



Open Science and Space Science: Some OECD Perspectives

Claire Jolly

Head, Innovation Policies for Space and Oceans (IPSO) Unit

Directorate for Science, Technology and Innovation

Organisation for Economic Co-operation and Development (OECD)

Email: claire.jolly@oecd.org

*Expert Meeting in preparation of the United Nations / Italy
Workshop on the Open Universe Initiative
ASI, Rome, Italy, 11-12 April 2017*



The age of “open science”

- Data and measurement have always been fundamental to science
- But new ICT instruments and methods of data-intensive exploration is prompting the arrival of “data-intensive scientific discovery”
- Data are collected and generated more quickly and in larger volumes than ever before



Defining open science

Open Science is a process that:

“make the primary outputs of **publicly funded research** results – **publications** and the **research data** – publicly accessible in digital format with **no or minimal** restriction”

(OECD, 2015)



Defining open data in research

Openness means:

- i) access should be granted on equal or non-discriminatory terms, and
- ii) access costs should not exceed the marginal cost of dissemination.

While marginal cost pricing is often considered the best option for the public sector, that option is seen as unattractive for the private sector, for which at least cost recovery is a necessity (e.g. subscription fees, freemium, voluntary donations, in combination with cross-subsidies).



The challenges of open data



Source: Deetjen, Meyer and Schroeder (2015).



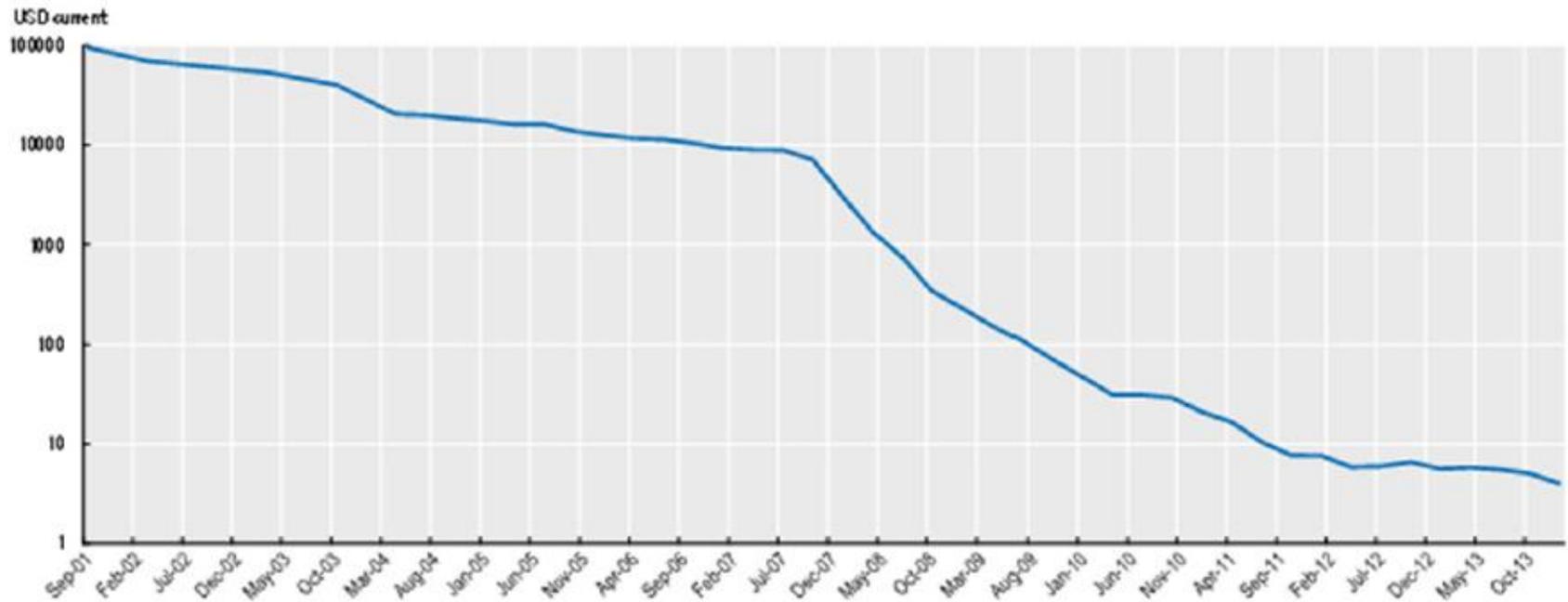
Diversity of scientific data

- **Observational data:** from telescopes, satellites, sensor networks, surveys (incl. social) and other instruments that record historical information or one-time phenomena.
- **Experimental data:** captured from high-throughput machines (such as accelerators), through clinical trials and biomedical and pharmaceutical testing, or through other controlled experiments.
- **Computational data:** generated from large-scale computational simulations (increasingly running on supercomputers).
- **Reference data sets:** highly curated data that are often in high demand by multiple scientific communities (i.e. mapping the human genome, documenting proteins, amassing longitudinal data on economic and social status).



