SPRINGER NATURE

Five essential factors for data sharing

Iain Hrynaszkiewicz

Head of Data Publishing

JOSS, Tokyo
27 May 2019

ADVANCING DISCOVERY

A global perspective on data sharing

We have undertaken and shared survey research to understand the impact of data sharing, and to understand researcher attitudes towards data sharing

To accelerate data sharing, we propose five essential factors:

1. Clear policy
2. Better credit
3. Explicit funding
4. Practical help
5. Training & education
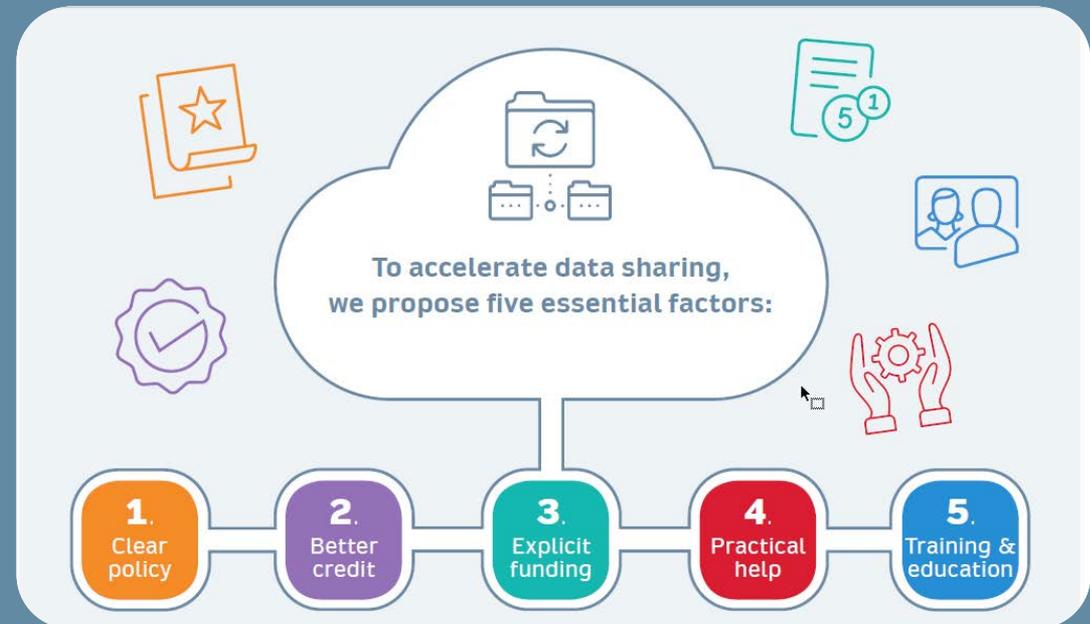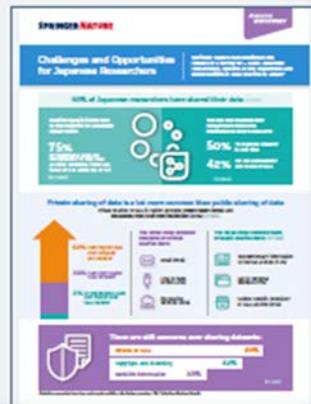
Published April 2019, our **5 essential factors** report brings together findings from surveys with more than 11,000 researchers worldwide.

SPRINGER NATURE

Many researchers are sharing data in some way

95% of Japanese researchers have shared their data (n=975)

Sharing data is important to the majority of Japanese researchers:

75% of respondents rated the discoverability of their data as being somewhat important (score of 6 or above out of 10)

(n=1062)
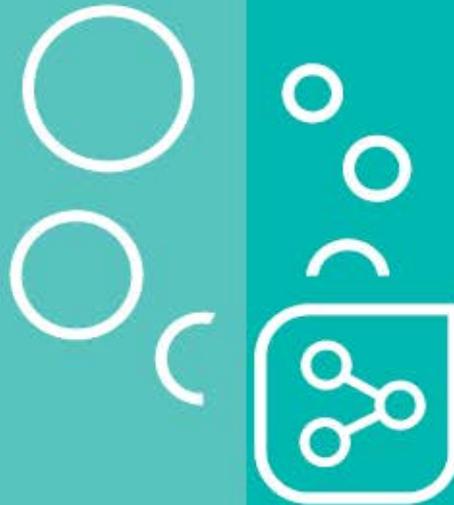
The top two reasons why researchers would be motivated to share data are:

50% 'To progress research' in their field

42% For the transparency and re-use of data

(n=994)

White paper published May 2019:
Challenges and opportunities for data sharing in Japan
Ludivine Allagnat, Katie Allin, Grace Baynes, Iain Hrynaszkiewicz, Mithu Lucraft
https://www.doi.org/10.6084/m9.figshare.7999451

SPRINGER NATURE

2019

But, the most common methods of data sharing are suboptimal despite the importance of finding data

How do you share your data privately (person to person)?

