DIRECTORATE FOR SCIENCE, TECHNOLOGY AND INNOVATION
COMMITTEE FOR SCIENTIFIC AND TECHNOLOGICAL POLICY

Cancels & replaces the same document of 27 February 2018

OECD Global Science Forum

Joint CSTP-GSF Workshop "Towards New Principles for Enhanced Access to Public Data for Science, Technology and Innovation" -- Draft Programme

13 March 2018
OECD Conference Centre, Paris

This workshop will be held back to back with the 112th meeting of the CSTP.

The aim is to explore several key areas and related principles and policy actions to promote enhanced access to data from public research. These are to be considered within the overall context of the existing OECD principles and guidelines for access to research data from public funding and the outstanding or new challenges to implementing this recommendation.

Contacts: Alan Paic, Senior Policy Analyst, Science and Technology Policy Division (STI/STP); Tel: +33 1 45 24 99 77; Email: alan.paic@oecd.org; Carthage Smith, Senior Policy Analyst – Lead Co-ordinator, OECD Global Science Forum, Tel: +33 1 45 24 92 80; Email: carthage.smith@oecd.org

JT03427833

Draft Programme: Towards New Principles For Enhanced Access To Public Data For Science, Technology And Innovation

13 March 2018

OECD Conference Centre, Paris

Introduction

At its 109th session in October 2016, the OECD Committee for Science and Technology Policy (CSTP) discussed and approved a proposed joint development of a possible new overarching recommendation on enhanced access to data, together with the Committee for Digital Economy Policy (CDEP) and the Public Governance Committee (PGC) [COM/DSTI/CDEP/STP/GOV/PGC(2016)1]. CSTP's central instrument in this domain is the Recommendation of the Council concerning Access to Research Data from Public Funding [C(2006)184] (referred to as "the Recommendation" in the text below).

As described in the Work Plan for the Project on Enhanced Access to Data and Terms of Reference (ToR) [COM/DSTI/CDEP/STP/GOV/PGC(2017)1], the work has the objective to identify gaps in current data governance frameworks, which will enable identifying the common elements that could be further developed as general principles on enhanced access to data, possibly resulting in an OECD umbrella legal instrument adopted by Council to serve as reference for any revision of existing OECD legal instruments concerning open data or for the development of new ones.

A survey was conducted in mid-2017 to assess the current use of the Recommendation, the results of which are summarised in 'Open Access To Data In Science, Technology And Innovation – Initial Survey Findings' [DSTI/STP(2017)25], discussed by the CSTP at its 111th Meeting in October 2017.
The key issues identified in the survey as requiring policy attention were:

- Data governance for trust - addressing privacy, confidentiality, quality and ethical issues
- Discoverability/findability, machine readability and data standards
- Recognition and reward system for data authors
- Definition of responsibility and ownership
- Business models for open data provision
- Building human capital and institutional capabilities at public agencies, to manage, create, curate and reuse data.

In parallel, the OECD Global Science Forum has recently completed two projects to inform policies to promote open data for science. The first of these addresses Business models for sustainable research repositories. The second project focuses on Coordination and support of international research data networks, which are necessary to support a global open science enterprise. These two projects are a follow-up to earlier GSF work on the use of data, including personal data, for social science research

(OECD, 2013) and (OECD, 2016). In addition some initial work has been carried out on incentives for sharing data in the specific area of research on dementia [DSTI/STP/MS(2015)16].

The CSTP is thus joining forces with the Global Science Forum and organising a workshop back to back with the 112th meeting of the CSTP with the objective to deepen the gap analysis already initiated through the recent survey report [DSTI/STP(2017)25].

DRAFT PROGRAMME

13 March 2018

Chairman: Dominique Guellec, Head of Science and Technology Policy, OECD

9:00-9:10

Introductory remarks

Andrew Wyckoff, Director Science, Technology and Innovation, OECD

9:10-9:20

Access to public data for science, technology and innovation – summary of learnings from OECD project

Alan Paic, Senior Policy Analyst, Science and Technology Policies, OECD

A summary will be given of the main issues identified in the framework of the joint development of a possible new overarching recommendation on enhanced access to data, together with the Committee for Digital Economy Policy (CDEP) and the Public Governance Committee (PGC). These issues are to be discussed in the following panel discussions.

9:20-10:30

Panel 1. Data governance and trust for science, technology and innovation

Data governance includes the broad range of guidelines, regulations, principles, standards of good practice and processes that determine how data is produced, collected, managed, distributed and re-used. Sound data governance is needed to ensure trust from both data providers and users and secure accessibility. In addition to provenance and quality, for sensitive data, this requires attention to privacy, confidentiality and ethical issues, including informed consent. Balancing the potential public benefits and risks of sharing data in science, technology and innovation is a critical issue for data governance.

