



Open Data @PLOS

Emma Ganley

Chief Editor, PLOS Biology

Lead for PLOS Data Program

Public Library of Science

Belmont Forum on Data Management | Jun 23rd 2017



@GanleyEmma

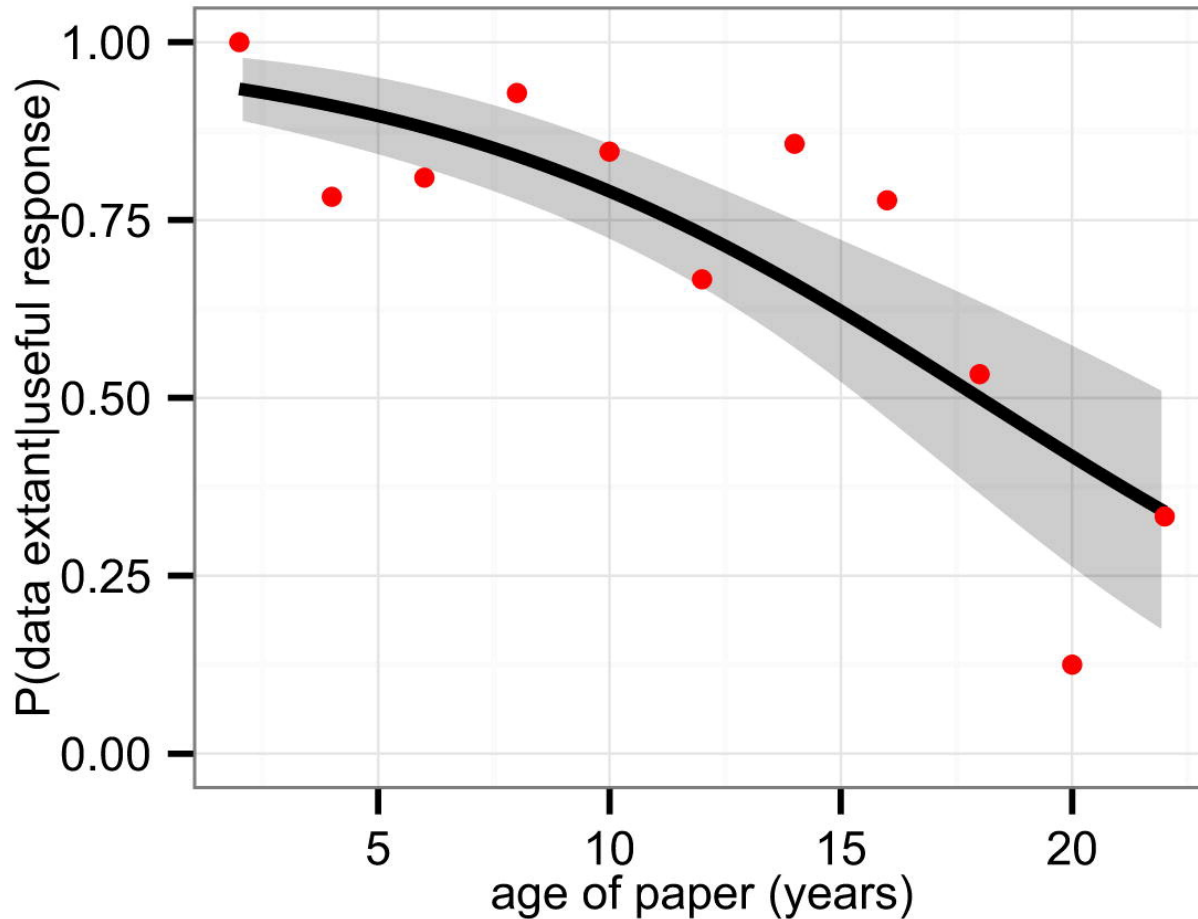
Open Access allows everyone to access research. Open Science and Open Data allow everyone to do research.

What Is the Value of Research Data?

- Replication/Validation
- New analysis
- Better interpretation
- Inclusion in meta studies
- Facilitate reproducibility
- Scrutiny post-publication
- Better financial and intellectual return on research investment

Data Availability Declines Over Time

Image: Nature doi:10.1038/nature.2013.14416



Probability of finding the data associated with a paper declined by 17% every year

Vines, Timothy et al. "The Availability of Research Data Declines Rapidly with Article Age." *Current Biology* 24, no. 1 (June 1, 2014): 94–97. doi:10.1016/j.cub.2013.11.014.



