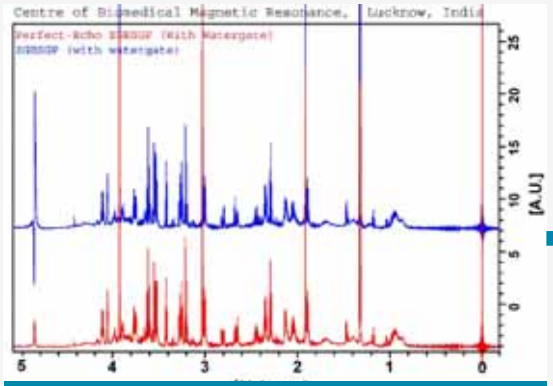
Perspectives on Open Data in Science

Open Data in Science: Challenges & Opportunities for Europe

Stephane Berghmans, DVM PhD

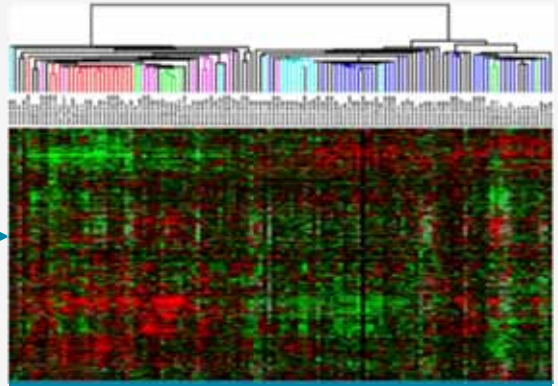
31 January 2018

When talking about data, we talk about…

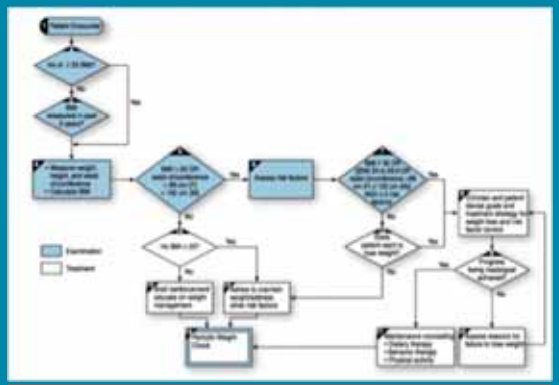


Raw data

All forms of research data, which includes everything needed to reproduce and reuse



