

# RDM AND DATA STEWARDSHIP:

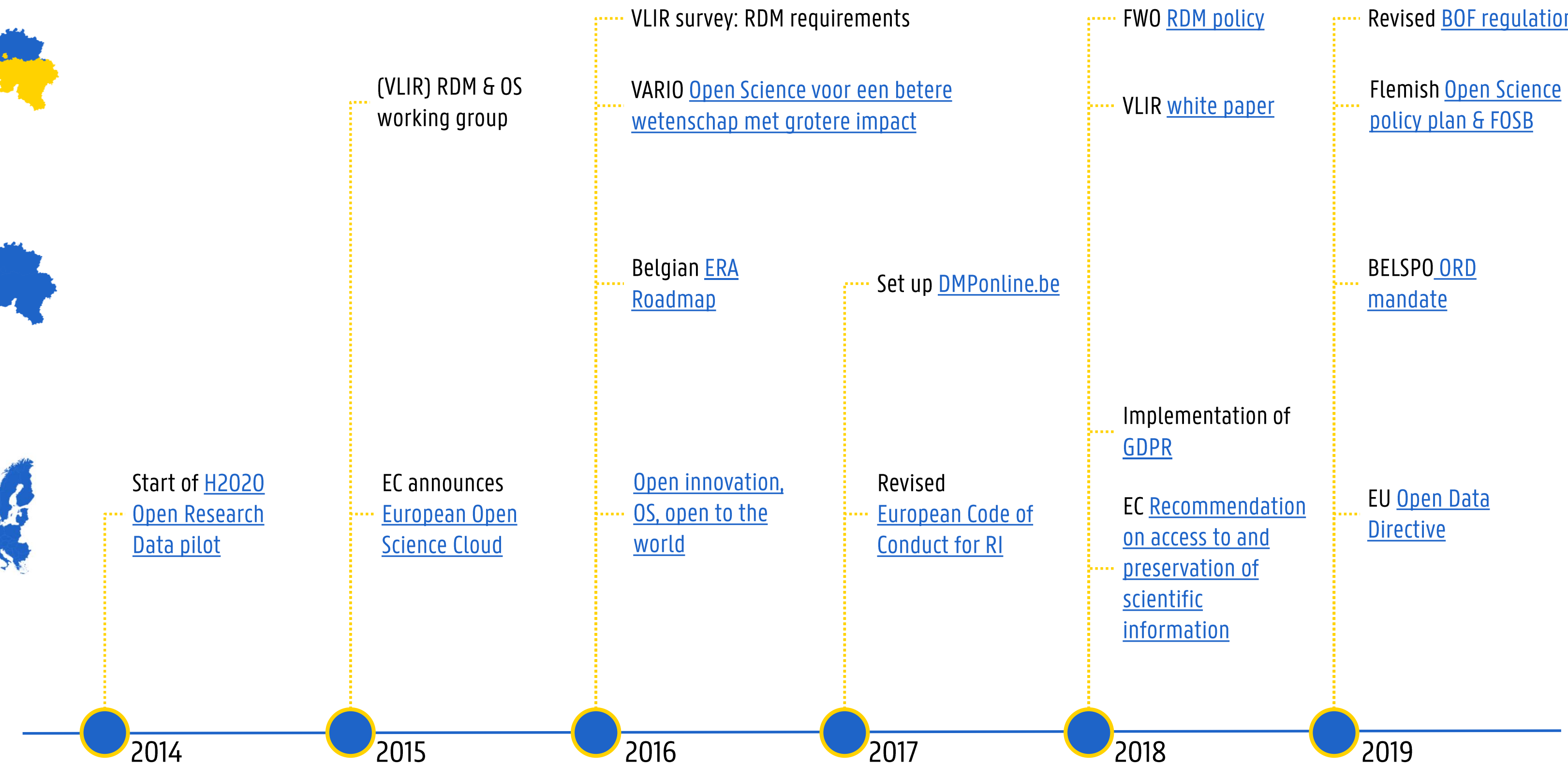
## KEY DRIVERS, TRENDS AND REQUIREMENTS

Advancing data stewardship: insights from environmental and life science research infrastructures  
2020-02-28, VLIZ (Ostend)