49% USB or flash drives

43% email

30% PC hard drive

China survey (n=1,441)

65% email

41% USB or flash drives

39% file sharing services

Japan survey (n=905)

Importance of data discoverability

76% of respondents highly rate the importance of their data being discoverable: **most popular ranking was 10/10**

Importance out of 10

7.7 South America
7.6 Asia
7.3 Europe
7.2 North America
6.9 Australasia

Importance out of 10

6.5 Other sciences
6.6 Physical sciences
**7.2** Medical sciences
7.7 Earth sciences
7.8 Biological sciences

SPRINGER NATURE

Clearer policies on data sharing are needed from all stakeholders

**1.
Clear
policy**

In 2018, the policies of more than 50 global funders mandated data sharing

85 (22%) funders mandate or encourage data sharing but over 300 have no stated policy yet.

Mandate
12%

Mandate: some data
types
1%

Mandate: some
projects
2%

No mandate
77%

Encouraged
8%

Forthcoming
1%
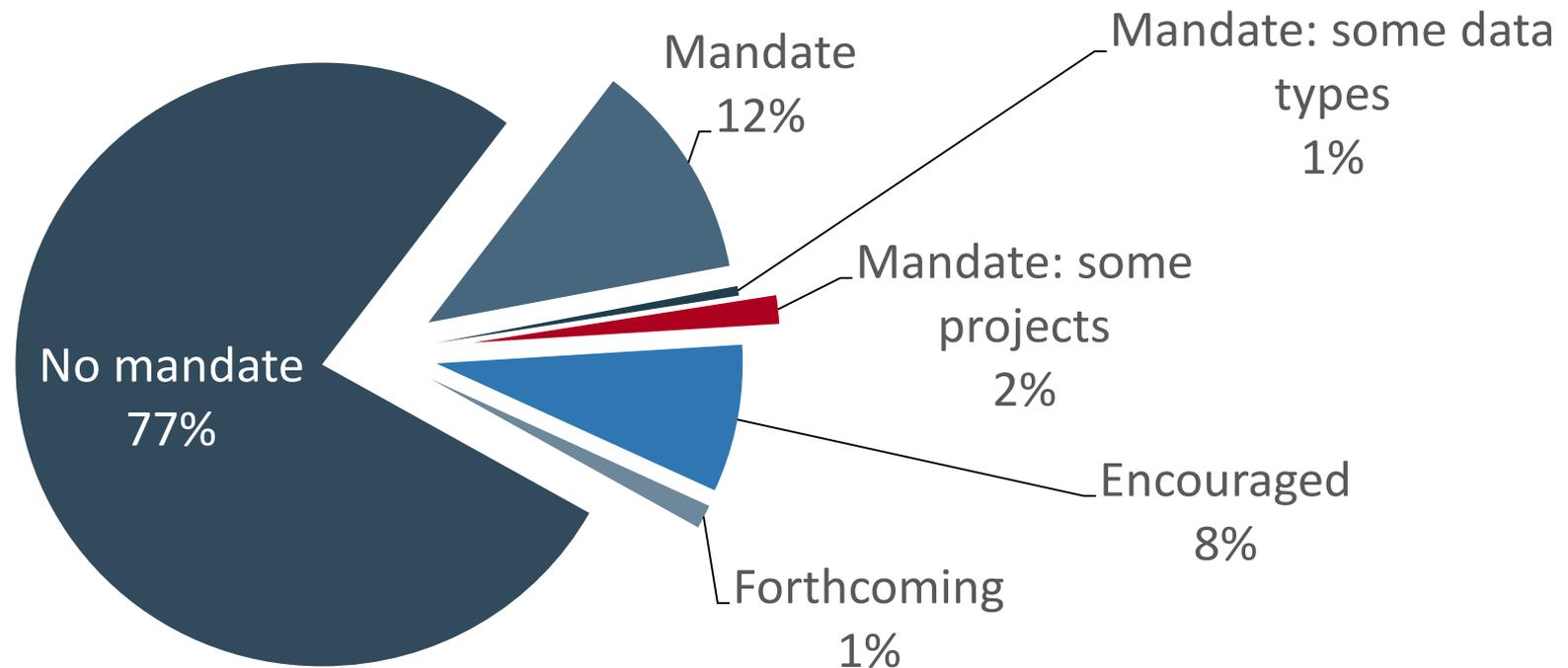
SPRINGER NATURE

Publisher policies on data sharing are becoming more consistent, since 2016 in particular

