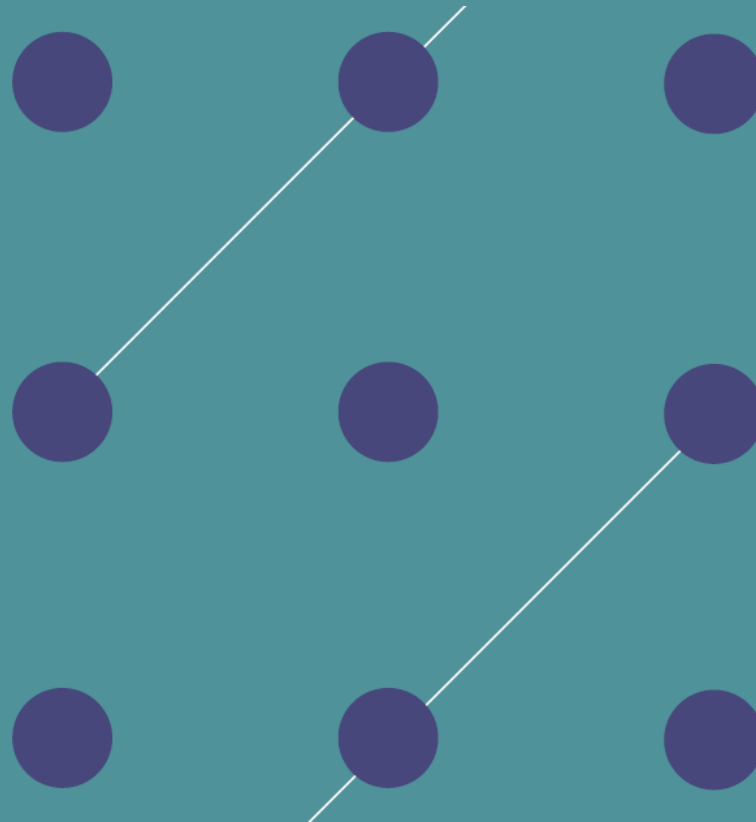


# BoF on an Open Research Data Survey

*Preparation of a WG Case Statement on an Open Research Data Survey*



10<sup>th</sup> RDA Conference

MONTRÉAL, CANADA

19 SEPTEMBER 2017

Stephane Berghmans, DVM PhD  
Ingeborg Meijer, PhD

# 01: Introduction



# Governments, funders, institutions, publishers are formulating Open Data policies



## OPEN RESEARCH DATA IN HORIZON 2020

### CHALLENGE

Wider access to scientific facts and knowledge helps researchers, innovators and the public find and re-use data, and check research results:

offers better value for EU research funds

encourages research across scientific fields



a public benefit



essential for solving today's complex societal challenges

### SOLUTION

Horizon 2020 already mandates open access to all scientific publications



From 2017, research data is open by default, with possibilities to opt out



National Coordination Point  
Research Data Management



Accessibility | Media Enquiries

search  Search

Home Funding Research Innovation Skills Public Engagement News, Events and Publications About Us

Home / Research / RCUK Common Principles on Data Policy

## RCUK Common Principles on Data Policy

Making research data available to users is a core part of the Research Councils' remit and is undertaken in a variety of ways. We are committed to transparency and to a coherent approach across the research base. These RCUK common principles on data policy provide an overarching framework for individual Research Council policies on data policy.

### Principles

- Publicly funded research data are a public good, produced in the public interest, which should be made openly available with as few restrictions as possible in a timely and responsible manner.

## VSNU calls for a national RDM strategy

06 MAY 2015

The Association of Universities in the Netherlands (VSNU) has asked SURF to set up a National Research Data Management Coordination Point. The VSNU's Research and Valorisation Steering Group notes that much work in the RDM field is already taking place at universities. They also see a need for more coordination.



contact calendar press vacancies Neder

Home News & events Funding Research & results

## Data management protocol

- Open Science
- Q&A Open Access at NWO
- Researchers about Open Access
- Data management Data management section Q&A

Responsible data management is part of good research. NWO wants research data that emerges from publicly funded research to become freely and sustainably available, as much as possible, for the use by other researchers (data management as part of Open Science). Due consideration is given to aspects such as privacy, public security, ethical limitations, property rights and commercial interests.

