. [11]

Wilkinson, M et al. (2016), "The FAIR Guiding Principles for scientific data management and stewardship", *Sci Data*, Vol. 3, http://dx.doi.org/10.1038/sdata.2016.18. [35]

World Data System (n.d.), *World Data System*, https://www.worlddatasystem.org/. [20]