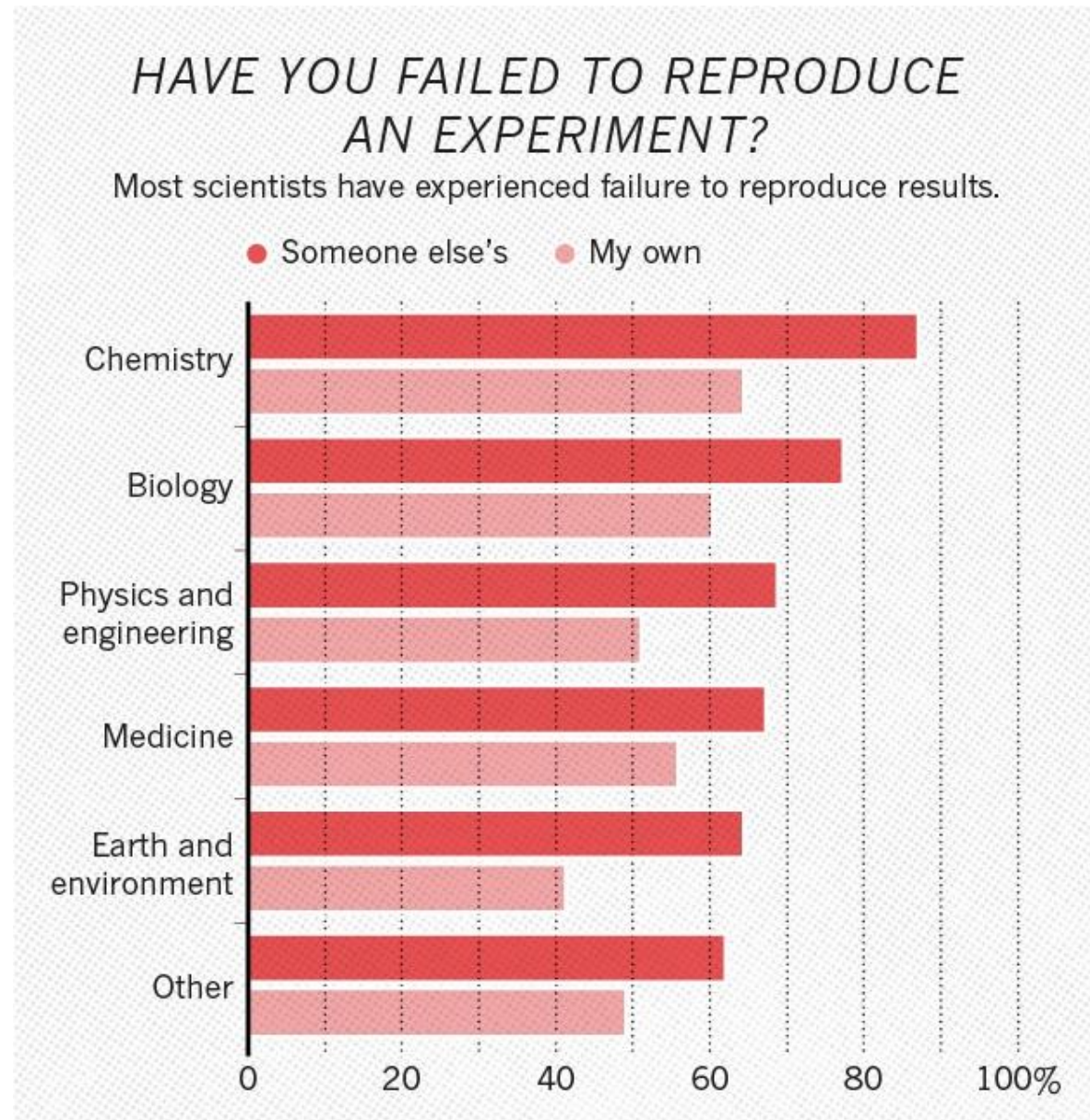
# OUTLINE

- RDM: key drivers and requirements
- The researchers' perspective
- Open Science enablers: trends
- Building a data stewardship community

# RDM: KEY DRIVERS & REQUIREMENTS



# RESEARCH INTEGRITY



> 50% researchers had not been able to reproduce their own experiments

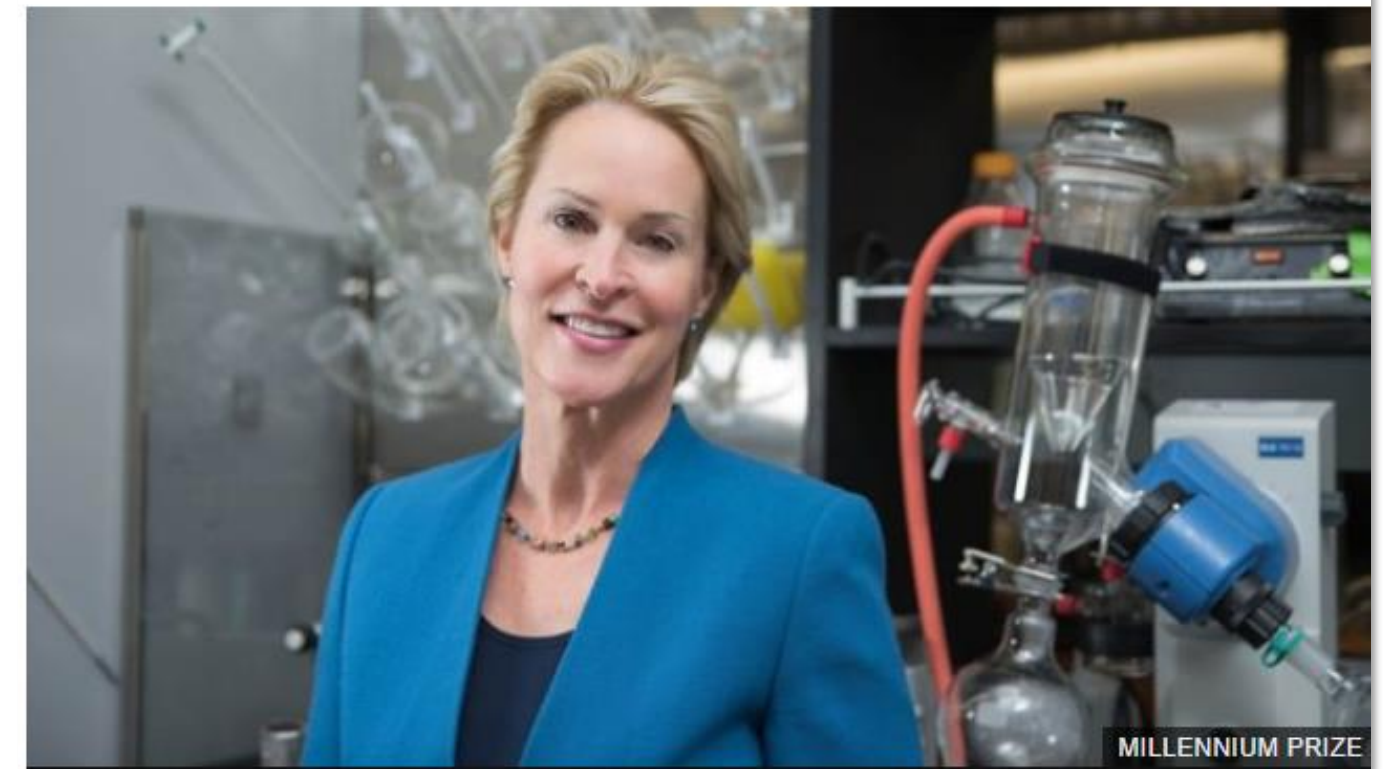
<https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>

## Nobel Prize-winning scientist Frances Arnold retracts paper

🕒 3 January 2020

f 🗨️ 🐦 ✉️ Share

Nobel Prize



MILLENNIUM PRIZE

“(...) the results were not reproducible, and the authors

found data missing from a lab notebook”.