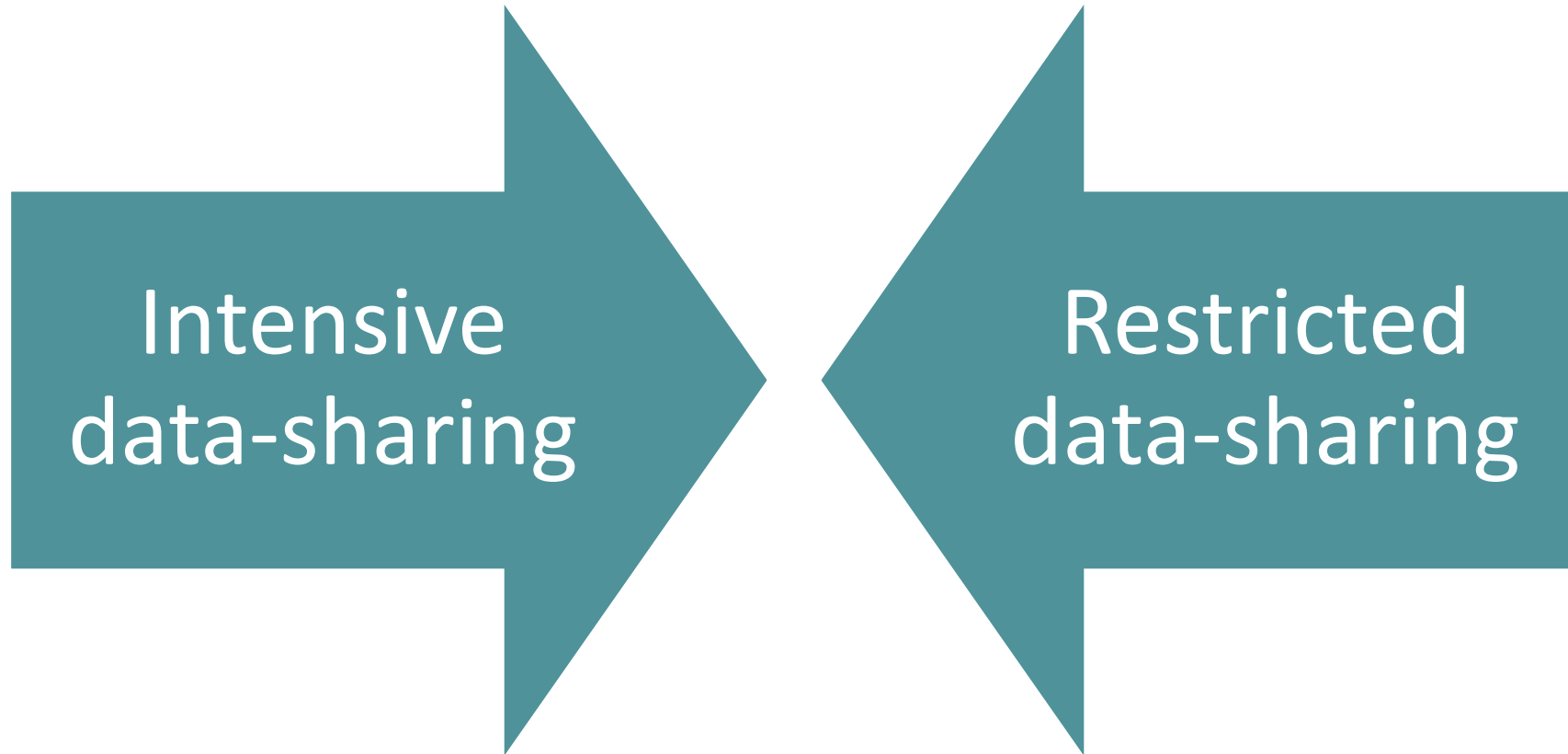


Home Expertise & Services Data Management Support Hub

## Data Management Planning

Data management planning is a matter of good research practice. At Wageningen University & Research PhD candidates and Chair Groups are required to have a Data Management Plan.