Introductory remarks by Duncan McIntyre, Assistant Secretary, Department of Prime Minister and Cabinet of Australia

Moderator: Alan Paic, Senior Policy Analyst, Science and Technology Policies, OECD

Panellists:

Marin Dacos, Open Science Counsellor, French Ministry of Higher Education, Research and Innovation

Andreas Ebert, Office of the CTO, Microsoft Corporation

Eunjung Shin, STEPI, Korea

Peter Elias, University of Warwick

10:30-11:00 – Coffee Break

11:00-12:00

Panel 2. Data standards, interoperability and re-use

There is insufficient information on what data is available for and from research, and when data can be found it is not always useable. There is a need for user-friendly and widely accessible catalogues for datasets, services and standards, based on machine readable metadata and common and persistent identification mechanisms. International standards for data documentation have been developed but are not always easy to adopt and thus are variably implemented. At the same time, interoperability and common standards are essential for ongoing efforts to establish open science clouds in Europe, Australia, US (NIH), Africa.

The FAIR[1] principles are being widely implemented to facilitate open access to data, and the Research Data Alliance issues recommendations addressing a broad range of issues related to interoperability, data citation, data catalogues, work-flows for research data publishing.

Moderator: Ross Wilkinson, Australian National Data Service, Research Data Alliance

Panellists:

Simon Hodson, Executive Director, CODATA

Tyler Walters, Dean, University Libraries and Professor, Virginia Tech

Thordis Sveinsdottir, Trilateral Research

Jean-Francois Abramatic, Consultant (formerly INRIA and IBM)

12:00-13:00

Panel 3. Definition of responsibility and ownership

Issues of copyright and intellectual property over data play a big role in open data access, especially with regards to text and data mining (TDM) in the context of scientific research and innovation. Issues of ownership can complicate data sharing and re-use even amongst different public sector actors. Co-operation with the private sector is an additional challenge to be addressed, with sensitive issues of data ownership for data created through public private partnerships or data from public research being offered on a private platform.

In this respect it is worth noting the specific case of the EU, which has created an exclusive "sui generis" right for database producers to protect the investment of time, money and effort, irrespective of whether a database is innovative or not.

Moderator: Ingrid Dillo, Research Data Alliance and Leiden University

Panellists:

Stan Matwin, Institute for Big Data Analytics, Dalhousie University

---

[1] FAIR stands for Findability, Accessibility, Interoperability and Re-use

Paul Uhlir, Consultant, USA

Ben White, Head of Intellectual Property at British Library

Marie Timmermann, Science Europe

13:00-14:00 – Lunch Break

14:00-15:00

Panel 4. Recognition and reward systems for data providers and stewards

Data sharing requires cultural change among researchers in many fields of science. Perceived barriers and risks of providing open access to data need to be counterbalanced by appropriate acknowledgement and reward systems. Researchers have incentives to publish scientific results, preferably positive ones. Incentives to publish data are less developed, and usually seen as a constraint imposed by funding agencies (threat of discontinuing funding) and/or publishers (data statements required). Data citation has not been widely implemented, and the prerequisites for it - standard formats, citation metrics - are not being broadly adopted. Open Science needs to be embedded in evaluation systems to ensure that researchers who provide high quality research data are rewarded.

Moderator: Simon Hodson, Executive Director, CODATA

Panellists:

Vanessa Proudman, Director, SPARC Europe

Brian Nosek, University of Virginia School of Medicine

Mustapha Mokrane, World Data System

Samuel Goeta, Datactivi.st

Beth A. Plale, National Science Foundation

15:00-16:00

Panel 5. Business models for open data provision

Costs of provision of open data are often borne by the providing institution, but benefits accrue to stakeholders around the world. Business models for research data repositories are restrained by mandates and incentives. The OECD Global Science Forum has recently published a report[2] on this subject and the field is changing rapidly with new private sector actors competing with and/or complementing public repositories.

Moderator: Carthage Smith, Senior Policy Analyst - Lead co-ordinator, OECD Global Science Forum

Panellists:

Barend Mons, Leiden University

Jean-Marc Lazard, CEO of Opendatasoft

---

[2] OECD (2017), "Business models for sustainable research data repositories", *OECD Science, Technology and Industry Policy Papers*, No. 47, OECD Publishing, Paris. http://dx.doi.org/10.1787/302b12bb-en

Ingrid Dillo, Research Data Alliance and Leiden University

Duncan Campbell, Director, Global Sales Partnerships at Wiley

16:00-16:30 – Coffee Break

16:30-17:30

Panel 6. Building human capital and institutional capabilities

Researchers often lack data management skills (funding agencies often make this a requirement for recurrent funding). Users (who may be from different sectors of academia or the private sector) do not always have appropriate skills for correct interpretation and analysis. Technical staff in data repositories needs training on data standards. Specific curricula including statistical skills, computer science and information science are needed. Many countries report limitation of current curricula in addressing those skills needs. Technical staff can be further broken down into various categories, including data engineers, data analysts and data stewards that require distinct skill sets.

Moderator: Bart Dumolyn, Department of Economy, Science and Innovation Flemish Government

Panellists:

Kazuhiro Hayashi, National Institute of Science and Technology Policy, Japan

Michelle Barker, Nectar/University of Melbourne

Michelle Willmers, University of Capetown

Pedro Fernandes, Instituto Gulbenkian

Victoria Stodden, University of Illinois

Steve Brewer, University of Southampton

17:30-18:00

Concluding remarks

Christian Reimsbach-Kounatze, OECD -CDEP

Barbara Ubaldi, OECD-GOV

Dominique Guellec, Head of Science and Technology Policy, OECD