A good case: Cost of genome sequencing

2011-14, cost per genome, logarithmic scale

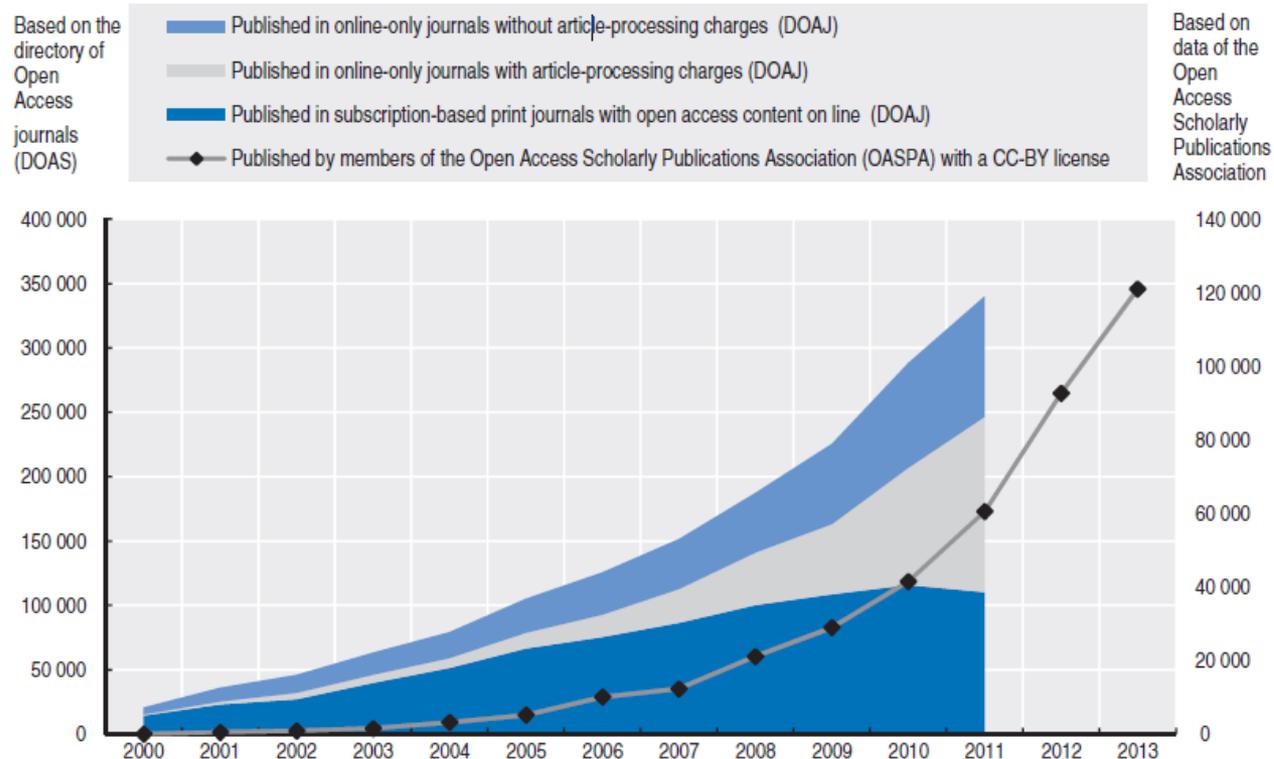


Source: OECD (2014), *Measuring the Digital Economy: A New Perspective*, OECD Publishing, Paris.



