

GAGE

National Science Foundation's
Geodetic Facility for the
Advancement of Geoscience

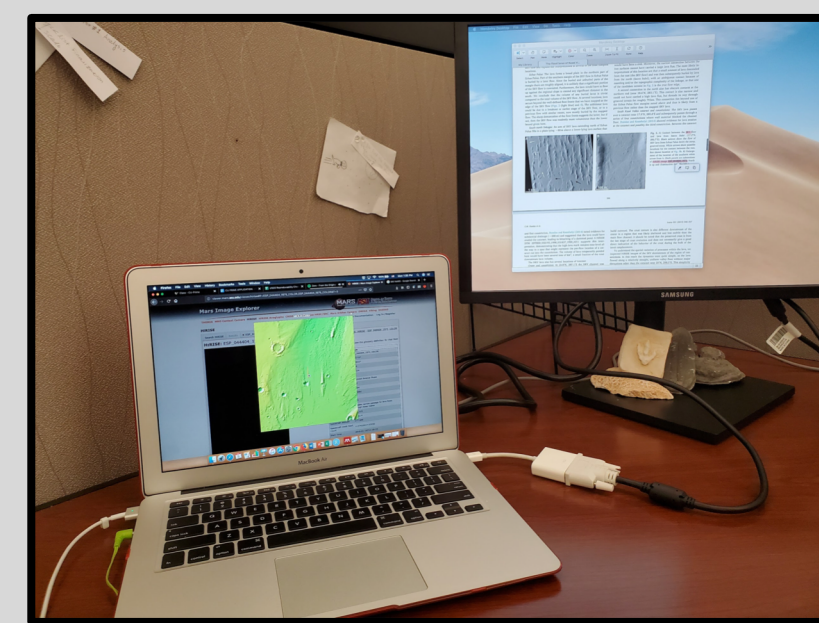
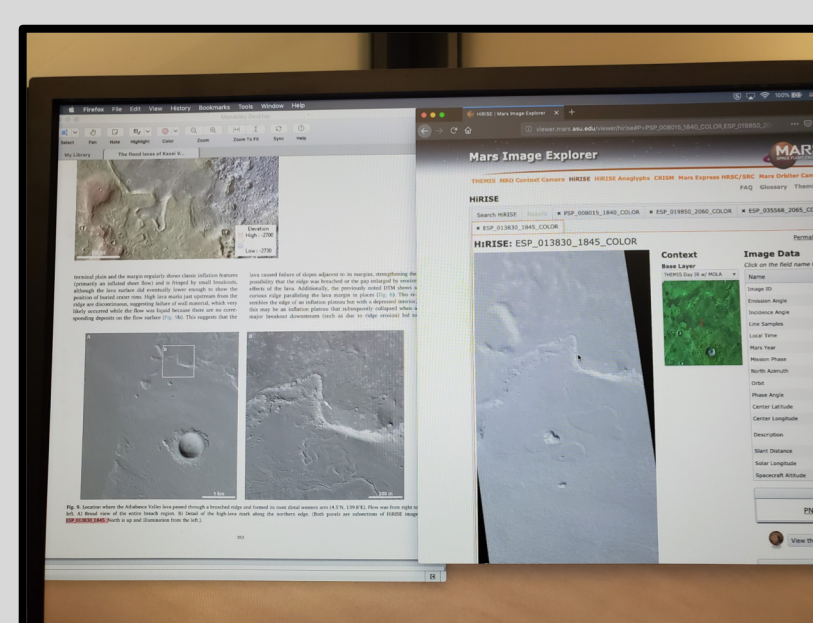
The Future of Published Papers: Improving Reproducibility Practices for USGS Publications

Alexandra M. Cohen^{1,2} and Shannon Q. O'Neill^{1,3}

UNAVCO

Background

Ideally, all published scientific works should be repeatable by another scientist reading the publication. Unfortunately, due largely to factors such as unfindable or unusable data, software, and samples, this is not always the case. As technology advances and scientists continue to tackle new questions, methods for maintaining reproducible publications must evolve to establish a basis for a more ideal and efficient scientific process.