1. Clear policy

Policy Types

Type 1	Type 2	Type 3	Type 4
Data sharing and data citation is encouraged but not required	Data sharing and evidence of data sharing encouraged	Data sharing encouraged and statements of data availability required	Data sharing, evidence of data sharing and peer review of data required

- Springer Nature began a journal policy standardisation initiative in 2016[1]
- Now, more than 1,700 (~68%) Springer Nature journals have adopted a standard research data policy
- There are several policy options, allowing journals to select the type or strength of policy that best meets their needs
- All policies and journals promote data citation and data repositories
- Similar initiatives since introduced at Elsevier[2], Wiley[3], Taylor & Francis[4]

1. *Standardising and harmonising research data policy in scholarly publishing*
Iain Hrynaszkiewicz, Aliaksandr Birukou, Mathias Astell, Sowmya Swaminathan, Amye Kenall, Varsha Khodiyar
International Journal of Digital Curation; doi: https://doi.org/10.2218/ijdc.v12i1.531
2. https://www.elsevier.com/authors/author-services/research-data/data-guidelines
3. https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-open-access/open-access/data-sharing.html
4. https://authorservices.taylorandfrancis.com/understanding-our-data-sharing-policies/

SPRINGER NATURE

Researchers want to receive credit for sharing data, through citations and publications

2. Better credit

58% Think researchers do not get sufficient credit for sharing data

46% Would be motivated to share by "getting proper credit"

55% Would value a data citation as much as an article citation

60% Note that re-use that resulted in credit as a co-author would motivate them "quite a lot" or "a lot"

https://figshare.com/articles/Five_Essential_Factors_for_Data_Sharing/7807949

SPRINGER NATURE

Collaborations and initiatives to promote credit and recognition for data sharing

2. Better credit

1. **Data citation**
   1. Journal policies promote citation of datasets in reference lists and publishers are collaborating to implement better ways to measure and track reuse of data through citations

2. **More recognition of data sharing in funding and promotion decisions**
   1. Datasets and software as well as papers are recognised by, for example, US National Science Foundation

3. **Data publishing options in peer-reviewed journals**
   1. Data journals, such as *Scientific Data*, provide a means to gain publication credit specifically for publishing datasets

SCIENTIFIC DATA

SPRINGER NATURE

Explicit funding for data management and data sharing, as well as data publishing are needed

3. Explicit funding

Policy must be coupled with dedicated funding and clear guidance about using grants on research data management costs

| Costs of data sharing	Source of funds
• 27% did not know how they would meet costs of making data open	• 37% likely to use own funds • 41% would use funds identified in their grant