<https://www.bbc.com/news/world-us-canada-50989423>



# EUROPEAN CODE OF CONDUCT FOR RESEARCH INTEGRITY

- Expectations regarding research data include:
  - secure preservation for a reasonable period
  - access = as **open** as possible, as closed as necessary
    - + in line with **FAIR** principles
  - data = legitimate & citable products of research



<http://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf>

# GDPR

- Generally perceived to have a negative impact on research.
  - Ambiguities and differences in interpretation yet to be clarified (e.g. consent and public interest, “disproportionate effort”).
- However: possible opportunities to improve practices (e.g. security), clarify responsibilities, EU harmonization may improve collaboration.
- Exposes the need for data management good practices
  - “To ensure that adequate data management practices are installed and respected, research institutions might consider the appointment of data stewards as part of their organisational chart”<sup>1</sup>.

<sup>1</sup> [How the general data protection regulation changes the rules for scientific research](#) (2019), Directorate-General for Parliamentary Research Services (European Parliament).

DOI 10.2861/17421

# EUROPEAN OPEN INNOVATION AND OPEN SCIENCE

- EC's investment plan for Europe: “fewer barriers to knowledge transfer, open access to scientific research” needed to boost innovation and EU competitiveness<sup>1</sup>.
- Some estimates:
  - Benefits of EMBL-EBI ~£1 billion p/year (>x20 operational costs)<sup>2</sup>
  - “The cost of not having FAIR research data costs the European economy at least €10.2bn every year”<sup>3</sup>.



<sup>1</sup> Communication ‘[An Investment Plan for Europe](#)’ COM(2014)903

<sup>2</sup> [The Value and Impact of the European Bioinformatics Institute](#)

<sup>3</sup> [Cost-benefit analysis for FAIR research data](#) (2019), Directorate-General for Research and Innovation (European Commission), PwC EU Services. DOI 10.2777/02999



# OPEN DATA DIRECTIVE (2019) AND NATIONAL POLICIES

- “Member States shall support the availability of research data by adopting national policies and relevant actions aiming at making publicly funded research data openly available (‘open access policies’) following the principle of ‘open by default’ and compatible with FAIR principles
  - Member States have to transpose Directive (EU) 2019/1024 by 16 July 2021

2014	Policy exists	Underway	No policy / unknown
	BE, CY, DK, EE, FR, FI, DE, LT, NL, PT, UK, NO, CH	AT, BG, HR, IE, IT, PL, SI, ES, SE, IS	CZ, EL, HU, LV, LU, MT, RO, SK, GR
2019	Policy exists	Underway	No policy / unknown
	BE, CY, DK, EE, FR, FI, DE, LT, NL, PT, UK, NO, CH, IS, RS	BG, EE, GR, HR, IE, IT, LU, PL, RO, SK, SE	LV

SPARC Europe, & Digital Curation Centre. (2019, August 28). An Analysis of Open Science Policies in Europe v4. Zenodo. <http://doi.org/10.5281/zenodo.3379705>

# REQUIREMENTS FROM RESEARCH FUNDERS



FWO

- Proposal stage: “Light” DMP.
- Post-award: DMP after 6 months + (updates) + final DMP with final report (evaluation).
- Emphasis on **long-term data preservation**: 5 year min.
- Gradual implementation. All applicants >2020.



BELSPO

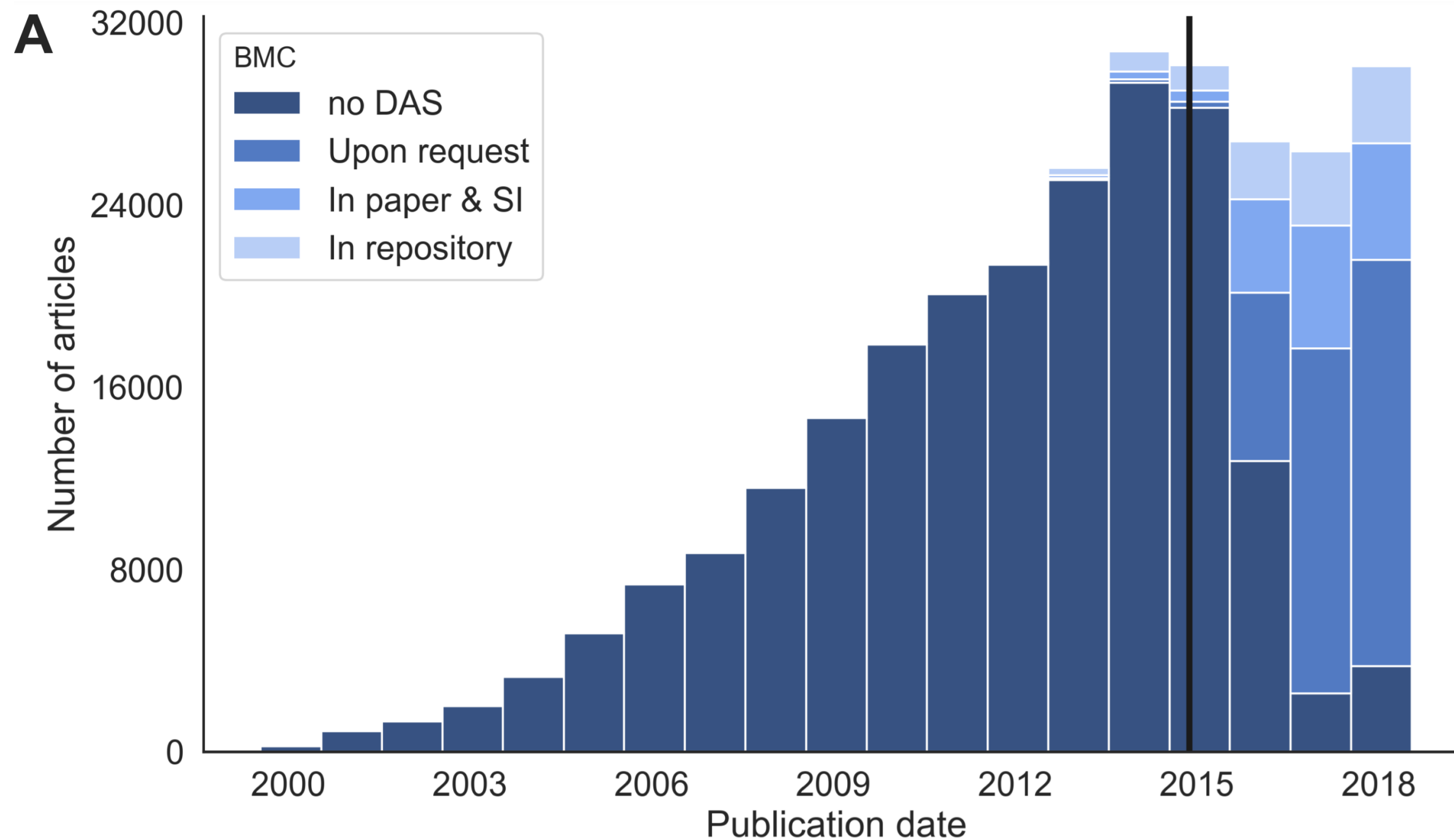
- Proposal stage: provisional DMP.
- Post-award: DMP after 6 months + (updates) + final DMP with final report (evaluation).
- Emphasis on **open data** (ORD mandate > 2020) & reuse.
- **FAIR**, CCO recommended. “As open as possible, as closed as necessary”.



