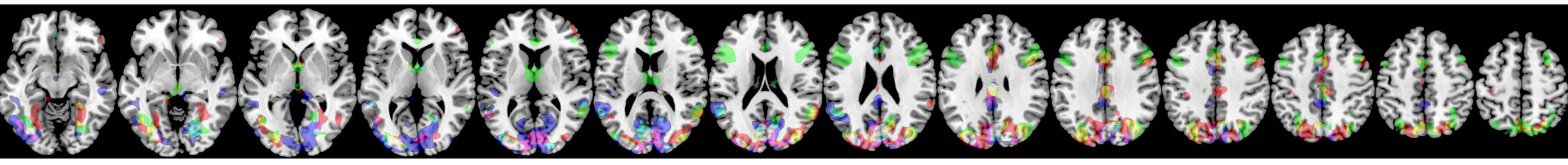
Reproducibility Begins at Home: Surveying Data Management Practices in Neuroimaging

John Borghi, California Digital Library
Ana Van Gulick, Carnegie Mellon University



Background

Because of the cost and complexity of the materials involved, effective research data management (RDM) is essential in human brain imaging (neuroimaging) research. Open science practices, including the regular sharing of data, code, and other materials have also been widely discussed as the field grapples with questions related to the rigor and reproducibility of its methods. However, comprehensive information about how active researchers manage and share their data over the course of a neuroimaging research project has - until recently - remained largely anecdotal.

Methods

In order to understand the RDM-related practices and perceptions of neuroimaging researchers, we designed and disseminated a survey structured around the progression of a research project involving magnetic resonance imaging (MRI). Questions addressed how data is saved, organized, and shared, the use of software tools, the perceived maturity of RDM-related practices, and the adoption of emerging scholarly communication practices. A total of 144 active neuroimaging researchers completed the survey, which was open June through September, 2017.

Results

A total of 144 neuroimaging researchers participated in our survey. Participants were a mix of trainees (graduate students and post-docs) and faculty and were affiliated with a range of academic disciplines, with the most common being cognitive neuroscience.

Open Materials

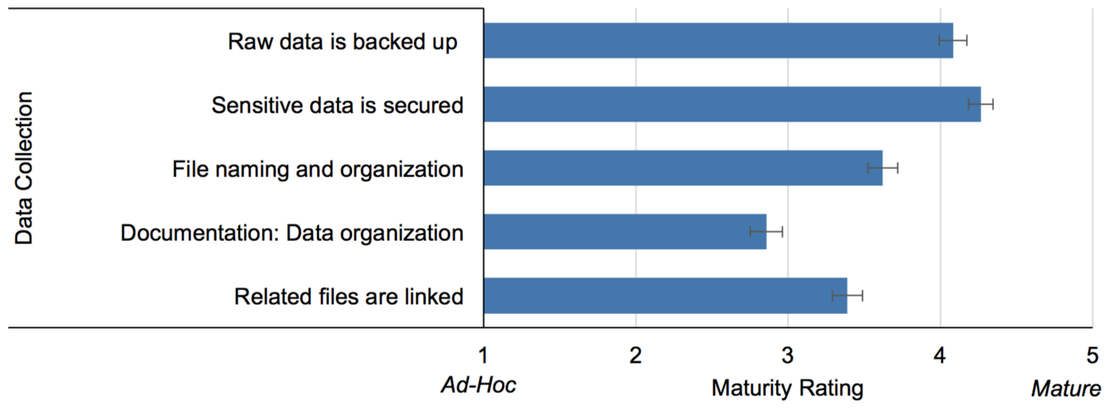
Our survey and dataset are openly available via figshare. For additional analysis and discussion, see our bioRxiv preprint.