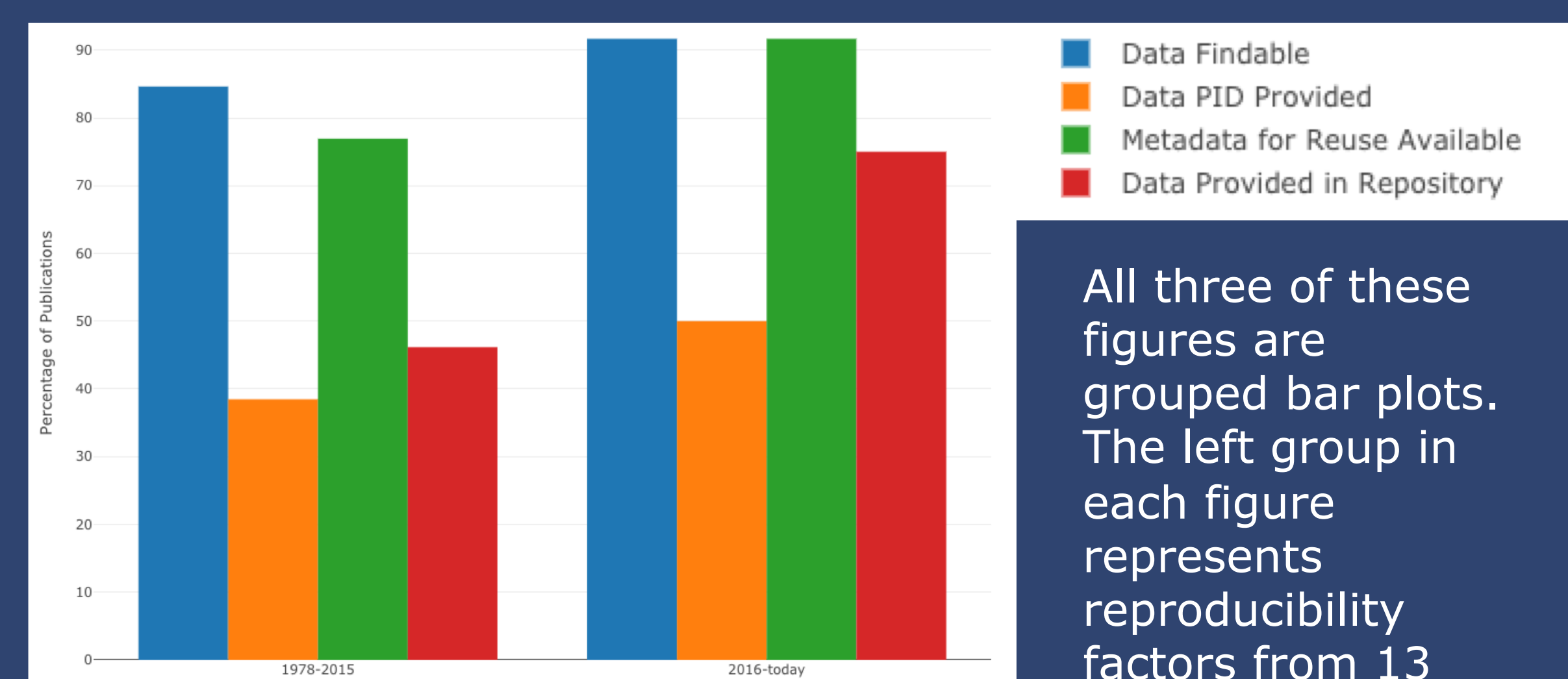


Study Focus

This project focuses on 25 publications associated with the USGS from the years of 1978-2019. Disciplines represented here include ecology, astrogeology, and geology. In 2016, the USGS enacted several policies aimed at increasing the reproducibility of their science. Through this project, we aim to identify where these policies have aided in the overall reproducibility of USGS research, and where improvement is still needed.