Open access publishing is on the rise

Number of papers, 2000-13

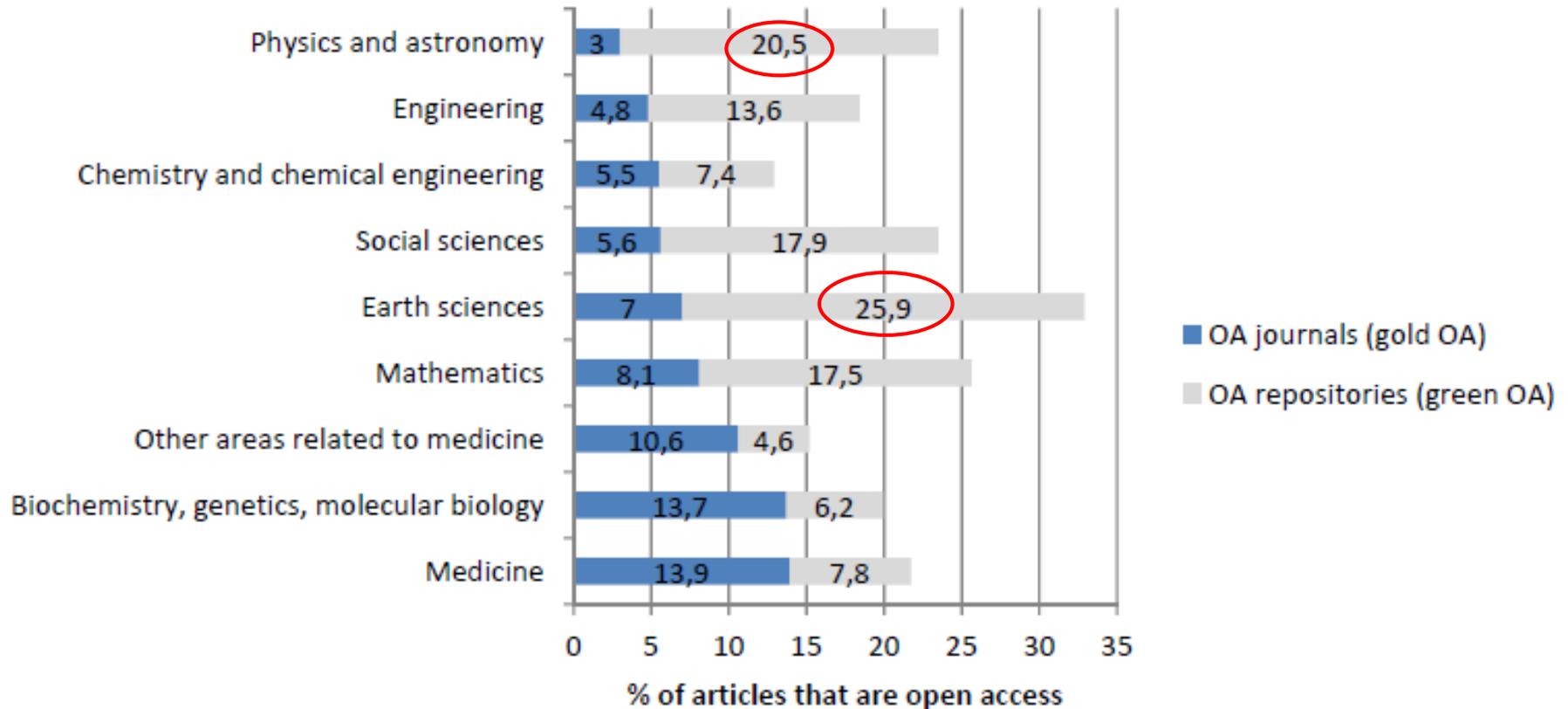


Note: Laakso and Björk describe the results of a study that focuses on measuring the longitudinal development of gold OA publication volume for the years 2000 to 2011. The study is founded on the assumption that the full population of OA journals is listed in the Directory of Open Access Journals (DOAJ). Figures of the Open Access Scholarly Publications Association (OASPA) include a total of 399 854 articles that were published with the CC-BY license by its members during the 2000-13 period. 30% of those articles (120 972) were actually published in 2013 alone. These OASPA numbers include only articles that were published in journals whose entire content is OA, so articles that were published in hybrid OA journals are not included.

Source: OECD (2016), STI Outlook 2016, OECD Publishing, Paris.