H2020

- Proposal stage: no DMP.
- Post-award: DMP after 6 months + (updates) + final DMP with final report.
- “Encouraged” when opting out of the ORD pilot.
- Emphasis on **open data**, data deposition and **FAIR** principles.

# JOURNALS: DATA AVAILABILITY REQUIREMENTS

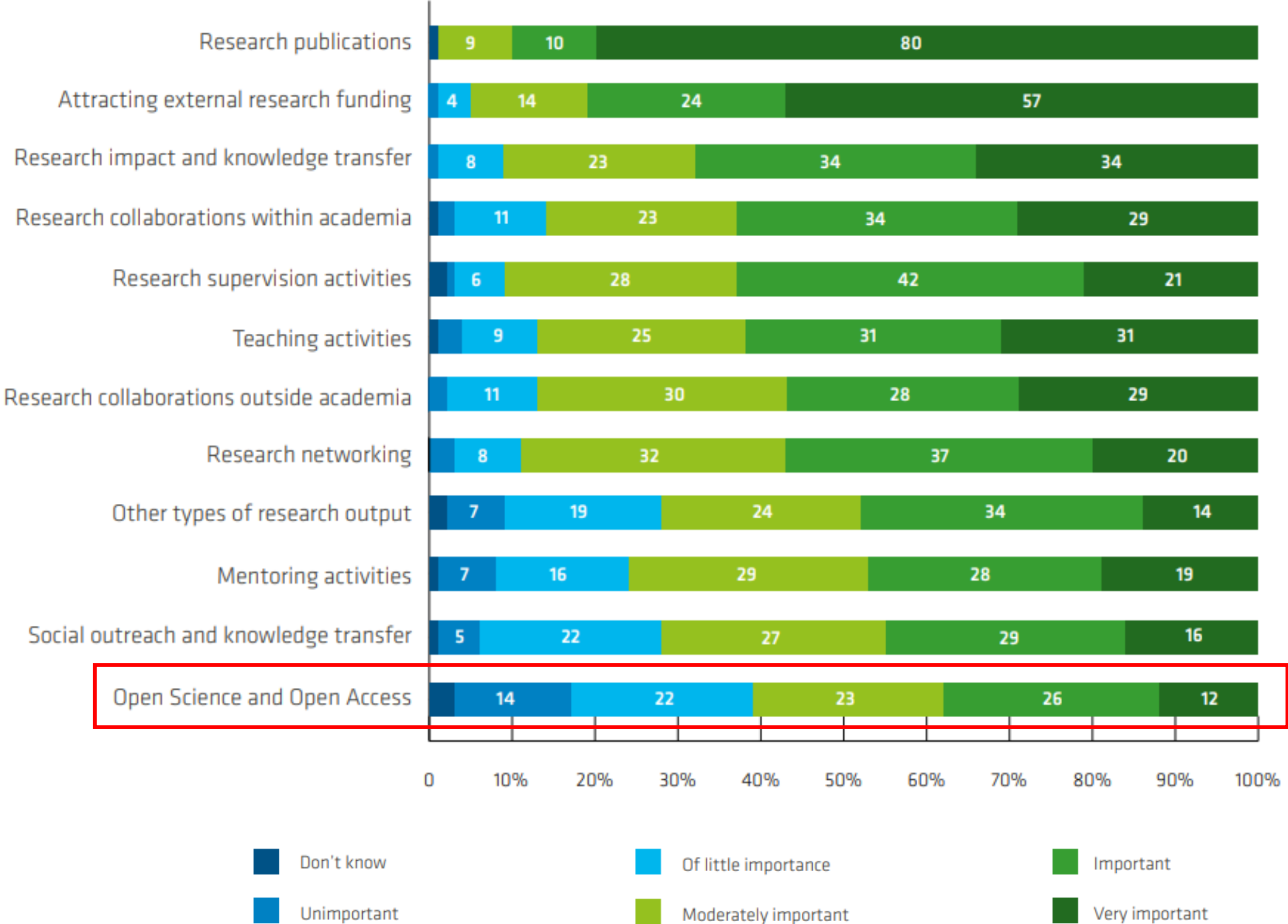


# RESEARCHERS' PERSPECTIVE

# PERCEPTION OF THE IMPORTANCE OF OPEN SCIENCE

## Importance of academic activities for research careers

[Research Assessment in the Transition to Open Science](#)  
[2019 EUA Open Science and Access Survey Results](#)