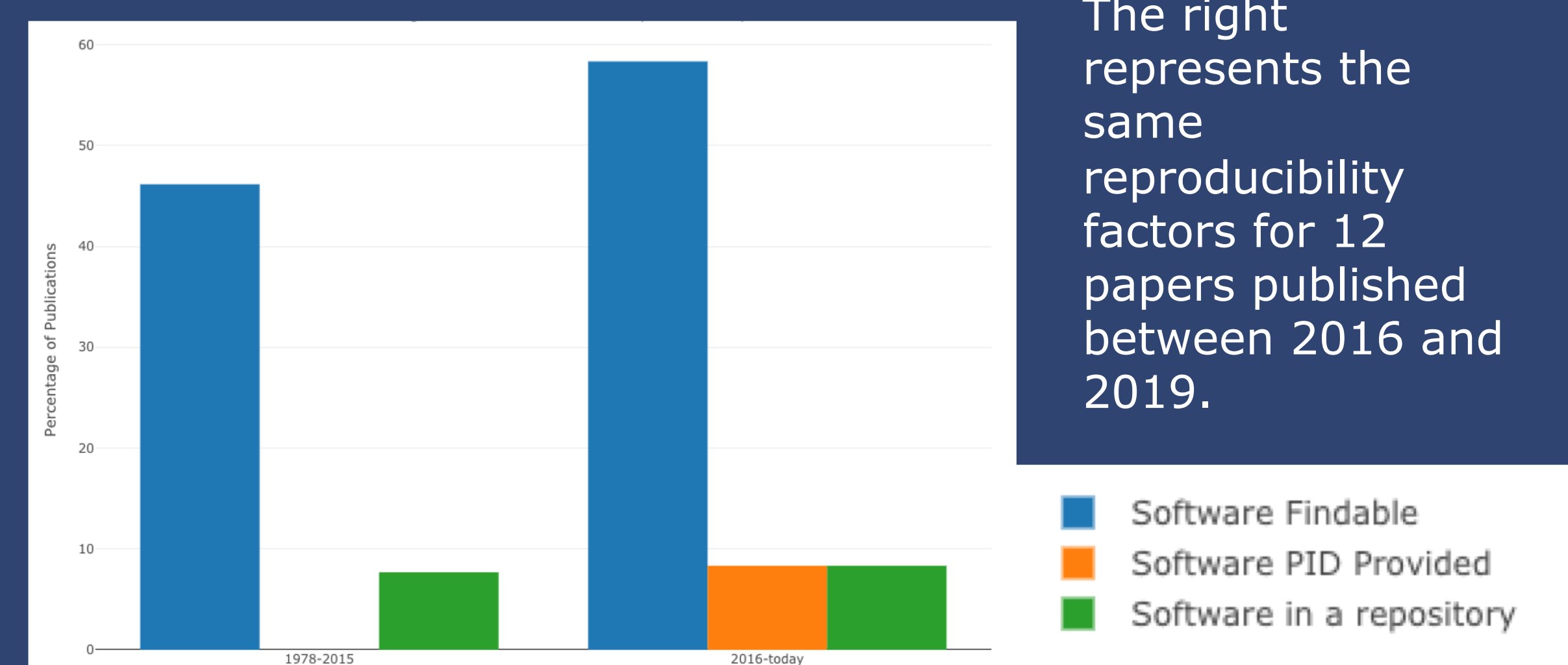
Analysis & Outcomes

Data Factors

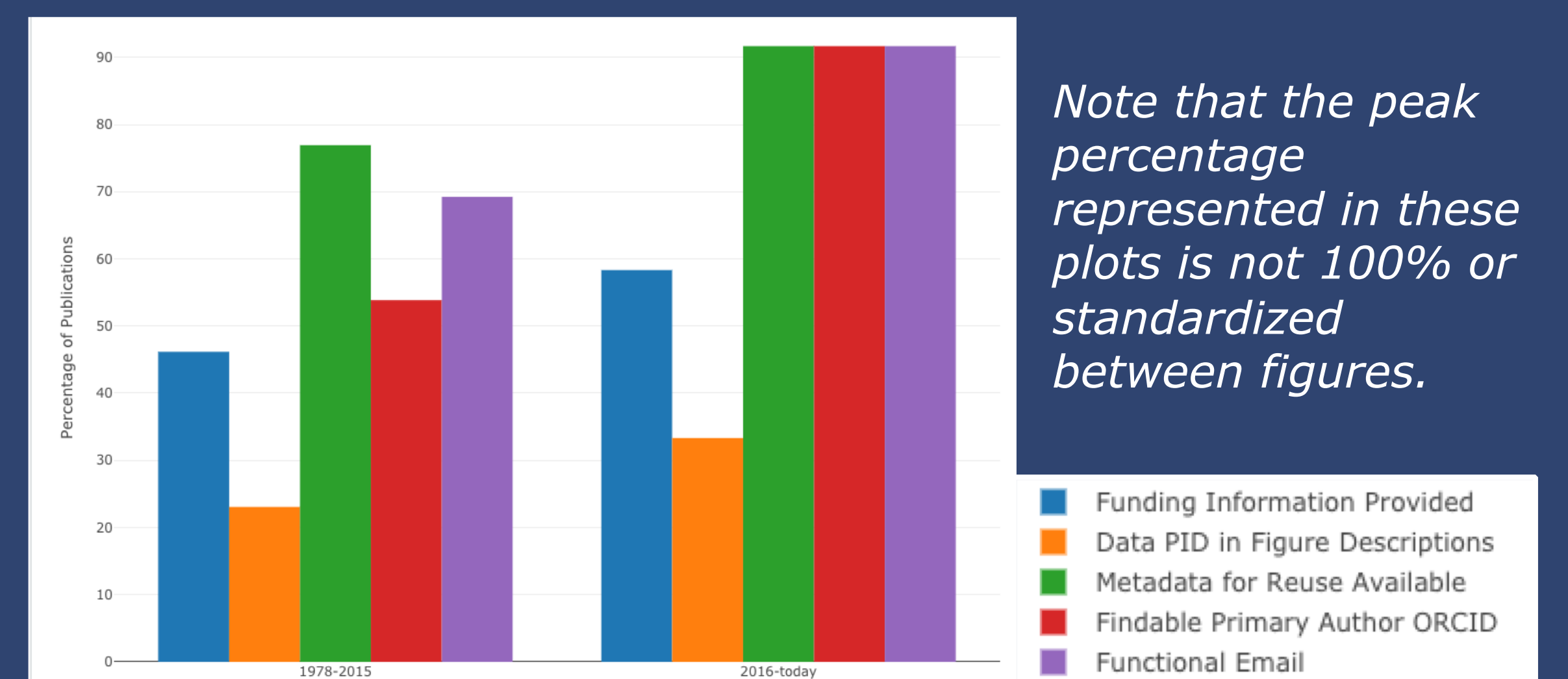


All three of these figures are grouped bar plots. The left group in each figure represents reproducibility factors from 13 papers published between the years of 1978 and 2015. The right represents the same reproducibility factors for 12 papers published between 2016 and 2019.