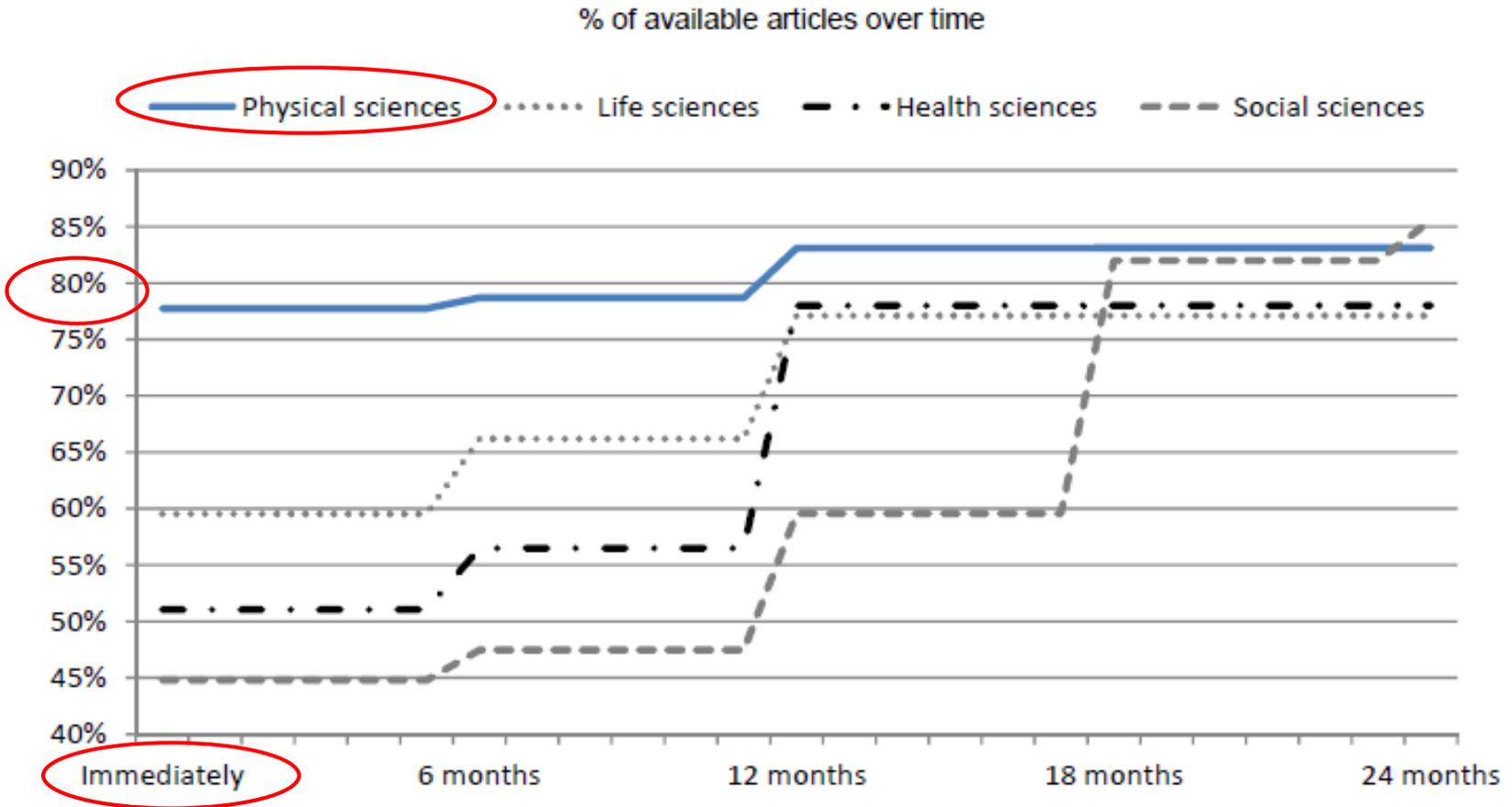
Open access varies by discipline



Source: UNESCO (2012), Policy Guidelines for the Development and Promotion of Open Access, UNESCO Publishing, and Björk et al. (2010), "Open Access to the scientific journal literature: Situation 2009", PloS ONE, Vol. 5, No. 6.



Embargo length of different disciplines



Source: Laakso, M. (2014), Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed, *Scientometrics*, Vol. 99, pp. 475-494.



Measuring the impact of open science and open data

Greater access to scientific inputs and outputs can improve the effectiveness and productivity of the scientific and research system

- by reducing duplication costs in collecting, creating, transferring and reusing data and scientific material;
- by allowing more research from the same data;
- by multiplying opportunities for domestic and global participation in the research process;
- and by ensuring more possibilities for testing and validating scientific results.

Increased access to research results (in the form of both publications and data) can not only foster spillovers to scientific systems, but also boost innovation systems more broadly (new products and services)



OECD Principles and Guidelines for Access to Research Data from Public Funding

- ✓ Science and Technology Ministers called on the OECD in 2004 to develop a set of guidelines based on commonly agreed principles to facilitate cost-effective access to digital research data from public funding.
- ✓ The Principles and Guidelines were attached to an OECD Recommendation and endorsed by the OECD Council on 14 December 2006.