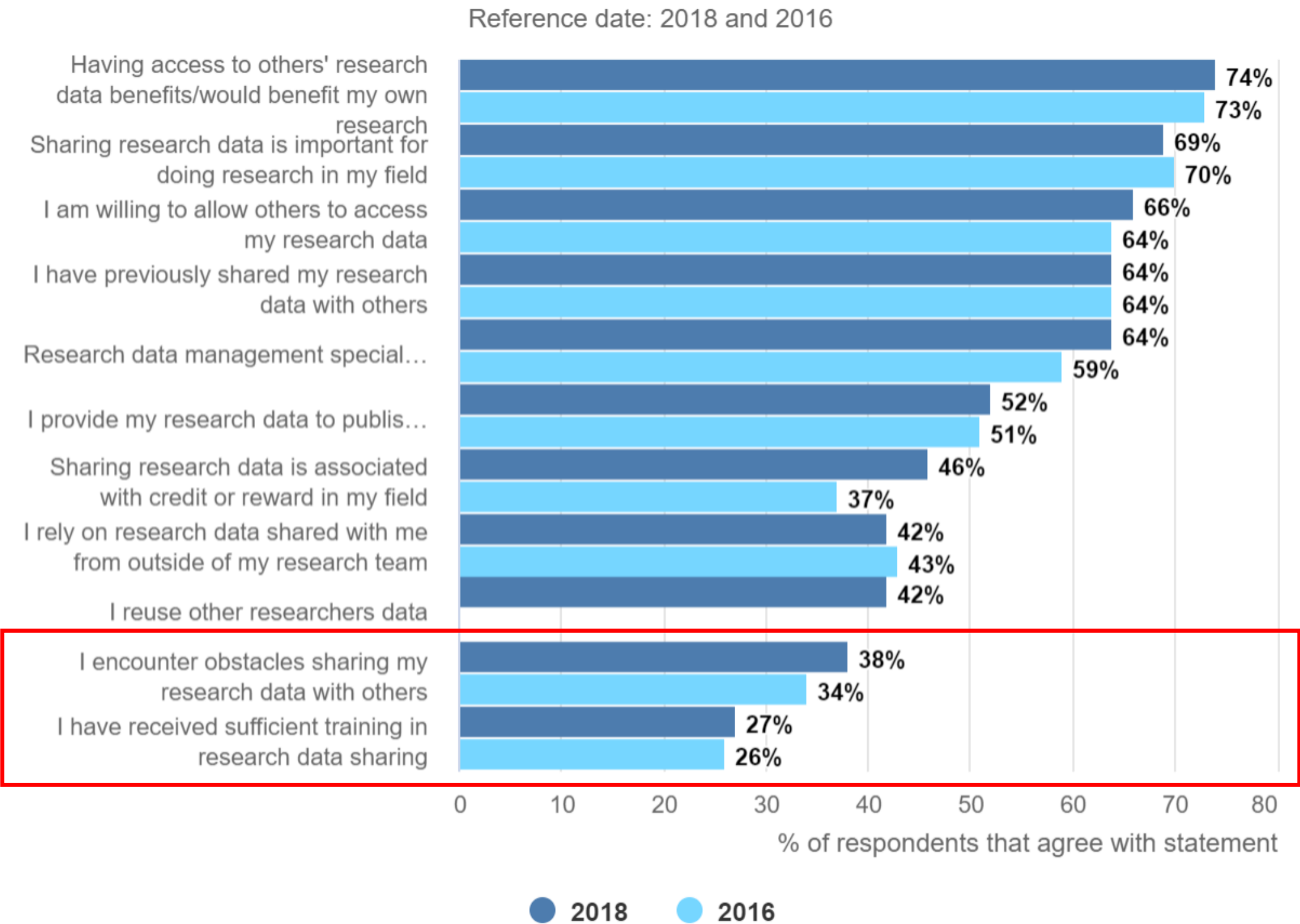




# DATA SHARING ATTITUDES

% of respondents (researchers) that agree with the statement

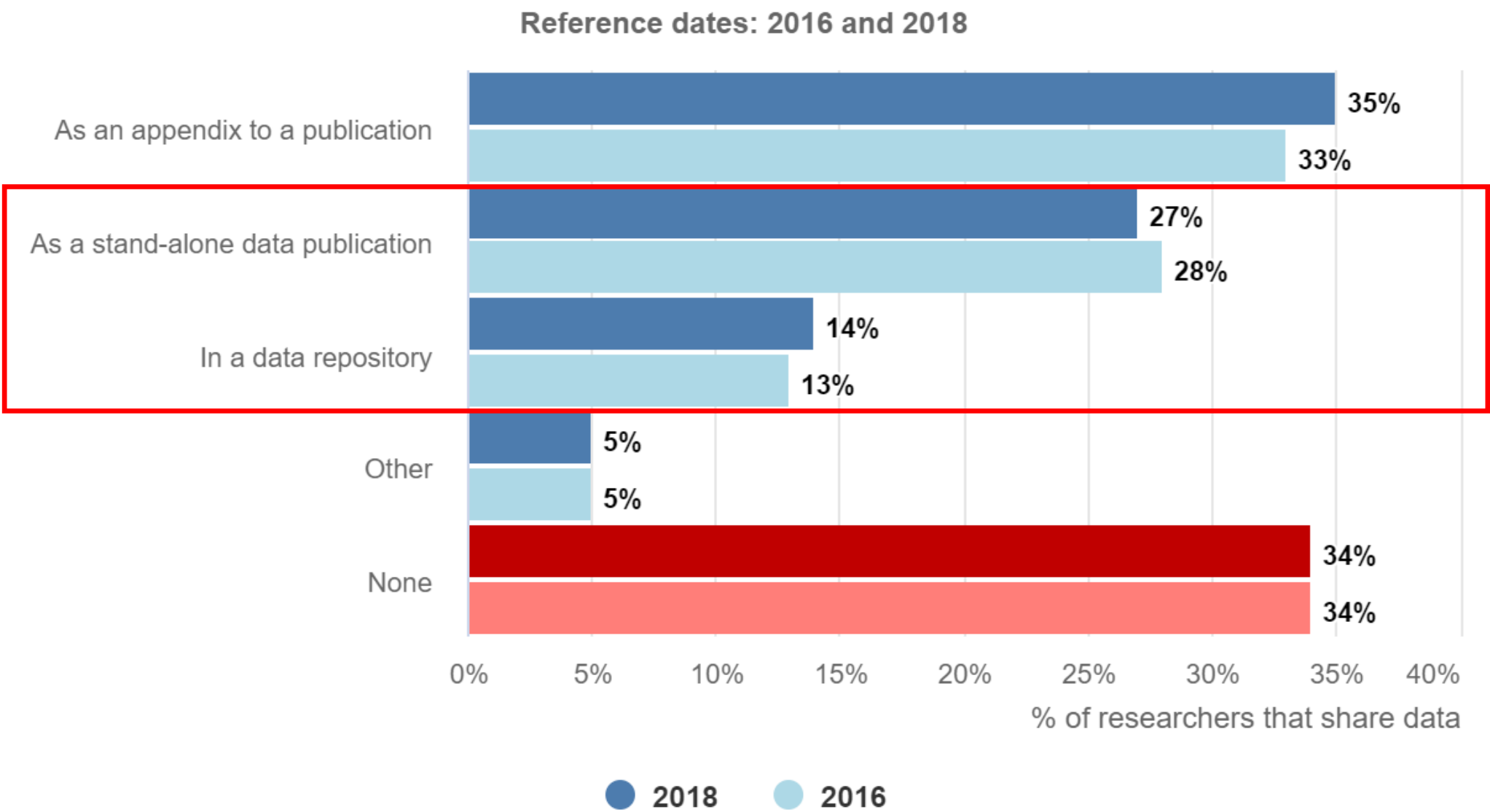
[Facts and Figures for open research data](#)



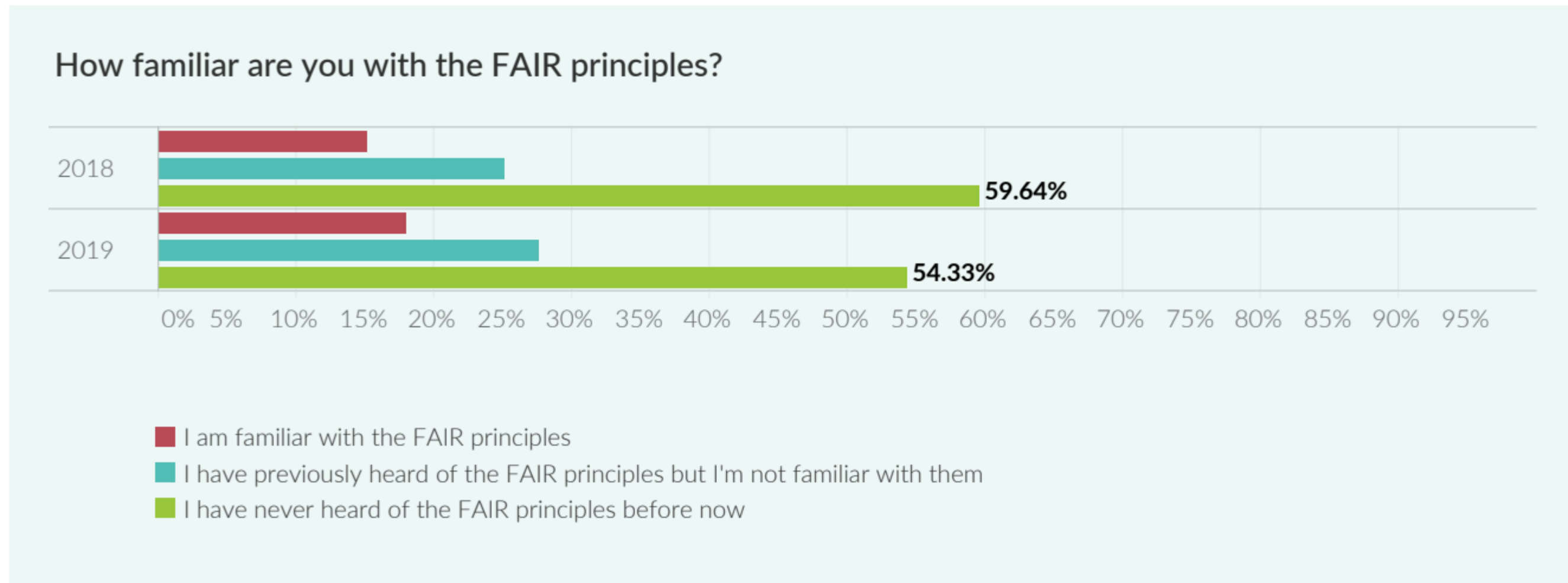
# DATA SHARING PRACTICES

% of respondents (researchers) that share data, by modality

[Facts and Figures for open research data](#)



# FAIR AWARENESS



“FAIR principles” are relatively unknown to the community with 52% of respondents who are frequent data-sharers never having heard of them.