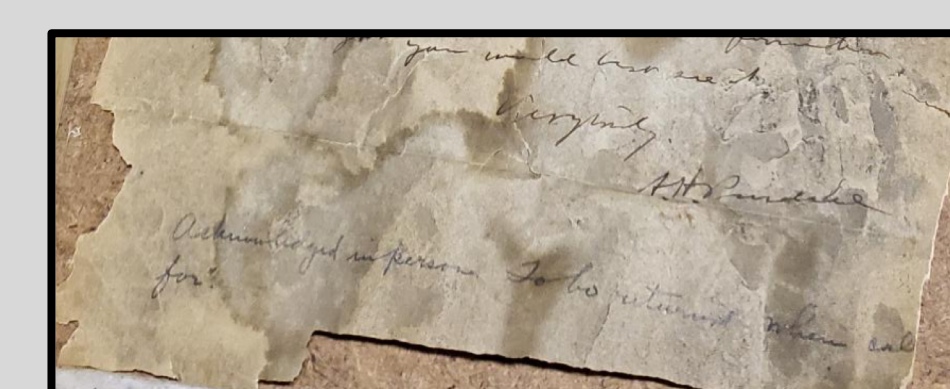
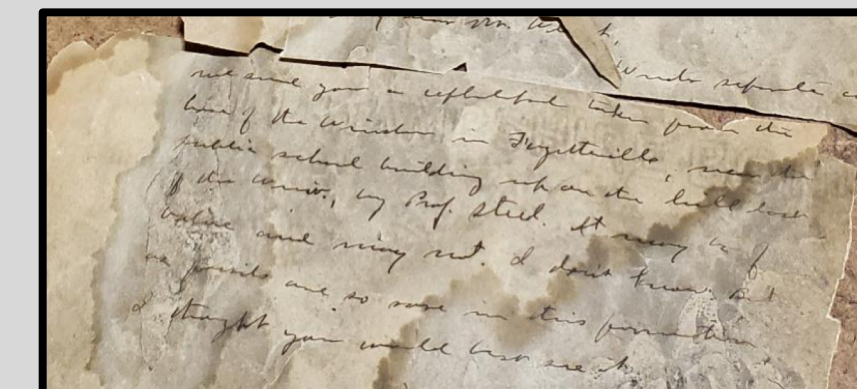
Software Factors



Other Factors



Note that the peak percentage represented in these plots is not 100% or standardized between figures.



Improvements In Progress

- Author Contact & ORCIDs
- Transparent Funding Info
- Findable Data & Software
- Extensive & Reusable Metadata

Improvements Still Needed

- Data Persistent Identifiers (PIDs) Provided
- Data Published in Repositories
- Software PIDs Provided
- Software Published in Repositories
- Relevant Data PIDs in Figure Descriptions
- Relevant Computational Provenance in Figure Descriptions

Methods Roadmap

We chose and analyzed 25 publications associated with USGS, based on publication year (pre-2016/post-2016) and discipline (Ecology/Astrogeology/Geology).



We created a rubric to assist in the determination of factors which contribute to the reproducibility of each paper's results. 18 of these factors were quantifiable and comparable.

PUBS WAREHOUSE		
ed		
USGS Offices		
ed on pubs warehouse landing page?		
Offices Listed on pubs warehouse?		
findable?		
u look?		
hey found?		
It take to find the pathway (MIN)?		
ave a license?		
here is where the data been provided in the publication?		

We compiled data for all 25 publications into a machine readable format. We used R-Studio Plotly package to create visual comparisons of reproducibility factors before and after 2016.

Attribute	Definition	Code
paper_id	Unique identifier for papers included in research	AGL, ECL, AG2, GEI
pub_year	Year in which the manuscript was published	YYYY
discipline	General discipline of the science described in the paper	AG-ASTROGEOLOGY, EC-ECOLOGY, GE-GEOLOGY
data_findable	Were data findable based on information provided in the publication?	0-No 1-Yes
data_repo	Were data provided in a repository	0-No 1-Yes
data_suppl	Were data supplement/appendix?	0-No 1-Yes
data_pid	Persistent identifier for data listed in the paper	0-No 1-Yes
data_metadata	Was enough metadata provided to reuse the data?	0-No 1-Yes
software_findable	Non-commercial software is findable online based on information provided in the paper	0-No 1-Yes

Acknowledgements

- UNAVCO Geo-Launchpad Intern Program
- Red Rocks Community College, Lakewood, Colorado 80228
<https://orcid.org/0000-0002-7798-6395>
- University of Colorado-Boulder, Boulder, Colorado
<https://orcid.org/0000-0002-4537-567X>

We would like to extend our gratitude to staff at USGS for their support and guidance throughout the project. Special thanks to our mentors at USGS; Madison Langseth, Vivian Hutchison, and Lindsay Powers for their invaluable aid. We would also like to thank the entire UNAVCO staff for creating and maintaining such an incredible learning and growth opportunity throughout this internship.

References

For a complete works cited, or for more information about this project, please visit
DOI: 10.6084/m9.figshare.8947310



This material is based upon work supported by the National Science Foundation under Grant No. 1540524 and Grant No. 1724794. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