Processed data



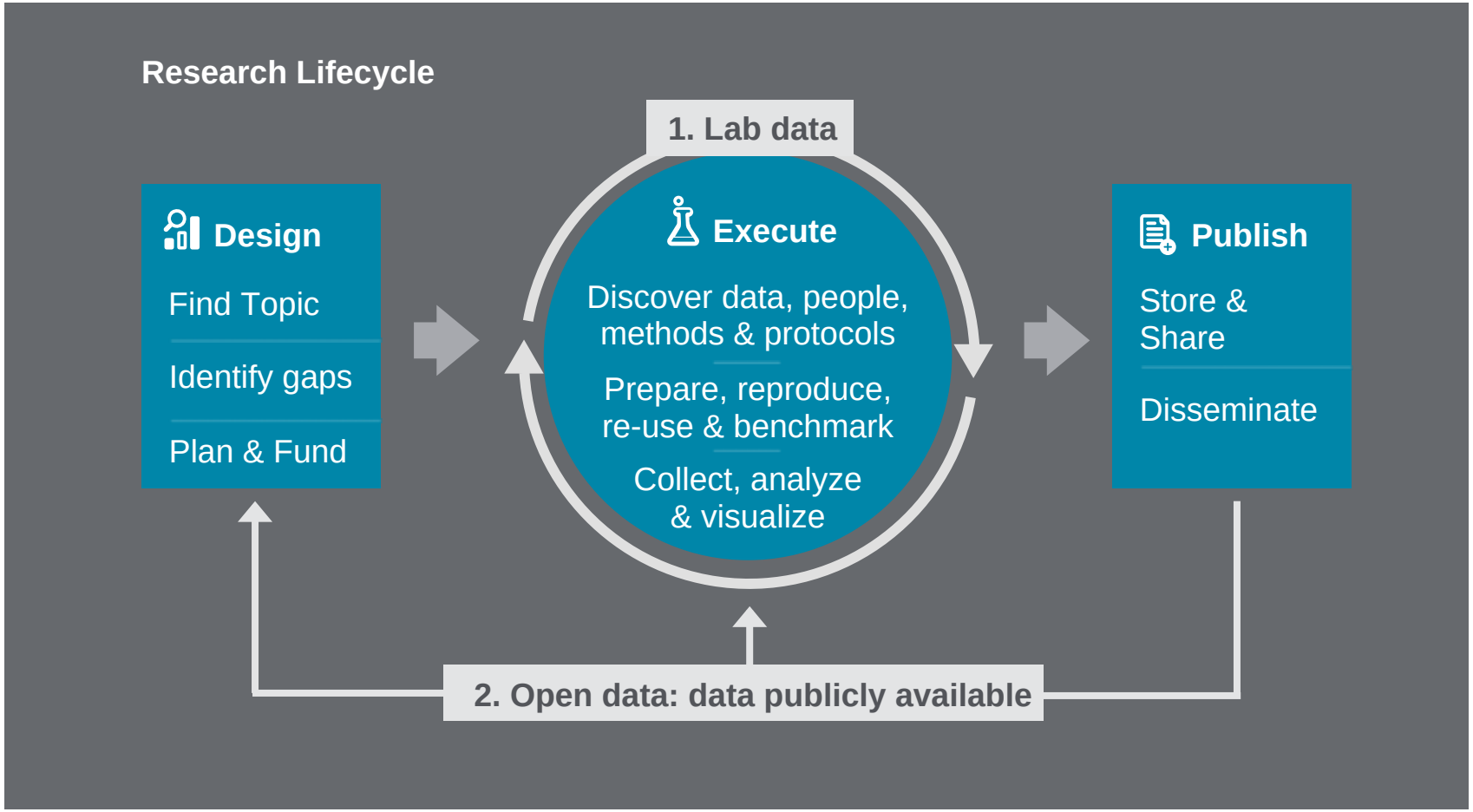
Protocols, methods, workflows

Machine & environment settings

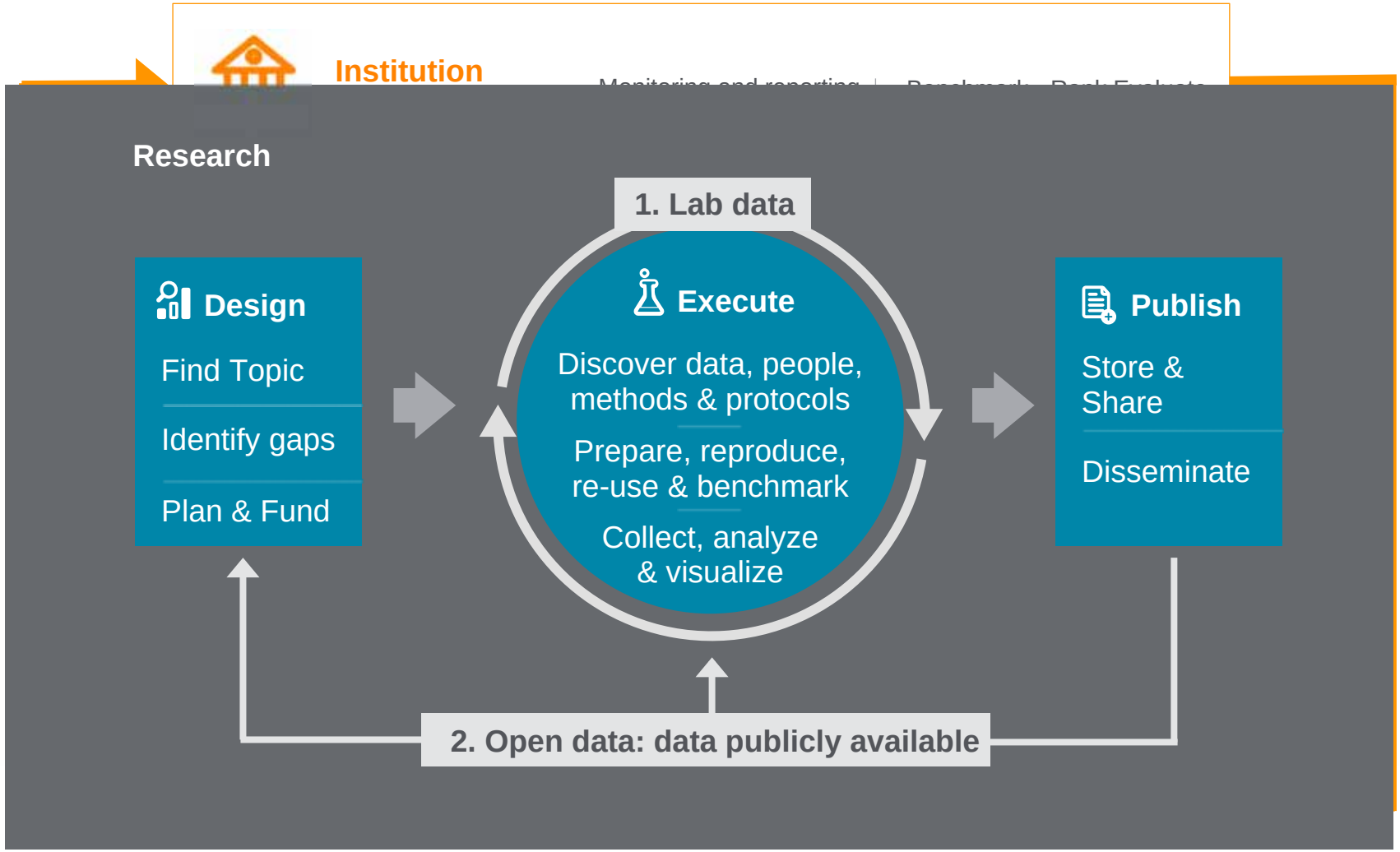
Scripts, analyses, algorithms

The research life cycle depends heavily two data life cycles:

1. Within the “lab”, covering all active data in all domains
2. Within the world at large



Taking the institutional lens, we can speak of 3 interlocking data life cycles

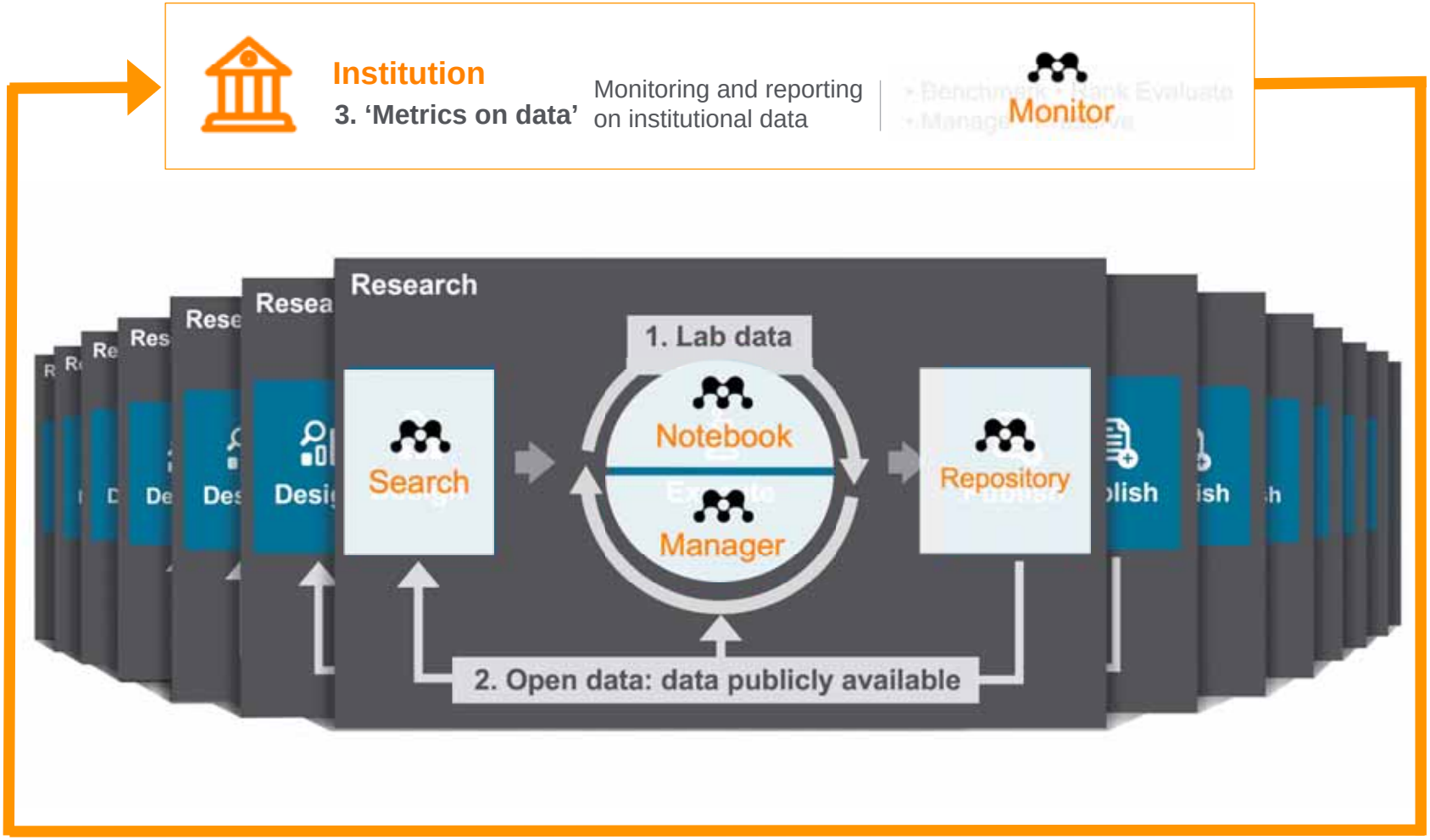


Re-using research data improves outcomes for the research life cycle

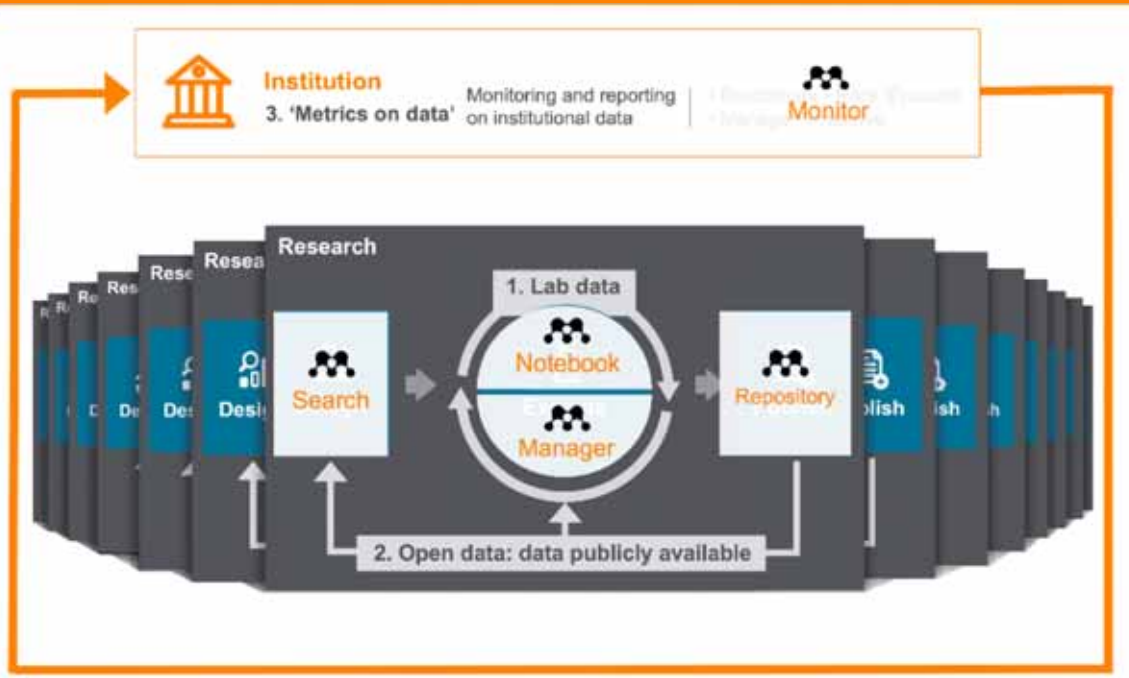
- This means improving the research data life-cycles: (1) within the lab and (2) to the world at large
- This also means keeping track of the institutional data lifecycles, and (3) reporting on them

Mendeley Data

A modular, cloud-based platform designed for research institutions, to manage the entire lifecycle of research data



Mendeley Data



Benefits for researchers:

- **Prevent re-work**: save time searching, collecting and sharing data
- **Comply** with funders' mandates
- **Improve impact**: increase data reuse

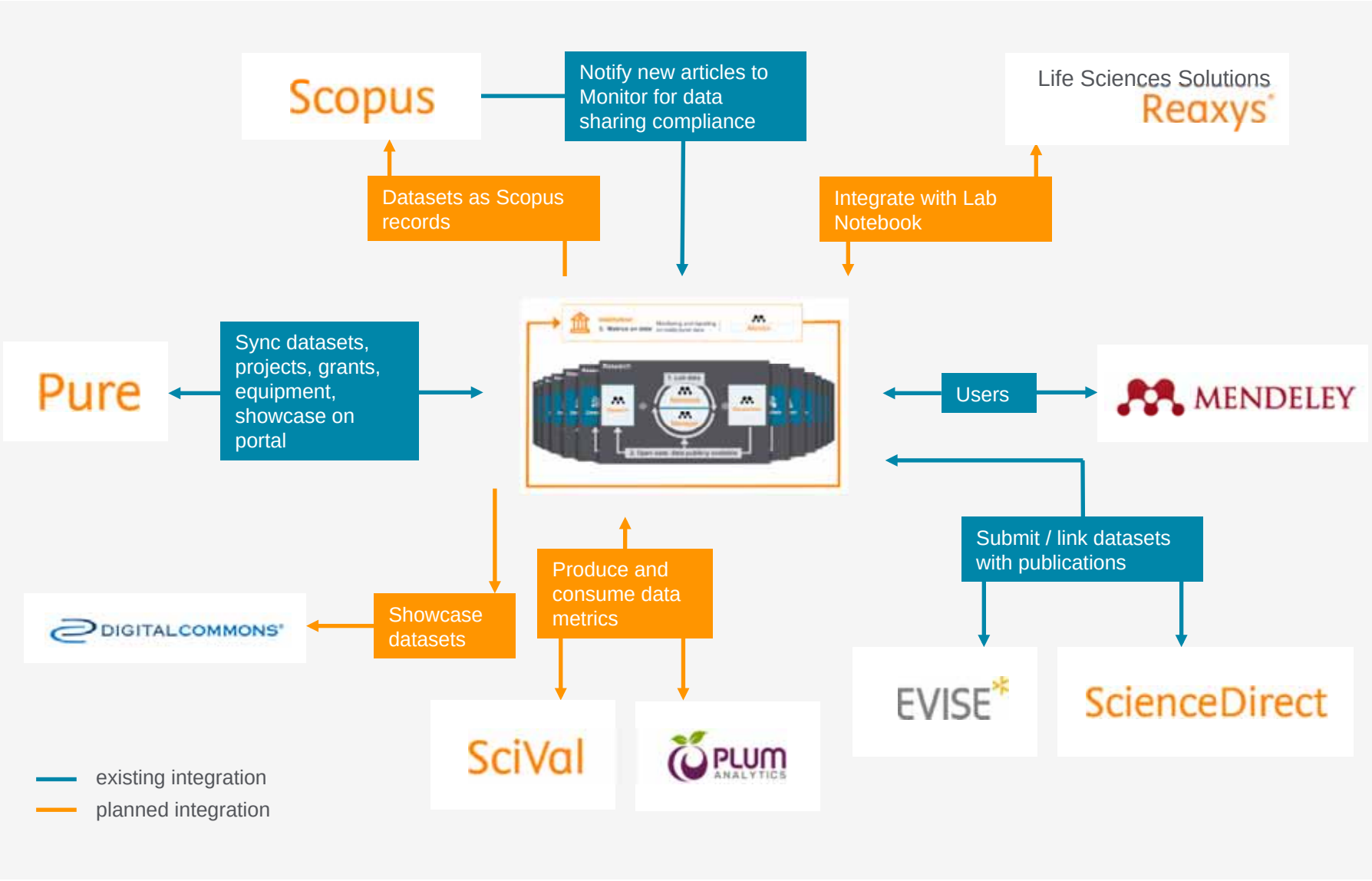
Benefits for institutions:

- **Keep track** of your data inside and outside your institution
- **Showcase** institutional research outputs
- **Improve** collaborations within/across institutions

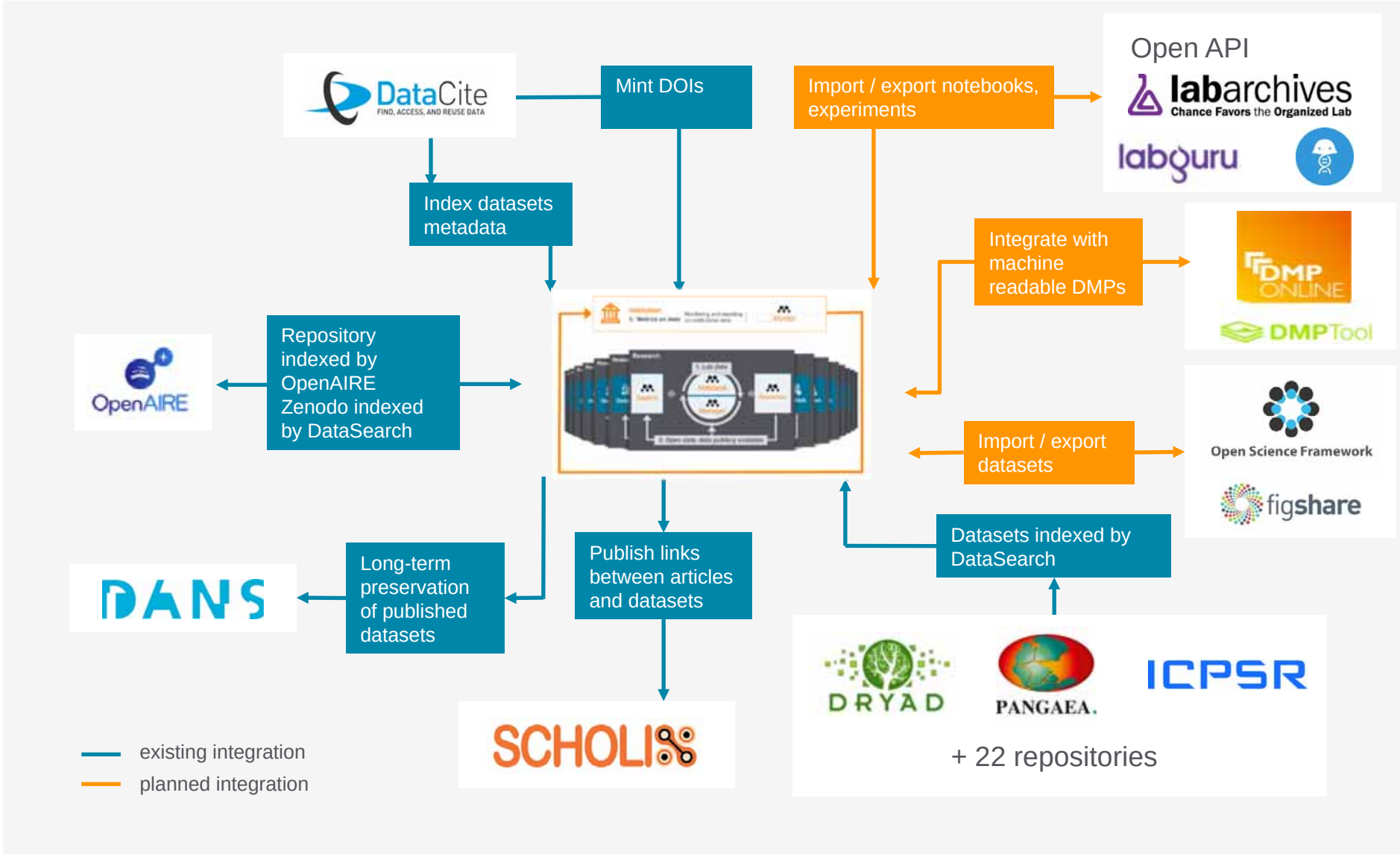
How we deliver:

1. **Open** system & open API's; modular approach enables integrations across many research data solutions
2. **Data** remains at/owned by institution
3. **System** is integrated with the researcher workflows: we make it simple & obvious
4. **Your researchers** keep working like they do today while avoiding additional bureaucracy & administration

Mendeley Data integrates with other Elsevier solutions



Mendeley Data already integrates through open APIs with the global Open Data ecosystem