Science, Digital; Fane, Briony; Ayris, Paul; Hahnel, Mark; Hrynaszkiewicz, Iain; Baynes, Grace; et al. (2019): The State of Open Data Report 2019. figshare. Report.  
<https://doi.org/10.6084/m9.figshare.9980783.v2>

# SKILLS AND KNOWLEDGE GAP

- >500,000 ‘Core Data Experts’ needed to effectively operate the EOSC within a decade<sup>1</sup>.
- There is an urgent need to develop skills in relation to FAIR data<sup>2</sup>
  - Data science + data stewardship
  - Researchers need a **foundational**-level set of data **skills**
  - Researchers should be provided with **professional support**
- Skills gaps is widely recognized at national and international level<sup>3, 4, 5</sup>
  - “The human resources challenge”<sup>6</sup>

<sup>1</sup> [Realising the European Open Science Cloud](#) (2016), EC High Level Expert Group on the EOSC. DOI 10.2777/940154

<sup>2</sup> [Turning FAIR into reality](#) (2017), EC Expert Group on FAIR data. DOI 10.2777/1524

<sup>3</sup> [VLIR white paper](#) (2018), VLIR Werkgroep Research Data Management & Open Science

<sup>4</sup> [COMMISSION RECOMMENDATION](#) (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information

<sup>5</sup> [Open innovation, open science, open to the world](#) (2016), Directorate-General for Research and Innovation (European Commission). DOI 10.2777/061652

<sup>6</sup> [Nota aan de Vlaamse Regering. Vlaamse beleidsplan Open Science en oprichting Open Science Board](#) (2019), EWI.

# OPEN SCIENCE ENABLERS: TRENDS

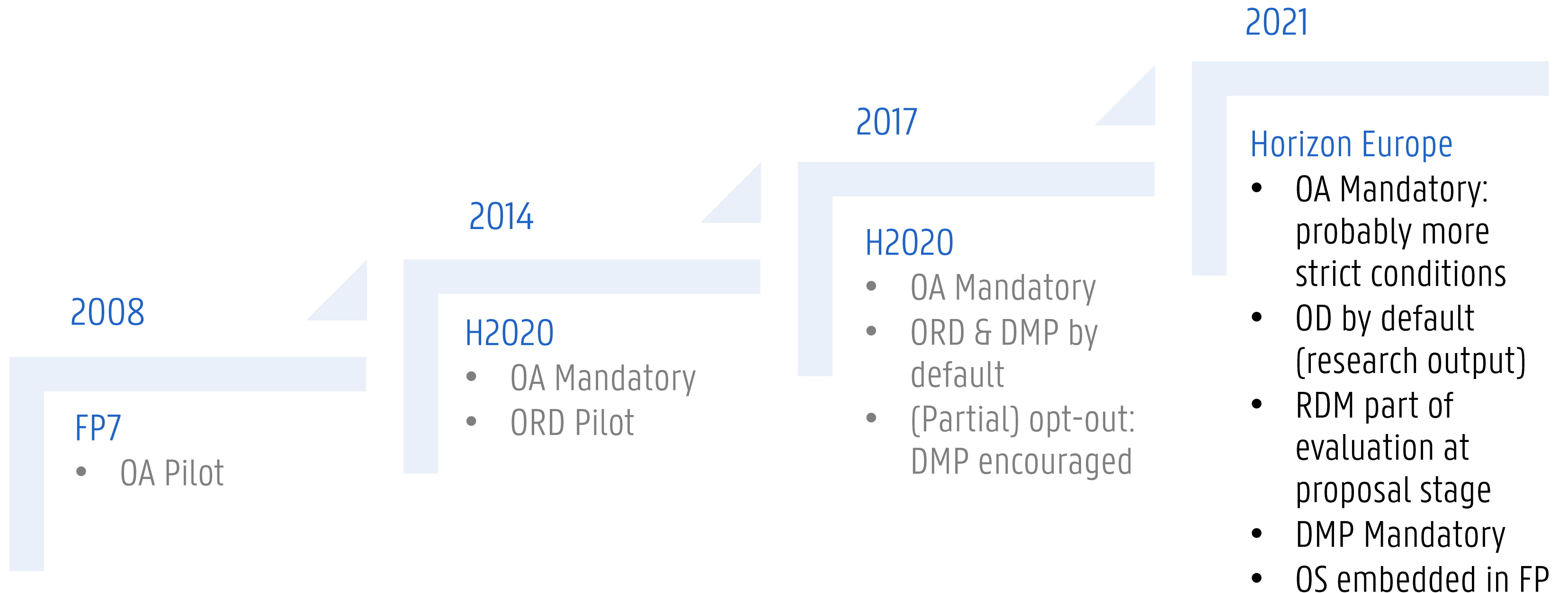




**EUROPEAN OPEN  
SCIENCE CLOUD**

# FUNDERS: HORIZON EUROPE

Glinos, Kostas. (2020, February). Making Open Science the new normal. Zenodo.  
<http://doi.org/10.5281/zenodo.3688670>



# JOURNAL POLICIES

- Many journals and publishers are adopting the TOP guidelines similar frameworks of [tiered data availability policies](#).
- In practice, data not fully available if only “recommended” or “encouraged”<sup>1</sup>.
- Publishers to consider how to create policy and review procedures that the proportion of articles for which complete data sets are available<sup>1</sup>.
- Need to [standardize research data policies](#) by publishers<sup>2</sup>.

+ Level 0 - Encourage

+ Level 1 - Disclose

+ Level 2 - Require

+ Level 3 - Verify

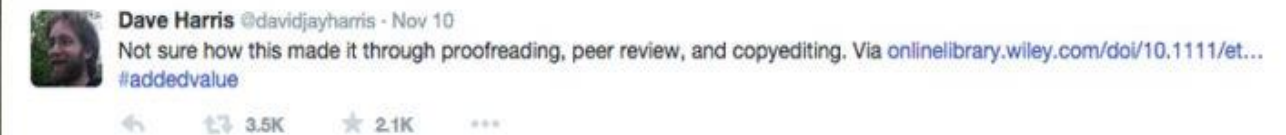
<sup>1</sup> Federer LM, Belter CW, Joubert DJ, Livinski A, Lu YL, et al. (2018) Data sharing in PLOS ONE: An analysis of Data Availability Statements. PLOS ONE 13(5): e0194768. <https://doi.org/10.1371/journal.pone.0194768>

<sup>2</sup> Hrynaskiewicz, I. et al. [Developing a Research Data Policy Framework for All Journals and Publishers](#) (2019), RDA [Data policy standardization and implementation Interest Group](#) (IG)

# REWARDS AND INCENTIVES FOR OPEN SCIENCE

- Shift from ‘publishing as fast as possible’ to ‘sharing knowledge as early as possible’.
- EC [Expert Group on Reward Systems](#) in Open Science
  - “Open Science must become part of the recruitment criteria, career progression and grant assessment procedures for researchers at all levels”.
- Initial [agreement](#) on Horizon Europe:
  - “Recognition and reward mechanisms [...], as well as incentives promoting the adoption of open science practices [...]”.