OECD Principles and Guidelines for Access to Research Data from Public Funding

Research data from public funding is defined as

“the research data obtained from research conducted by government agencies or departments, or conducted using public funds provided by any level of government.”



Effective access to research data, in a responsible and efficient manner...

... is increasingly required to take full advantage of the new opportunities and benefits offered by ICTs.

Accessibility to research data has become an important condition in:

1. The good **stewardship** of the public investment in factual information;
2. The creation of strong value chains of **innovation**;
3. The **enhancement of value** from international co-operation.



More specifically, improved access to, and sharing of data:

- ✓ Reinforces open scientific inquiry;
- ✓ Encourages diversity of analysis and opinion;
- ✓ Promotes new research;
- ✓ Makes possible the testing of new or alternative hypotheses and methods of analysis;
- ✓ Supports studies on data collection methods and measurement;
- ✓ Facilitates the education of new researchers;
- ✓ Enables the exploration of topics not envisioned by the initial investigators;
- ✓ Permits the creation of new data sets when data from multiple sources are combined.



OECD Principles and Guidelines for Access to Research Data from Public Funding

- **Openness** – Open access to research data from public funding should be easy, timely, user-friendly and preferably Internet-based.
- **Flexibility** – Flexibility requires taking into account the rapid and often unpredictable changes in ICTs, the characteristics of different research fields, and the diversity of research systems, legal frameworks and cultures among member countries.
- **Transparency** – Information on research data and data-producing organisations, documentation on the data, and conditions attached to the use of data should be internationally available in a transparent way, ideally through the Internet.
- **Legal conformity** – Data access arrangements should respect the legal rights and legitimate interests of all stakeholders in the public enterprise. Restriction to access may be for reasons of national security; privacy and confidentiality; trade secrets and intellectual property rights; protection of rare, threatened or endangered species; or legal processes.
- **Protection of intellectual property** – Data access arrangements should consider the applicability of copyright and other intellectual property laws that may be relevant to publicly funded research databases (as in the case of public-private partnerships).
- **Formal responsibility** – Access arrangements should promote the development of rules and regulations dealing with the responsibilities of the various parties involved; should be developed in consultation with representatives of all parties affected; and should be responsive to factors such as the characteristics of the data and their potential value for research purpose. Data management plans and long-term sustainability should also be considered.
- **Professionalism** – Institutional arrangements for the management of research data should be based on the relevant professional standards and values embodied in the codes of conduct of the scientific communities involved.
- **Interoperability** – Access arrangements should consider the relevant international data documentation standards.
- **Quality** – The value and utility of data depend to a large extent on the quality of the data. Particular attention should be paid to ensuring compliance with explicit quality standards.
- **Security** – Attention should be devoted to supporting the use of techniques and instruments to guarantee the integrity and security of research data.
- **Efficiency** – One of the central goals of promoting data access and sharing is to improve the efficiency of publicly funded scientific research so as to avoid expensive and unnecessary duplication of effort. This also involves cost and benefit analysis to define data retention protocols, the engagement of data management specialist organisations, and the development of new reward structures for researchers and database producers.
- **Accountability** – The performance of data access arrangements should be subject to periodic evaluation by user groups, responsible institutions and research funding agencies.
- **Sustainability** – Due consideration should be given to the sustainability of access to publicly funded research data as a key element of the research infrastructure.



Policy responses to support open data

1. **Mandatory rules (sticks)** = often implemented in the form of requirements in research grant agreements, or in some cases defined in national strategies or institutional policy frameworks.
2. **Incentive mechanisms (carrots)** = financial incentives to cover the release of data sets; proper acknowledgment of open data efforts of researchers and academics (e.g. data set citations or career advancement mechanisms that take into account open science or data-sharing initiatives)
3. **Enablers** = infrastructure developed to share data; initiatives undertaken to develop an open science and open data culture; amendments to the legal framework to make it increasingly open science-friendly; or development of the skills necessary for researchers to share and reuse the research outputs produced by others.



Looking forward: OECD ongoing activities

- Large-scale OECD activity on the digitalisation of science, technology and innovation (with direct inputs from the OECD Space Forum)
- Ongoing review of two major OECD instruments with possible impacts on space science data:
 - *Recommendation of the Council concerning Access to Research Data from Public Funding of 14 December 2006,*
 - *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information of 30 April 2008*

<https://www.innovationpolicyplatform.org/oecd-space-forum>



**Thank you for
your attention**



<https://www.innovationpolicyplatform.org/oecd-space-forum>