Culture is changing ... not uniformly





# Track changes and benchmark

## Towards Open Research

Practices, experiences, barriers and opportunities

October 2016

Veerle Van den Eynden, Gareth Knight, Anca Vlad, Barry Radler, Carol Teng, David Leon, Frank Manista, Jimmy Whitworth and Louise Corti

CSUC  
Consorci de Serveis Universitaris de Catalunya

CSUC  
Consorci de Biblioteques Universitàries de Catalunya

## Gestió de les dades de recerca: resultats de l'enquesta prospectiva a gener de 2016

(Doc 1656) (B&B)ROR(enquesta)01610 Resultats(enquesta)Prospectiva02020.doc | OTIR 22.11.16

El Grup de Treball de Suport a la Recerca del CSUC fou creat per representants de la Universitat de Barcelona, l'Universitat Autònoma de Barcelona, l'Universitat Politècnica de Catalunya, l'Universitat Pompeu Fabra, l'Universitat de Girona, l'Universitat de Lleida, l'Universitat Rovira i Virgili, l'Universitat Oberta de Catalunya, l'Universitat de València, l'Universitat Central de Catalunya, l'Universitat Ramon Llull i l'Universitat Jaume I va elaborar entre l'octubre i el desembre de 2015 una enquesta destinada als responsables dels diferents processos de les Universitats de Catalunya que rebria el finançament del programa Horizon 2020 de la Comissió Europea, independentment de si fossin o no de l'Open Research Data Pilot, amb la finalitat de conèixer les necessitats dels investigadors en matèria de gestió de dades de recerca.

### 1. Mostra

L'enquesta va ser enviada des de les diferents biblioteques i oficines de recerca de les universitats de Catalunya entre desembre de 2015 i gener de 2016 a una mostra de 164 projectes europeus de diferents àmbits de coneixement.

Aquesta informació està basada en la base de les 73 respostes rebudes fins al 31 d'agost de 2016 i, tot i que el grau de resposta és desigual entre les diferents universitats (vegeu taula 1), respon a un percentatge de resposta del 45%.

Open Capital, 2 | D'Info Mèdia - 8984 Barcelona - T: 93 203 6461 - www.csuc.cat | W: 93 203 6111

## OPEN DATA THE RESEARCHER PERSPECTIVE

CWTS  
University of London

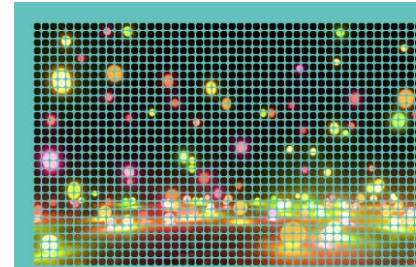
National Science Foundation  
WHERE DISCOVERIES BEGIN

## Final Report:

## Workshops to Gauge the Impact of Requirements for Public Access to Data Produced by NSF-funded Research in Mathematics and the Physical Sciences<sup>1,2</sup>

<sup>1</sup> Editors: Robert Haselich (NIET), Michael Hildreth (Notre Dame), Leah McIlwain (Cornell), Victoria Stoddin (UIUC), Gordon Watts (IOW), Daniel S. Katz (UIUC), Natalie Meyers (Notre Dame), Ashley E. Sands (OSU)

<sup>2</sup> Funded by NSF-PHY-1457413



Digital Science Report

## The State of Open Data

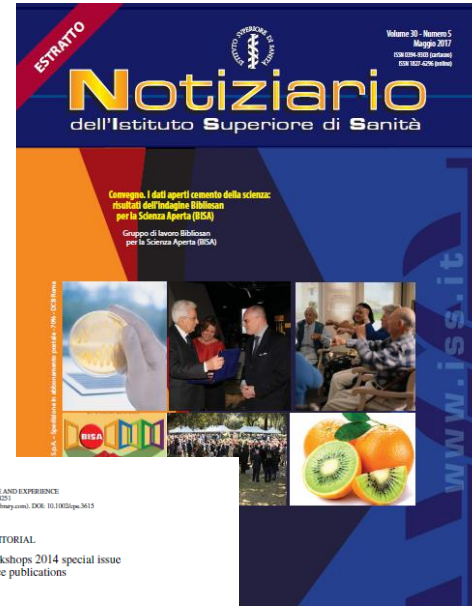
A selection of analyses and articles about open data, curated by Figshare

Foreword by Professor Sir Nigel Meade

OCTOBER 2015

DIGITAL  
SCIENCE

figshare



## Notiziario dell'Istituto Superiore di Sanità

Convegno: I dati aperti cemento della scienza: risultati dell'indagine Biblioman per la Scienza Aperta (BISA)

Gruppo di lavoro: Biblioman per la Scienza Aperta (BISA)



CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE  
Concurrency Comput. Pract. Exper. (2015) 27, 467-471.  
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.2615