Although association preferences documented in our study theoretically could be a consequence of either mating or shoaling preferences in the different female groups investigated ([should we cite the crappy Gabor paper here?](#)), shoaling preferences are unlikely drivers of the documented patterns both because of evidence from previous research and inconsistencies with *a priori* predictions. Our methods closely followed those of published mate choice experiments in this system (Tobler et al. 2009a,b; Plath et al. 2013),





# SKILLS AND SUPPORT

- Researchers need training on data science and data stewardship:
  - Data skills should be core curricula for researchers.
  - Domain specific knowledge and community best practices: important role of ESFRIs.
- Need to professionalize support / data stewardship roles:
  - EC projects have worked on curricular frameworks for data stewardship & curation, data science, skills for EOSC
    - [DigCurV](#)
    - [EDISON](#)
    - [EOSCPilot](#)
- Data stewards / RDM support roles are on the rise.

## Setting up a **data steward team** at Ghent University

### WHAT?

Since October 2019, Ghent university has a team of **six full-time data stewards**. Their mission:

-  Promote & facilitate good RDM practices
-  Contribute to cultural change towards FAIR & open data, data reuse

### WHY?

- Human, not just technical RDM infrastructure required
- Offer professional, more tailored support
- Ensure consistent messages & optimal information exchange

### HOW?

#### 1 Making the case

To get buy-in from our Board of Governors, we pointed to:

- the changing legal & data policy landscape
- the RDM knowledge & skills gap<sup>1</sup> in our researchers
- the emergence of data stewardship programs elsewhere (especially in the Netherlands)<sup>2</sup>

#### 3 Recruiting

With no formal learning for data stewards yet, we looked for:

- research experience and keen interest in RDM (support)
- variety of disciplines & skill sets
- relevant attributes as identified in other institutions' RDM job ads

#### 4 Getting started

We set out with an initial 3-phased work plan:

- 1 **Onboarding** via self-study, group discussions & meetings with key central services
- 2 Start **faculty outreach** to introduce data stewards & gather input for phase 3
- 3 Launch more **proactive support** with advisory & training services

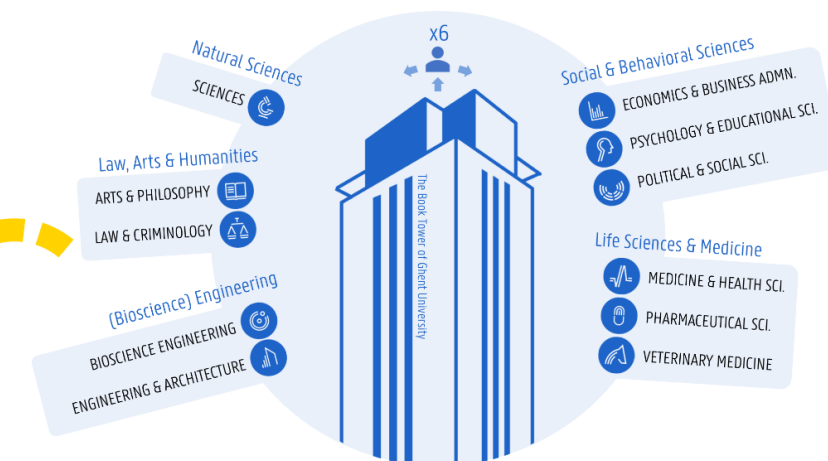
-  DMP support
-  RDM helpdesk
-  Consultations
-  Online information & guidance
-  Training, info-sessions and other events
-  User support for Library RDM tools

#### 2 Shaping the local implementation of data stewardship

How to implement this new role within our institution? We drew inspiration from Dutch data steward landscape studies<sup>3,4</sup>.

##### 2.1 Between central & faculty level

Team within **Ghent University Library** (Book tower), with one team coordinator and one data steward for each **faculty cluster**:



##### 2.2 Three main task areas

*Advisors, not operational data managers!*

- 1 Raise awareness, inform, train & advise
- 2 RDM infrastructure-related projects
- 3 Monitor & shape RDM policy and strategy

#### Author:

Myriam Mertens  
myriam.mertens@ugent.be

#### More information:

<https://osf.io/bhy8w/>  
@UGentRDM

#### References:

1. VLIR Werkgroep RDM & OS (2017), Handle: 1854/LU-8555410
2. Teperek et al. (2018), DOI: [10.17605/OSF.IO/5W9PJ](https://doi.org/10.17605/OSF.IO/5W9PJ)
3. Verheut et al. (2019), DOI: [10.5281/zenodo.2642066](https://doi.org/10.5281/zenodo.2642066)
4. Scholtens et al. (2019), DOI: [10.5281/zenodo.3474789](https://doi.org/10.5281/zenodo.3474789)



# BUILDING A DATA STEWARDSHIP COMMUNITY

# BUILDING A DATA STEWARDSHIP COMMUNITY



- Share experiences, share training resources.
- Keep the network informed about RDM developments.
- Identify common challenges & bottlenecks and find shared solutions.
- Advance towards professionalizing data stewardship and give visibility to our efforts.
- Experience in the NL:
  - [Data Stewards Interest Group](#). “Hands-on and solution-oriented approach”.
  - [LCRDM Tasks Groups](#)

# UGent Data Stewards

RESEARCH DEPARTMENT - UNIVERSITY LIBRARY



[rdm.support@ugent.be](mailto:rdm.support@ugent.be)



[@UGentRDM](https://twitter.com/UGentRDM)



[ugent.be/en/research/datamanagement](https://ugent.be/en/research/datamanagement)