Mendeley data integrates into your own ecosystem

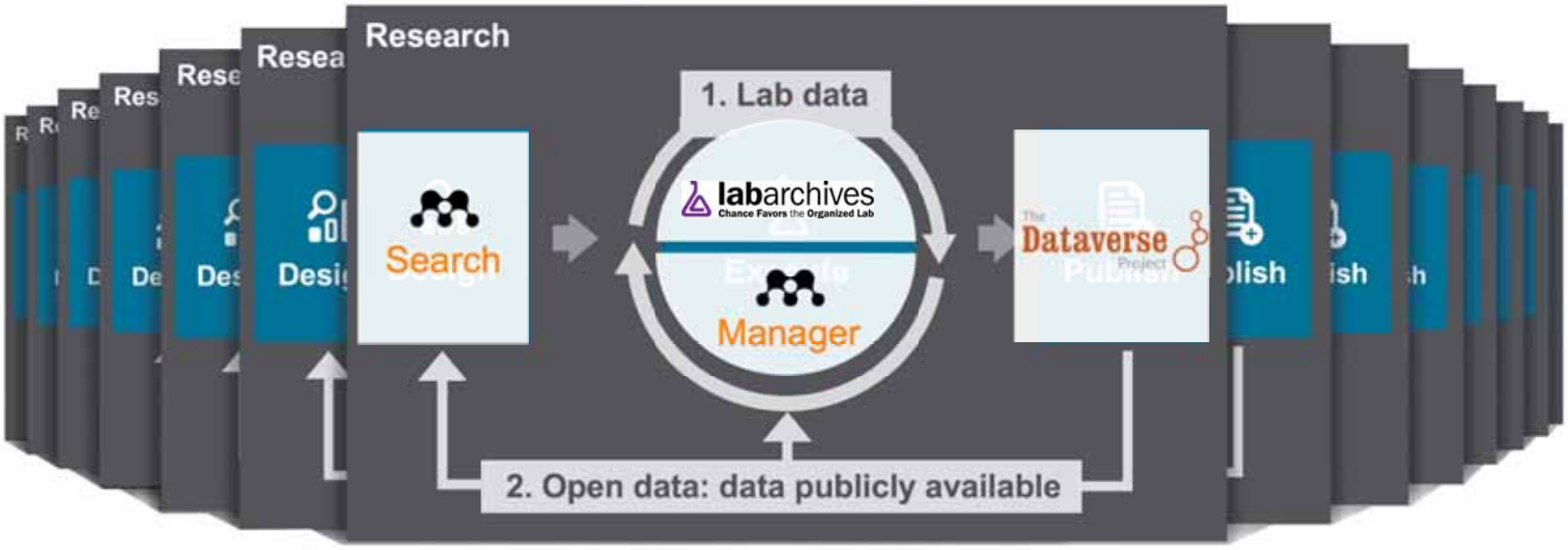
EXAMPLE

Institution

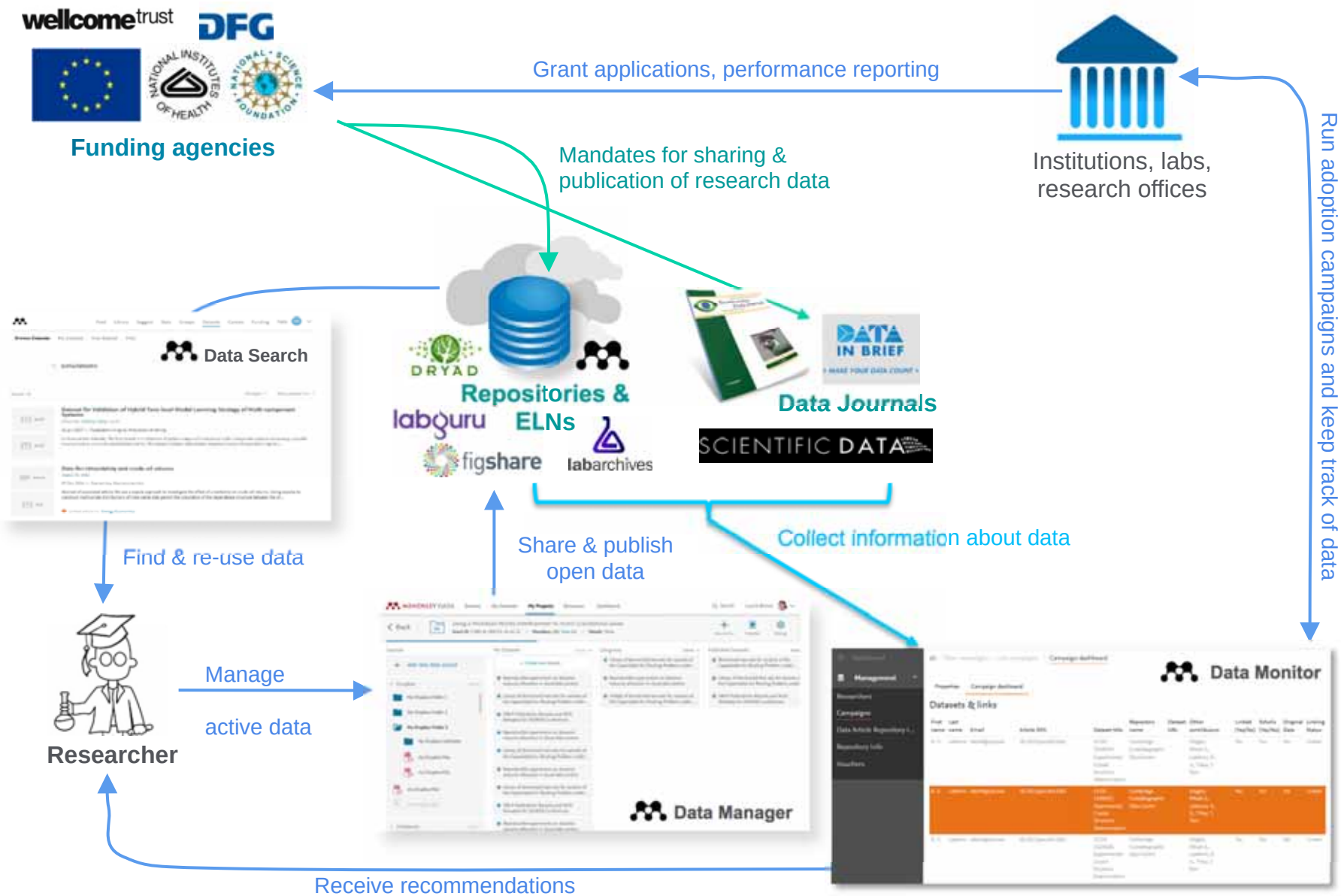
3. ‘Metrics on data’