### EDITORIAL

## Science gateway workshops 2014 special issue conference publications

### 1. INTRODUCTION

Science gateways are a solution for user communities to access applications and data via a graphical user interface. These graphical user interfaces hide the underlying infrastructure, as far as feasible and as far as desired by the users. In general, science gateways offer a single point of entry to create and/or analyse domain-specific data. Their core goal is to increase the usability and accessibility of computational tools and digital data as well as to leverage reusability of scientific processes. While the user interfaces are especially tailored to the specific demands of a user community, the underlying infrastructures, for example, national or international distributed computing infrastructures (e.g., XSEDE), are mainly applicable for a wide range of use cases. Thus, science gateway frameworks and science gateway APIs, which offer building blocks for the management of jobs and data within such infrastructures, ease the implementation of science gateways for developers. The latter can focus on the domain-specific demands while reusing or extending available building blocks.

The contributions in this special issue present the current state-of-the-art research and academic trends in the area of science gateways as well as demonstrate available solutions for the users. Submissions are grouped in five general areas: science gateway use and sustainability, generic development frameworks, novel workflow-oriented approaches, data management, and use cases from diverse domains. The statistics illuminate among other topics the increased usage of science gateways, which is also reflected in high number of submissions demonstrating specific use cases. Consequently, sustainability approaches have found their way into the special issue reflected not only in a submission about a model for sustainability but also in numerous submissions on development and enhancements for generic building blocks of diverse existing science gateway frameworks and APIs. While novel approaches for workflow management and data management can be also considered under the enhancements for generic building blocks addressing new technologies such as mobile applications, they have already been core subjects for a couple of years and are presented in own sections emphasizing their importance for the science gateway community.

### 2. SCIENCE GATEWAY USE AND SUSTAINABILITY

The close collaboration between user communities and developers as well as provides is crucial for developing and offering effective science gateways widely used by science communities. Insights about the demands of user communities and about possibilities to support science gateway developers can immensely improve both the efficiency of creating science gateways and their long-term sustainability. Lawrence *et al.* (1) present results from a large-scale user survey of nearly 5000 researchers from diverse research domains and computer science departments involved in science gateway provisioning, who answered a questionnaire about demands on and existing concerns for developing science gateways. Major topics include the support of user communities, aiding developers in choosing a suitable science gateway technology and involving specific expertise. The paper also notes that for the first time in the National Science Foundation (NSF) sponsored program, more users have accessed resources via gateways than by using the command line. The manuscript 'Reflections on Science Gateways Sustainability Through the Business Model Canvas: Case Study of a Neuroscience Gateway' (2) goes into detail on sustainability approaches for science

Copyright © 2015 John Wiley & Sons, Ltd.

# Increasing number of reports and surveys

- *Workshops to Gauge the Impact of Requirements for Public Access to Data Produced by NSF-funded Research in Mathematics and the Physical Sciences* (2017)
- *Open Data Research – a researcher perspective* (2017)
- *The State of Open Data – a selection of analyses and articles about open data* (2016)
- *Towards Open Research – practices, experiences, barriers and opportunities* (2016)
- *Gestió de les dades de recerca: resultats de l'enquesta prospectiva a gener de 2016* (2016)
- *e-infrastructures Austria: Forschende und ihre Daten. Ergebnisse einer österreichweiten Befragung* (2015)
- *Science Gateways Today and Tomorrow* (2015)
- *Wiley Researcher Data Insights Survey* (2014)
- *I dati aperti cemento della scienza* (to be published)
- *American Physical Society report* (to be published)
- ...

# 02: Reports

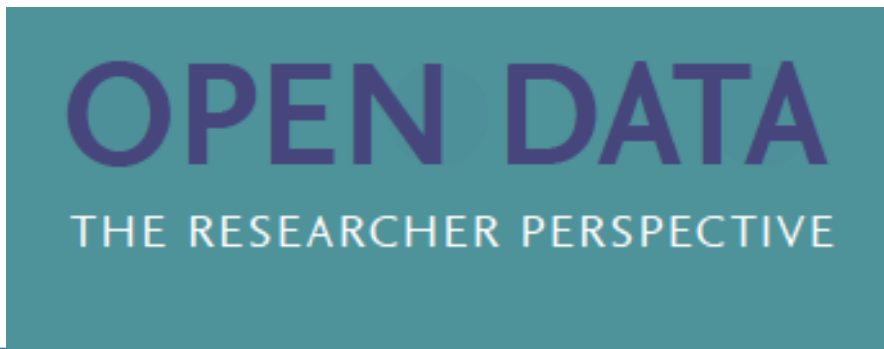
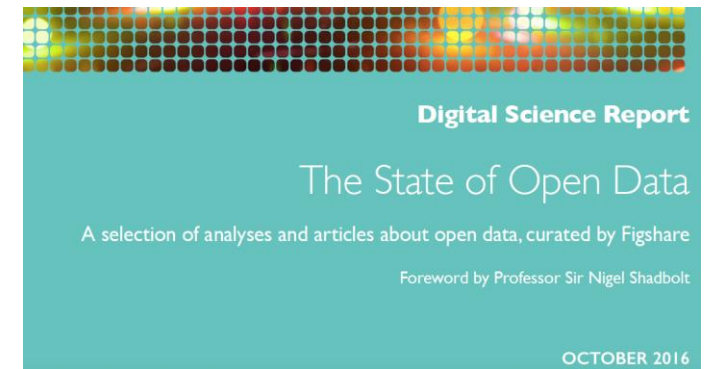