Survey **Data** **Preprint**

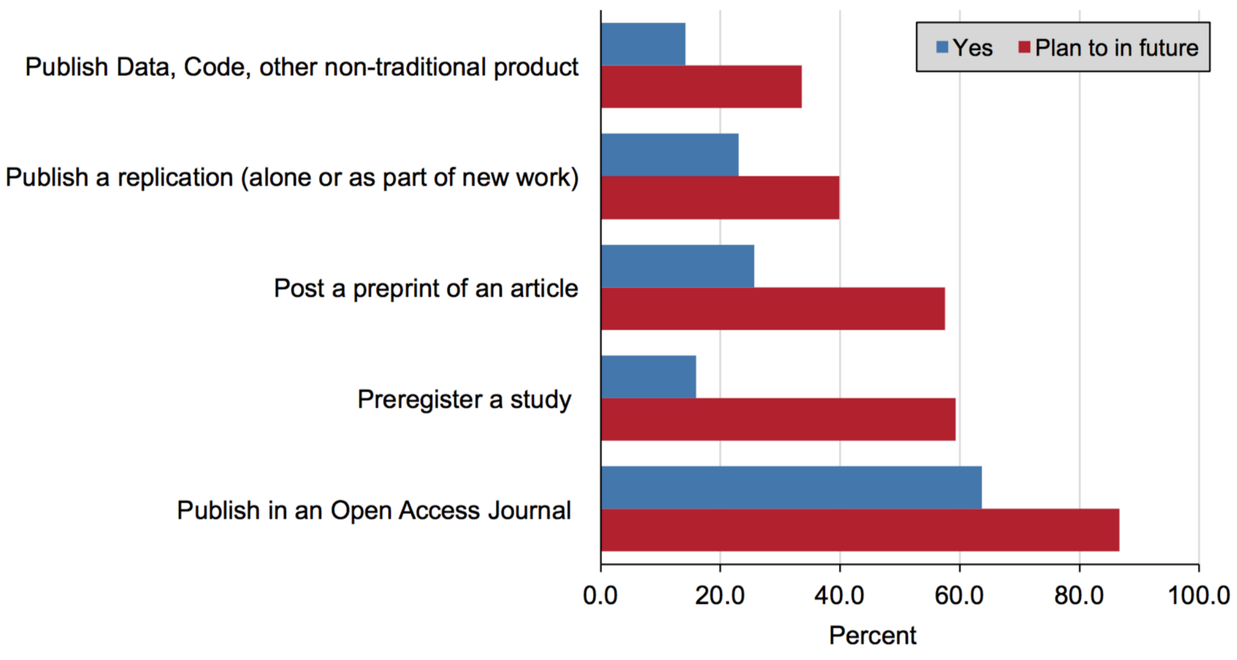
RDM limits and motivations

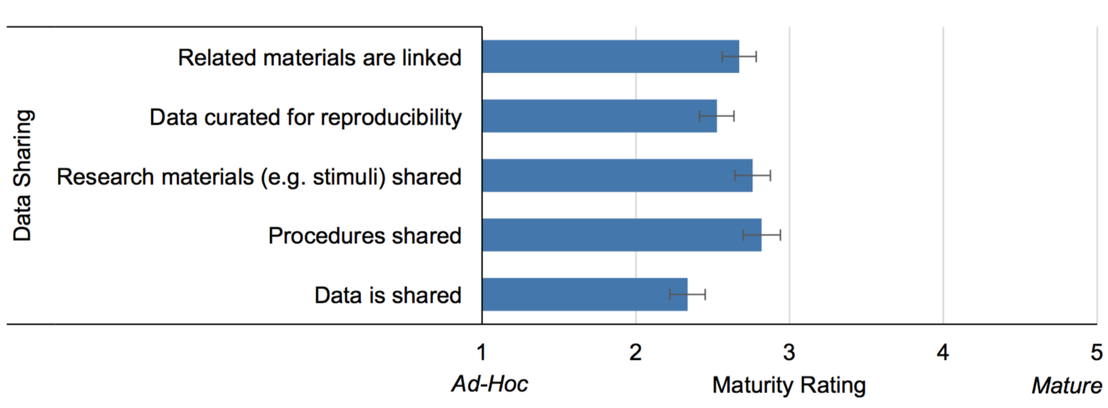
		Data Collection	Data Analysis	Data Sharing
Limits	The amount of time it takes	69.60%	71.30%	79.46%
	Lack of best practices	43.20%	48.70%	49.11%
	Lack of incentives	36.80%	32.18%	37.50%
	Lack of knowledge	32.80%	40.87%	41.07%
	The financial cost	17.60%	8.70%	22.32%
	Other	7.20%	6.09%	5.36%
Motivations	Prevent loss of data	100.00%	85.83%	78.57%
	Ensure access for collaborators	76.80%	73.33%	70.53%
	Openness and reproducibility	63.20%	64.17%	66.96%
	Institutional data policy	52.00%	39.17%	47.32%
	Publisher/funder mandates	35.20%	28.33%	41.96%
	Availability of tools	12.00%	9.17%	8.93%
	Other	3.20%	3.30%	0.00%

Self Report RDM Maturity Ratings

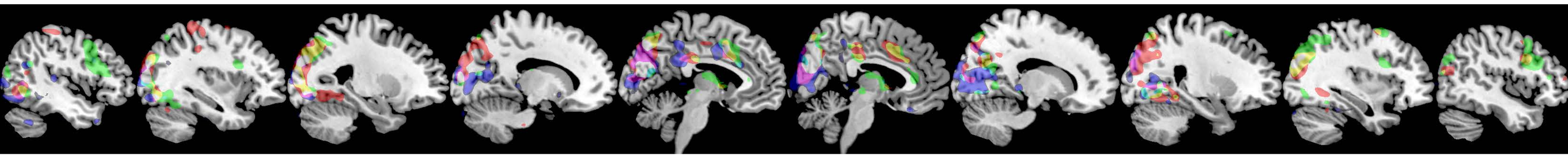


Emerging Practices





About the brain images: The two sets of brain slices on this poster were created using open materials from Neurovault (https://neurovault.org). Illustrating the importance of RDM in brain imaging, the data underlying these images can not be interpreted, reproduced, or re-used without access to extensive documentation related to the characteristics of study participants and image acquisition, processing, and analysis parameters.



orcid.org/0000-0001-9570-4163
orcid.org/0000-0002-8133-6395

Sponsored in part through grants from the Alfred P. Sloan Foundation #G-2014-13746 and from the National Science Foundation NSF ACI #1349002.

University of California CDL California Digital Library · UC3 UC Curation Center · Carnegie Mellon University · RDA Research Data Alliance