Monitoring and reporting on institutional data

Monitor

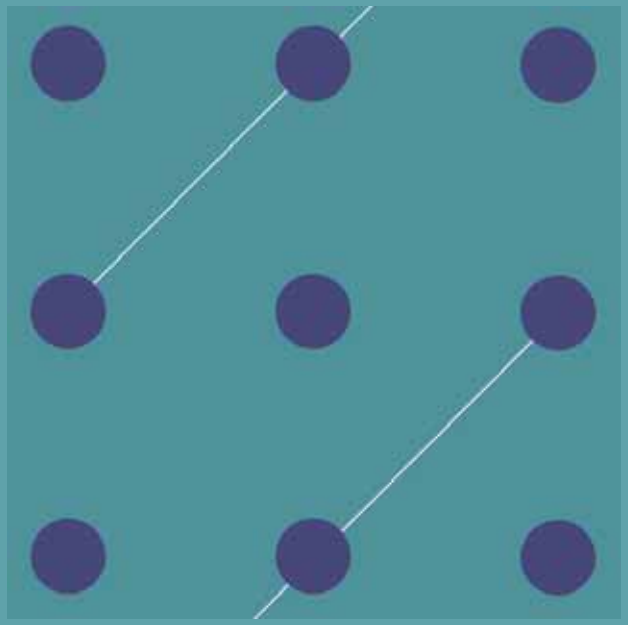


Mendeley Data allows cross-platform tracking of data



OPEN DATA

THE RESEARCHER PERSPECTIVE



Stephane Berghmans
Helena Cousijn
Gemma Deakin
Ingeborg Meijer
Adrian Mulligan
Andrew Plume

Alex Rushforth
Sarah de Rijcke
Clifford Tatum
Stacey Tobin
Thed van Leeuwen
Ludo Waltman

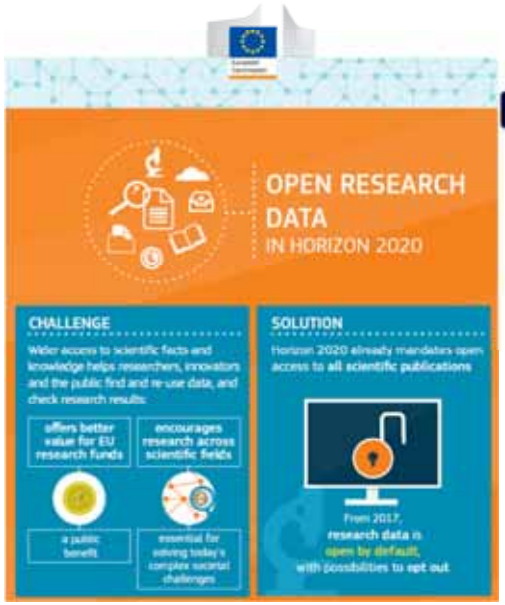
CWTS Meaningful metrics

ELSEVIER

Universiteit Leiden

Data sharing is important for science and society

Funders, associations, and institutes increasingly require data sharing



National Coordination Point Research Data Management

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.

vsnu

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.

Research Questions – *the researcher's perspective?*

1. How are researchers sharing data?

2. Do researchers themselves actually want to share data and/or reuse shared data?

3. Why might researchers be reticent to share their own data openly?

4. What are the effects of new data-sharing practices and infrastructures on knowledge production processes and outcomes?

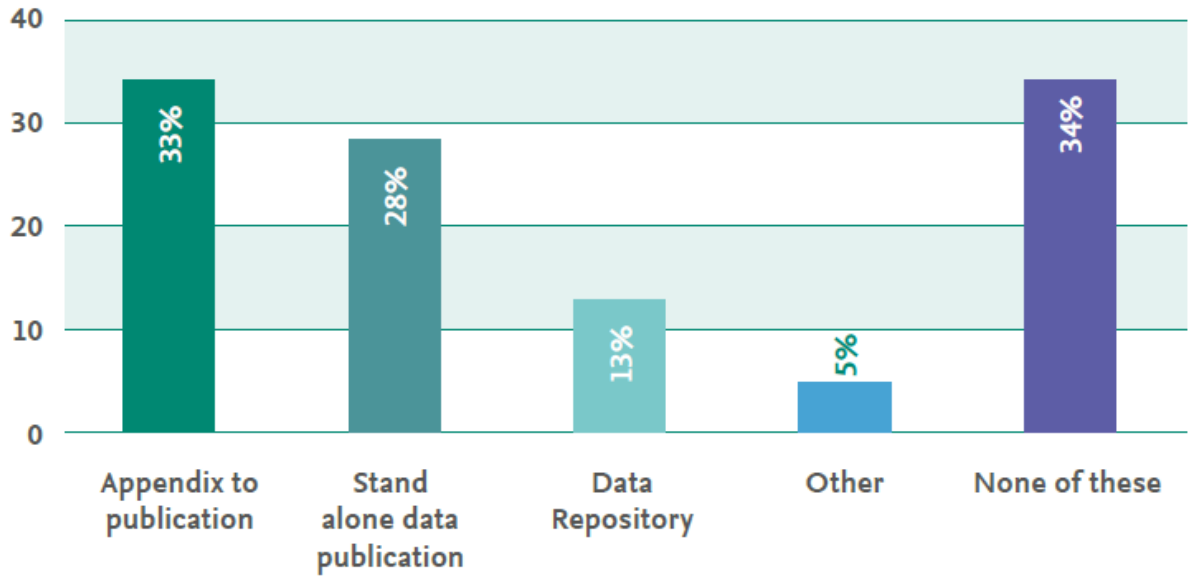
Complementary methods approach

Large-scale global survey

- How is data shared?
- How is data managed?
- How do researchers perceive data sharing?
- How do researchers perceive reusability?

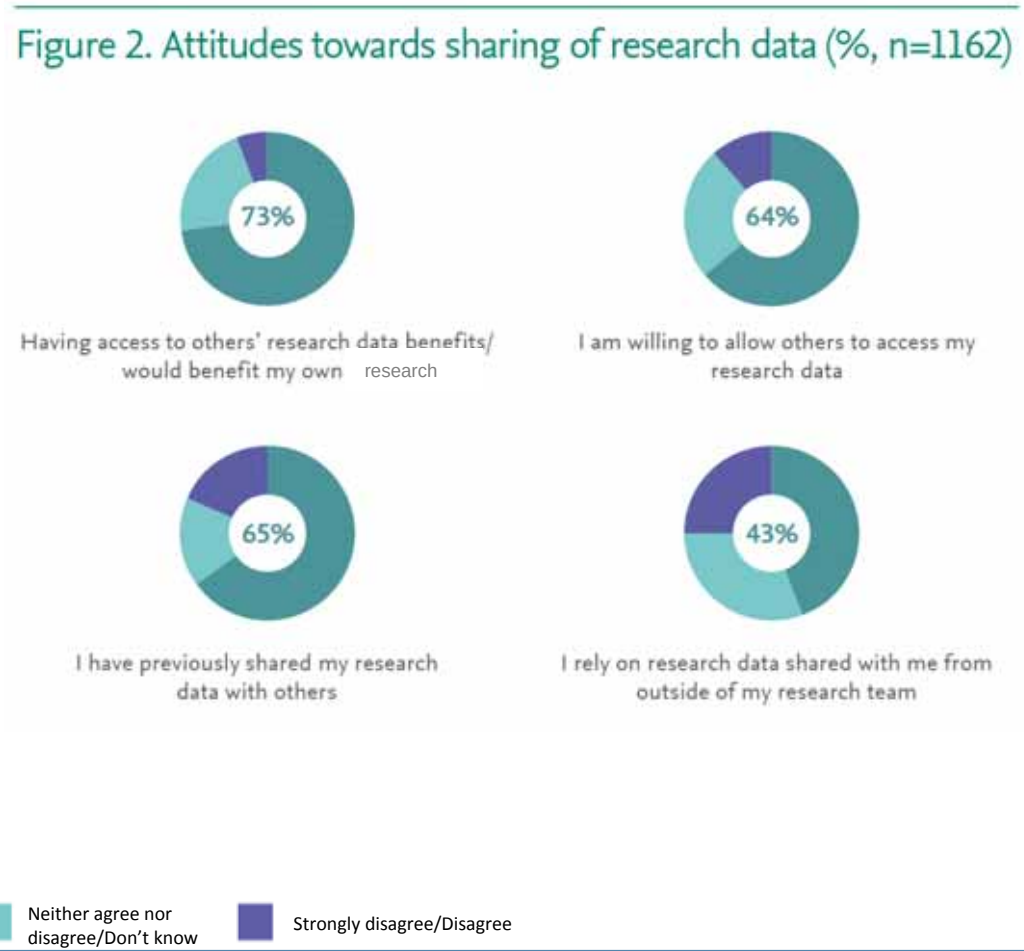
A third of respondents do not publish research data