# What is in this increasing number of reports and surveys?

Summary of the context, not the content:

- Coverage
- Disciplines
- Selection
- Number of respondents
- Topics



# UK survey of academics (2015)

- Repetition of 2012; open data is new
- Initiated from library
- Coverage: UK
- Disciplines: all
- Selection: Academic staff at UK higher education institutions. In total, 64,259 academics were invited
- Number of respondents: 6,679 completed responses, response rate of 10.4 percent
- Topics: library, access, research dissemination, dissemination services and research data management



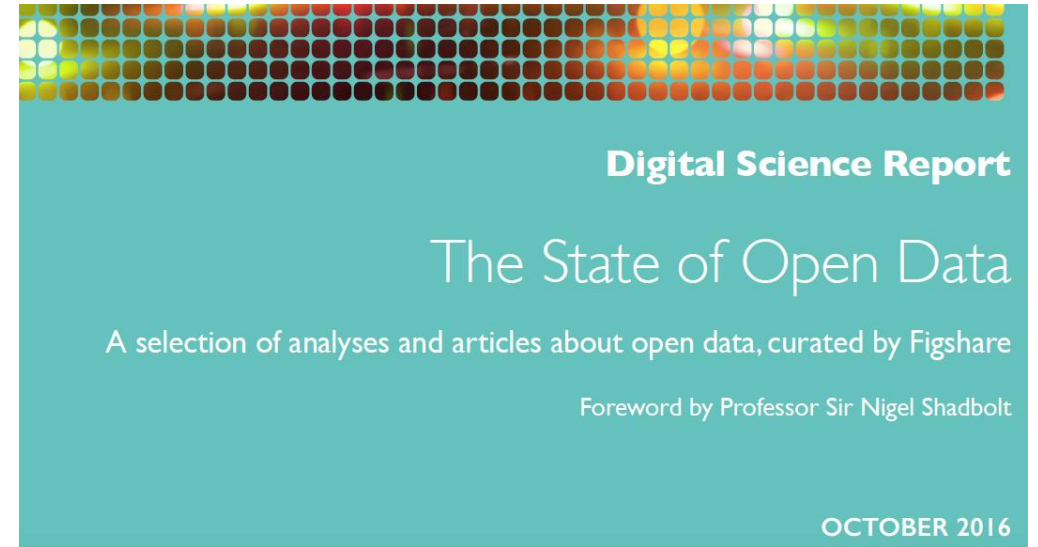
# Researcher data sharing insights (2014)

- Coverage: global
- Disciplines: 4 health sciences/life sciences/physics/social science
- Selection: ??
- Number of respondents: > 2,250
- Topics: ways of sharing, motivations, hesitations, actual sharing practices



# The State of Open Data (2016)

- Coverage: global
- Disciplines: all
- Selection: ? Figshare users
- Number of respondents: ~2,000
- Topics: Awareness on freeness, value on credit for making data open data, reuse of open data, uncertainty and gaps in knowledge on how to make it open (e.g. Licensing); institutional, funder or publisher requirements, citation practices



# Open Data; the researcher perspective (2017)

- Coverage: global
- Disciplines: all
- Selection: ~26,800 authors from Scopus
- Number of respondents: 1,164 (2.3%)
- Topics: how & why researchers are sharing, attitudes towards sharing, research data ownership, aspects of research data management, perceptions on reusability