ALMOST ALL DATA LOST 10-15 YRS AFTER PUBLICATION

PLOS Data Policy

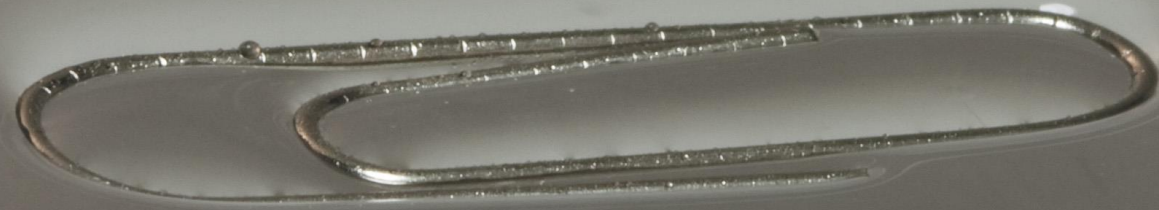
PLOS journals require authors to make **all data underlying the findings** described in their manuscript fully available without restriction, with rare exception.

When submitting a manuscript online, authors must provide a **Data Availability Statement** describing compliance with PLOS's policy.

Exceptions

- **Ethical or legal reasons**, e.g., public availability would compromise patient confidentiality or participant privacy.
- Data deposition could present some **other threat**, such as revealing the locations of fossil deposits, endangered species, or farms/other animal enclosures etc.

Natural Tension & Resistance



>65,000

papers published with a data
statement at PLOS

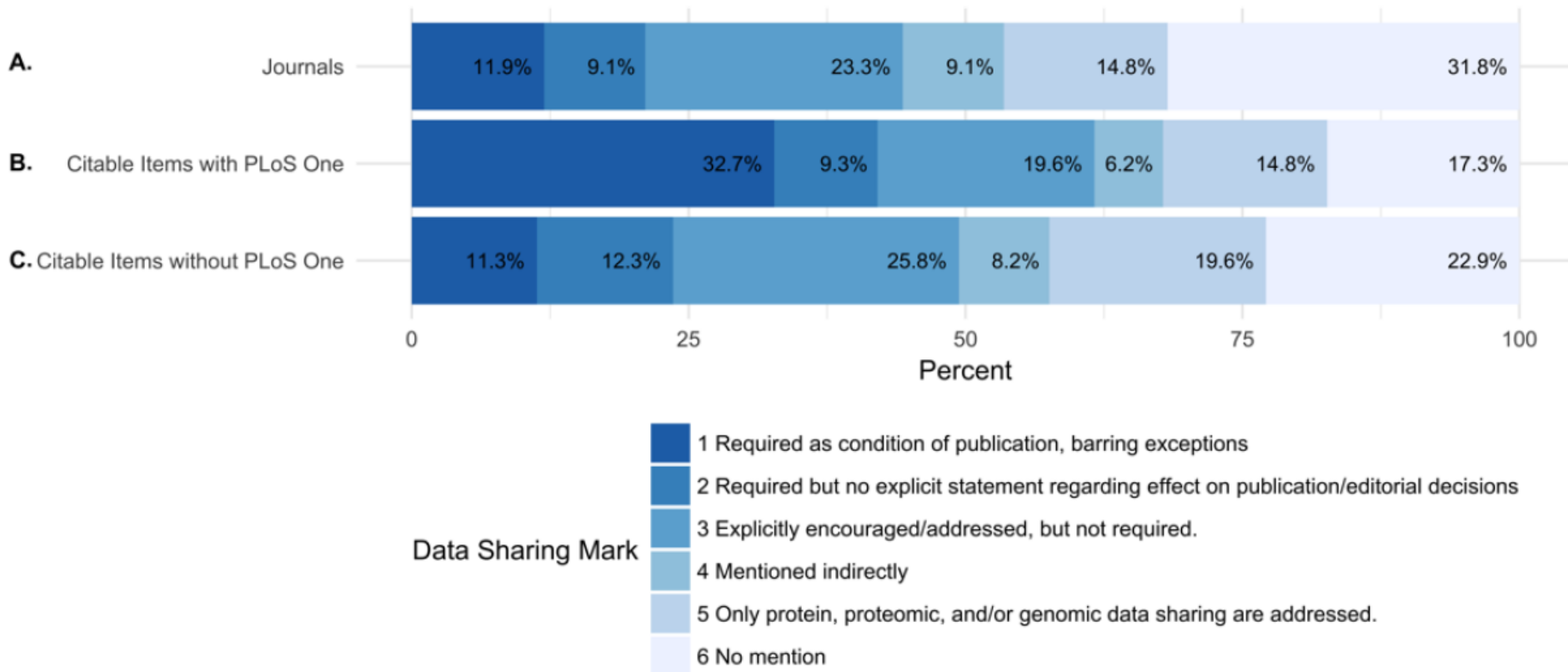
<0.1%

of submissions rejected due
to authors' unwillingness or
inability to share data

~20%

of submissions use data
repositories

Data sharing policies make a difference



Challenges

- Research areas such as clinical studies require more complex data sharing considerations and data release mechanisms. **Community input** is important for policy implementation
- Data citations and mechanisms to provide **author credit** need to see a stronger uptake
- **Metadata** of published data sets is often lacking, needs community-agreed standards
- It is not always clear what constitutes **compliance**



Thank you!