

# The Future of Data in Research Publishing: *From Nice to Have to Need to Have?*

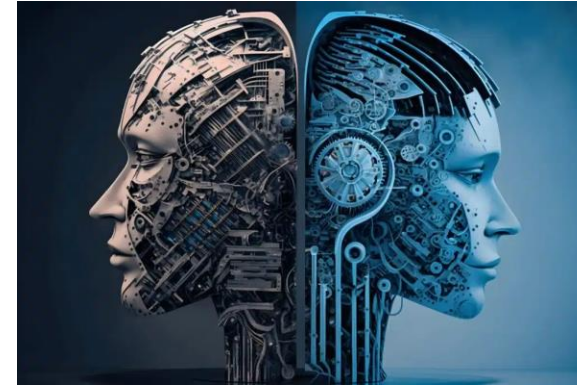
Borgman, C. L., & Brand, A. (2024). [The Future of Data in Research Publishing](#): From Nice to Have to Need to Have? *Harvard Data Science Review, Special Issue 4*. [Webinar May 22, 2024](#)

Christine L. Borgman

Distinguished Research Professor  
University of California, Los Angeles

& Amy Brand

Director and Publisher  
The MIT Press



[Image source](#)



[Image source](#)

# Open Research Data: Policy and Practice

## Why share research data?

- Transparency
- Reuse
- Reproduce
- Educate
- Required
  - Funding agencies
  - Journals



## How to share research data?

- Describe dataset; add metadata
- Deposit in data archive
- Publish data documentation
  - Research protocols
  - Codebooks
  - Software
  - Algorithms
- Link datasets to publication
- Cite data and software
- Develop instructional materials



Borgman, C. L., & Groth, P. T. (2024). *From Data Creator to Data Reuser: Distance Matters*. <https://doi.org/10.48550/arXiv.2402.07926>

# Open Data

- Transparency
  - Trust
  - Verification
  - Accountability
  - Reuse
  - Reproducibility
  - Replication

# Artificial Intelligence

- Opacity
  - Black box
  - Algorithmic
  - Generate text
  - Answer questions
  - Minimal domain expertise
  - Hallucinatory responses
    - Non-existent facts
    - Non-existent publications

# Open Data: Researcher/Publisher Perspectives

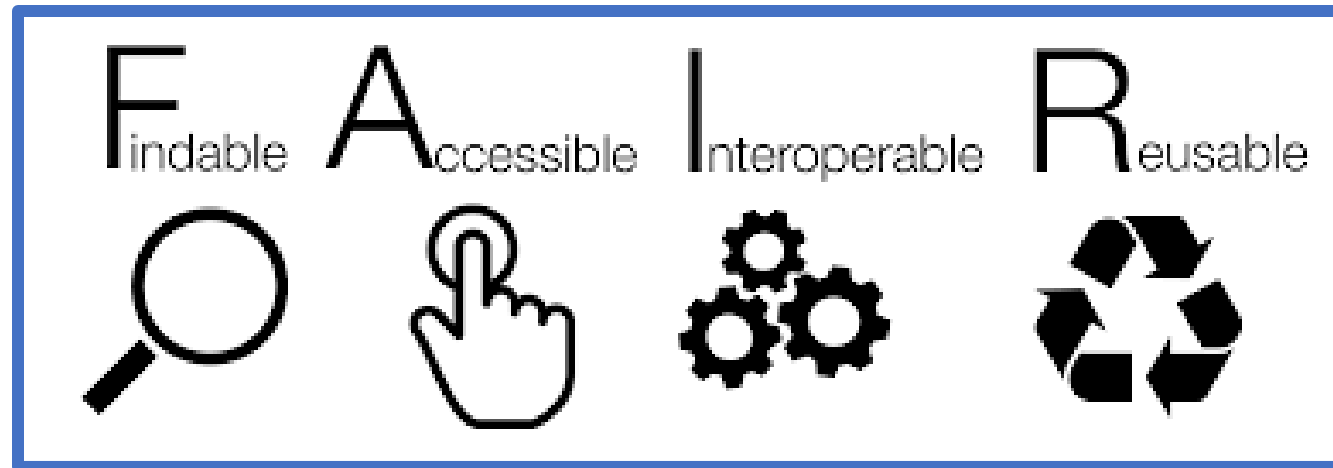
- Benefits

- Recognition
- Citation
- Collaboration
- Transparency
- Evidence
- Trust
- Reuse
- Reproducibility
- Replication

- Risks and costs

- Labor skills, time
- Lack of infrastructure
- Benefits to others
- Intellectual property control
- Reuse without attribution
- Fake data sources
- Deliberate misuses of data
  - Trolls
  - Climate skeptics
  - Vaccine skeptics...

# Data Sharing and Stewardship: The Ideal



Wilkinson, et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. [Scientific Data](#), 3,

# Data Stewardship: The Reality



We just need to migrate the data from these systems to fit into that hole over there.

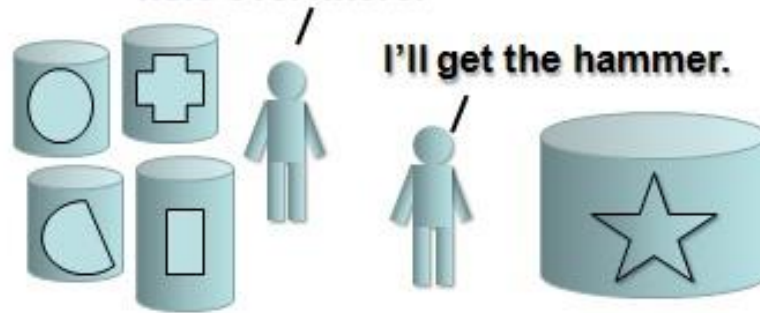


Image source



Graduate students



Post-doctoral fellows <sup>6</sup>

# Governance: Building the Village

- Data sharing is a ‘collective action problem’
- Holistic approaches to sharing infrastructure
  - Distribute responsibility among stakeholders
  - Invest in data management expertise
- Fund the commons
  - Public support for data repositories
  - International exchange of best practices
- Invest in sustainable strategies for data
- Address benefits and risks of open data