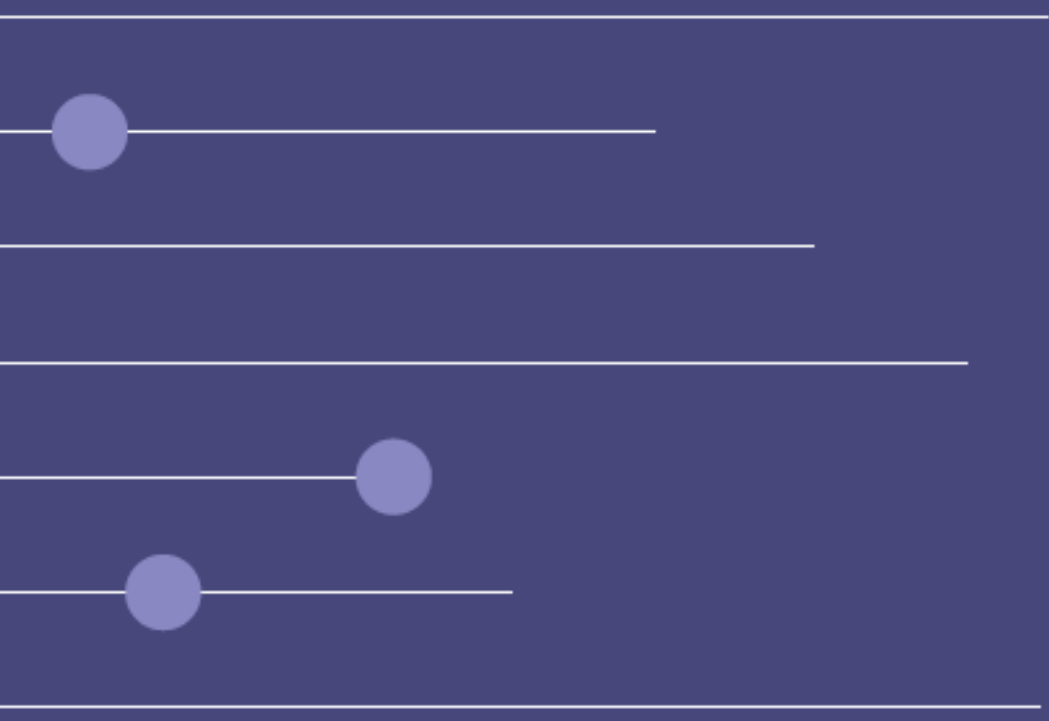


# Summary: what do we want in a joint survey?

- Coverage: Researchers, global, all disciplines
- Selection & Number of respondents; as high as possible
- Topics:
  - Actual sharing
  - Actual reuse
  - Locations of archiving
  - Data intensive vs. data restricted practices
  - Data publishing?
  - Change in issues for not-sharing
  - AOB



# 03: Next steps



# an Open Research Data Survey

**Coordinated approach** using a common survey  
→ co-developed and community-designed **open survey**

## Advantages

- To make use of **community sourcing** to design a comprehensive and broadly accepted survey
- To rationalize the use of resources by **avoiding duplication**
- To design a **flexible survey adapted** to all stakeholders, regions, environments, etc.
- To offer all stakeholders a **tool** to track the response to open data policies and measure their potential effect within scientific communities
- To standardize a survey whose multi-source results could be combined and/or aggregated to perform **meta-analysis** at various levels (institutional, disciplinary, regional, etc.)
- To create the opportunity to systematically **track open data developments** over time
- ...



# Today's objectives

- ✓ To present the concept of an open research data survey on open data and its potential benefits
- To discuss this concept and supportive arguments
- To finalize the arguments and objectives for a case statement  
[45 min]
- To identify interested stakeholders and specific partners  
[10 min] (in and outside this room)

# Today's objectives

- **What?** an Open Research Data Survey
- **How?** an RDA Working Group
- **When?** asap
  
- **Why?** advantages (and pitfalls)
- **Who?** Interested stakeholders  
(organisations & individuals)

# Project Supporters

**Stephane Berghmans**, Elsevier, EuroScience

**Jean-Claude Burgelman**, DG Research & Innovation, European Commission

**Gabi Lombardo**, European Alliance for Social Sciences & Humanities

**Ingeborg Meijer**, CWTS, University Leiden

**Martin Müller**, SwissCore

*to be continued*

Feel free to contact us:

[stephane.berghmans@elsevier.com](mailto:stephane.berghmans@elsevier.com)

[i.meijer@cwts.leidenuniv.nl](mailto:i.meijer@cwts.leidenuniv.nl)

# Thank You