Figure 1. Dissemination of research data (%, n=1162)



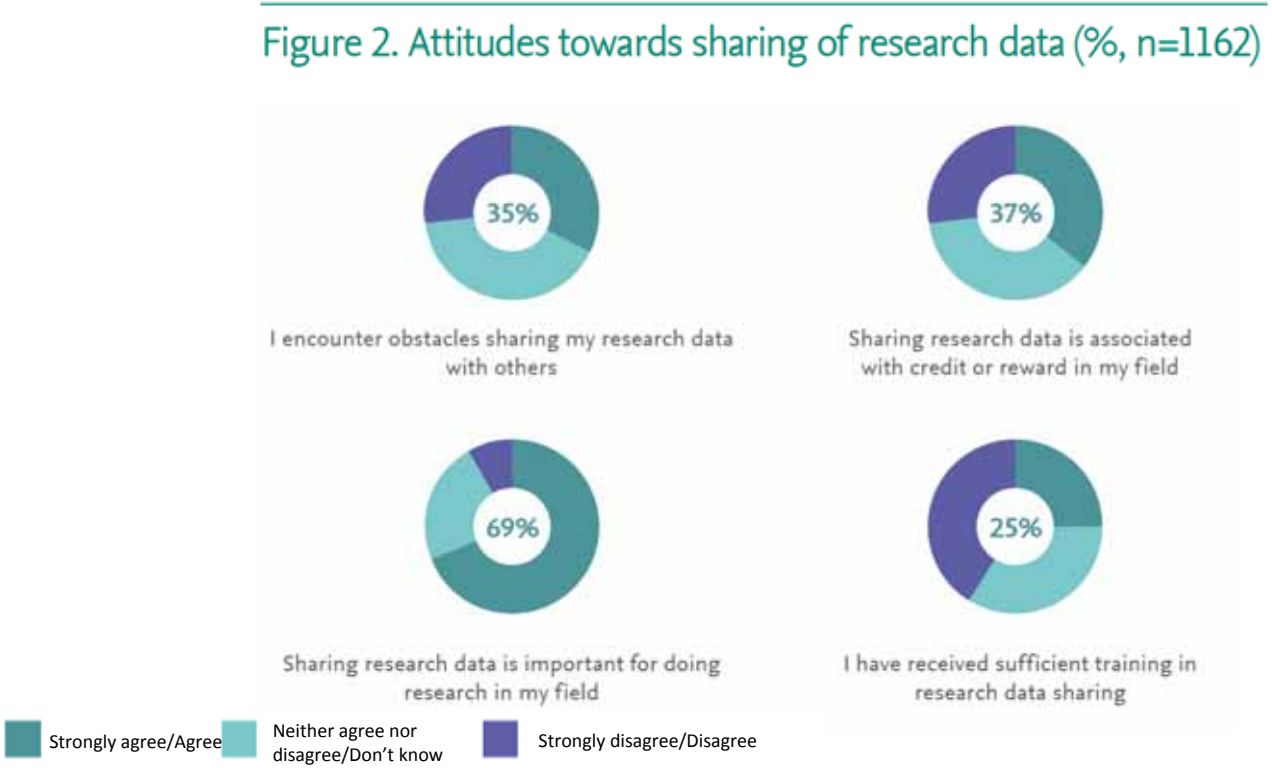
Q: Have you published the research data that you used or created as part of your last research project in any of the following ways?
Note: placing data in a repository is counted as publication

The benefits of sharing research data are clear…



Q: To better understand your attitudes towards research data access, please think about the research data that typically is not published (e.g. not summary charts, tables or images), and indicate how much you agree or disagree with the following statements.