https://figshare.com/articles/Five_Essential_Factors_for_Data_Sharing/7807949

SPRINGER NATURE

Researchers have numerous practical challenges in sharing their datasets

4. Practical help

Research Data
PRACTICAL CHALLENGES FOR RESEARCHERS IN DATA SHARING
White paper

Costs of sharing data	Lack of time to deposit data	Not knowing which repository to use	Unsure about copyright and licensing	Organising data in a presentable and useful way
12%	16%	20%	23%	28%
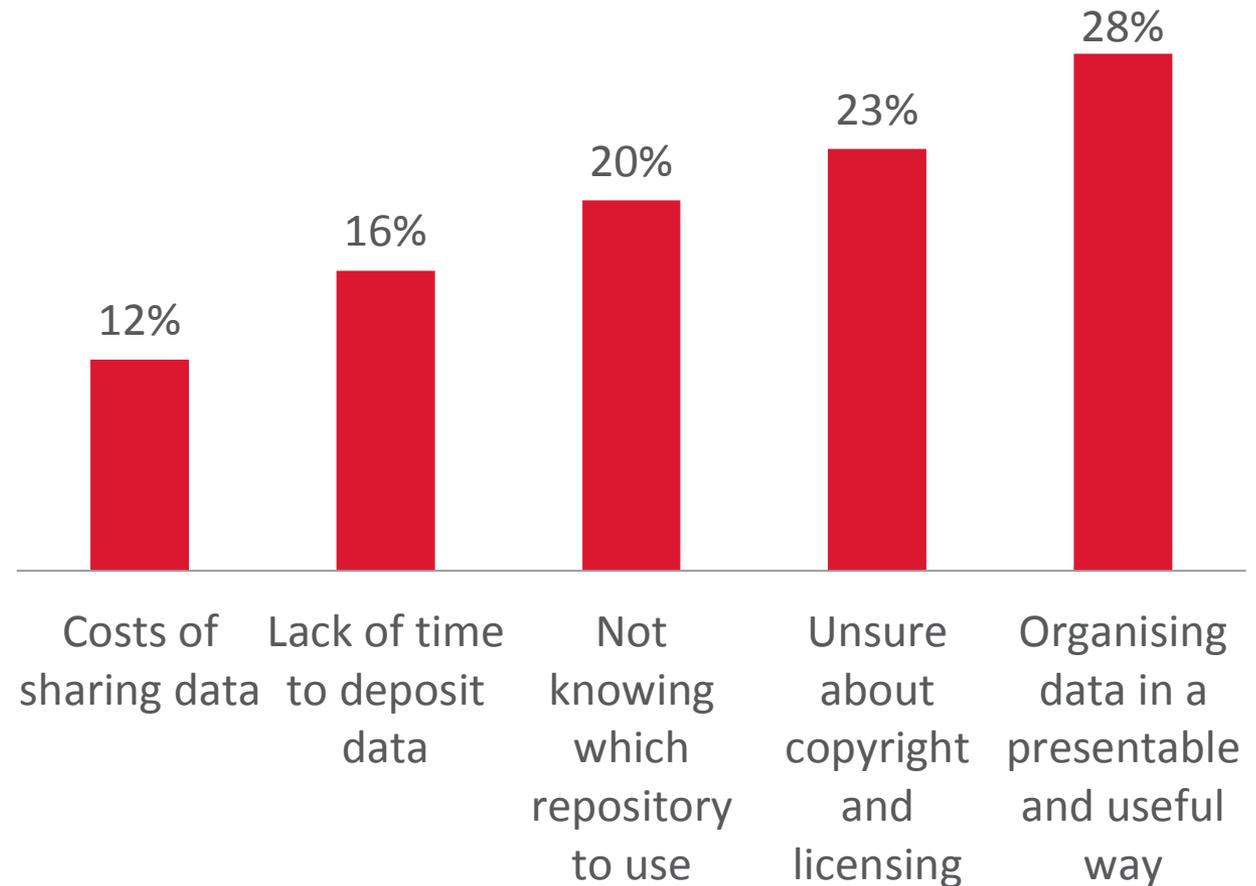
From a Springer Nature researcher survey. Total respondents: 7719

https://doi.org/10.6084/m9.figshare.5975011

SPRINGER NATURE

Practical help: organising data, finding appropriate repositories and faster easier routes for sharing

4. Practical help

- **Promoting data repositories provides infrastructure for data sharing and makes data more findable**
  - Need to increase awareness and use of repositories over supplementary information files
  - Recommended repository lists
- **Advice, resources and guidance to help researchers find the best solution for sharing and managing their data**
  - Springer Nature offers a free Research Data Helpdesk
  - Institutional libraries have an important role
  - Other third parties such as Digital Curation Centre
- **Optional services such as Springer Nature's Research Data Support**
  - Solution for researchers needing more "hands on" assistance to organise and curate their data and metadata

41%
researchers share their data via a repository

42%
as supplementary information at the point of submitting a publication

https://figshare.com/articles/Five_Essential_Factors_for_Data_Sharing/7807949

Training and education required to support best practice and answer questions

5. Training & education

Who is best placed to support data management?

57% peers

52% publishers

42% libraries

Copyright and licensing of data?

Which repository to use?

65% of researchers feel there is not sufficient training, support and advice in regard to data management

How to prepare a DMP?

How to share sensitive data?

How to comply with journal and funder policies?

How to guard against misuse of data?

Nature Research Academies – Research Data Principles and Practices

Available from Nature Research Academies: workshops in Research Data Principles and Practices. Nature Research Academies is a series of training workshops for researchers, developed by Nature Research.

Contact Us

5 essential factors JOSS May 2019

SPRINGER NATURE

https://figshare.com/articles/Five_Essential_Factors_for_Data_Sharing/7807949

Thank you

Iain Hrynaszkiewicz

iain.hrynaszkiewicz@nature.com

@iainh_z

For more information on research data-related activities at Springer Nature:

Email: researchdata@springernature.com

Website:
http://go.nature.com/ResearchDataServices

Slide acknowledgements:
Mithu Lucraft

The story behind the image

Chien Shiung Wu (1912–1997)

Chien Shiung Wu was a Chinese American experimental physicist best known for conducting The Wu experiment that bears her name. This experiment showed that the conservation of parity was violated by a weak interaction and it was possible to distinguish between a mirrored variation of the world and the mirror image of the current world. This discovery earned Wu the Wolf Prize in Physics in 1978.

SPRINGER NATURE