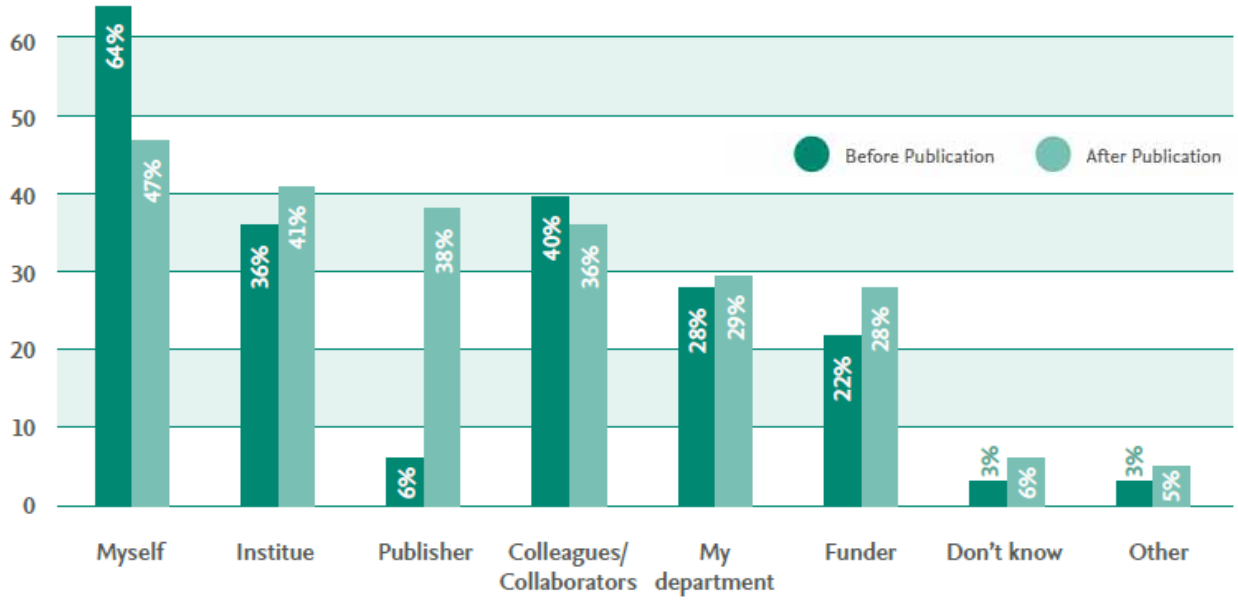
…but obstacles remain



Q: To better understand your attitudes towards research data access, please think about the research data that typically is not published (e.g. not summary charts, tables or images), and indicate how much you agree or disagree with the following statements.

Whose data is it anyway?

Figure 3. Research data ownership before and after publication (%, n=1162)



Q: Who do you believe 'owns' the research data that you have made or will make available to others as part of your last research project?

https://www.elsevier.com/about/open-science/research-data/open-data-report

searcher reprot elsevier

Most Visited DuckDuckGo Trello Gmail GMaps Translate URL Libre Monde NYT Guardian Mendeley Insightly Météo Garmin LinkedIn Twitt

ELSEVIER

SEARCH CART

OPEN DATA

CWTS Meaningful metrics

THE RESEARCHER PERSPECTIVE

Open data: The researcher perspective

The Open Data report is a result of a year-long, co-conducted study between Elsevier and the Centre for Science and Technology Studies (CWTS), part of Leiden University, the Netherlands. The study is based on a complementary methods approach consisting of a quantitative analysis of bibliometric and publication data, a global survey of 1,200 researchers and three case studies including in-depth interviews with key individuals involved in data collection, analysis and deposition in the fields of soil science, human genetics and digital humanities.

Download the report or scroll down to read some of the report's key

www.elsevier.com/about/open-science/research-data/open-data-report

Next steps…

- ☒ 7 April 2017 – 9th RDA Conference, Barcelona

  Launch of the report & website

- ☒ Further dissemination (e.g. ERCEA Seminar)
- ☒ Can we use your survey questionnaire?
- ☒ September 2017 - BOF 10th RDA Conference, Montreal

  'An Open Research Data Survey'

- ☒ RDA Interest Group proposal on an

  'Open Questionnaire for Research Data Sharing Survey'

- ☐ March 2018 – 11th RDA Conference, Berlin

  Kick-off IG Session *(proposal)*

Open Questionnaire for Research Data Sharing Survey

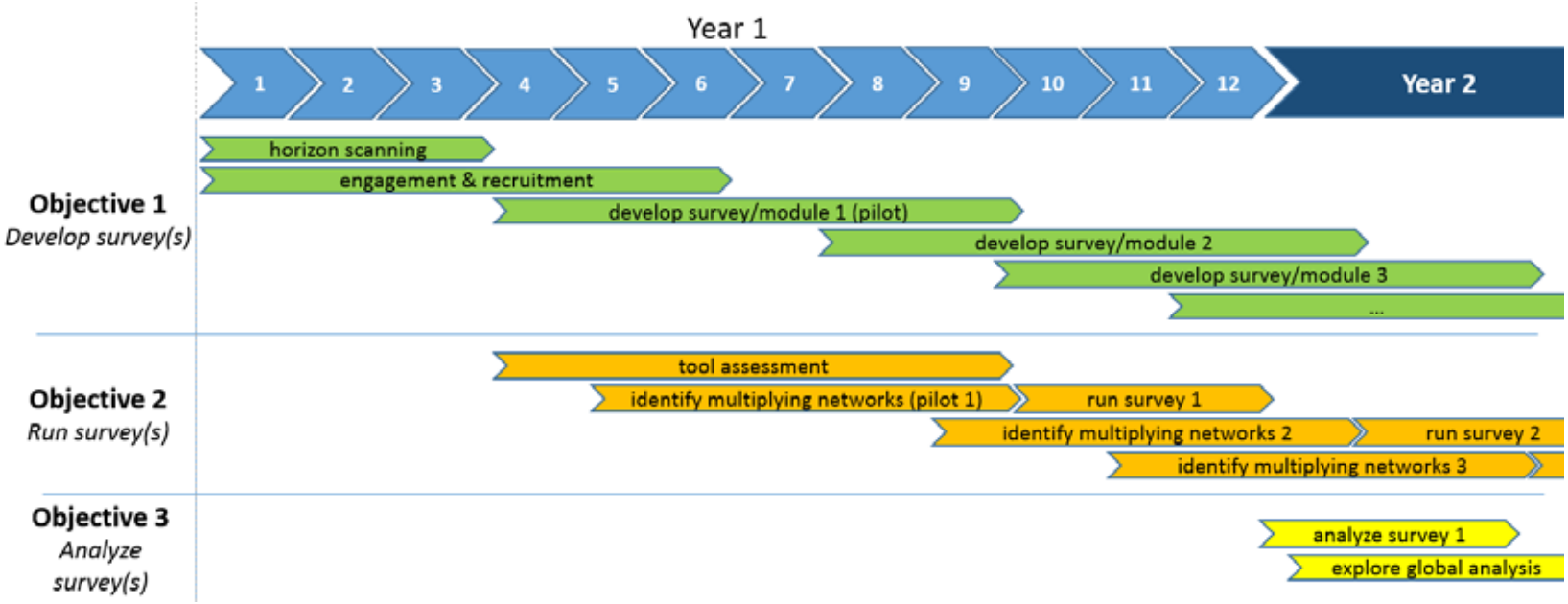
Objective

- Develop a community-designed modular and interoperable open survey(s) questionnaire(s)

Outcomes

- Track changes in practice and policy overtime
- Articulate better policies, identify existing gaps, prioritize research funding, initiate initiatives

RDA Interest Group



Co-Chairs

Ingeborg Meijer, CWTS, Leiden University, The Netherlands

David O'Brien, IDRC (International Development Research Center), Canada

Kazuhiro Hayashi, NISTEP (National Institute of Science and Technology Policy), Japan

Stephane Berghmans, Elsevier, Belgium

Thank you

Stephane Berghmans
Vice President, Academic & Research Relations
Elsevier
s.berghmans@elsevier.com
+32-471-781900