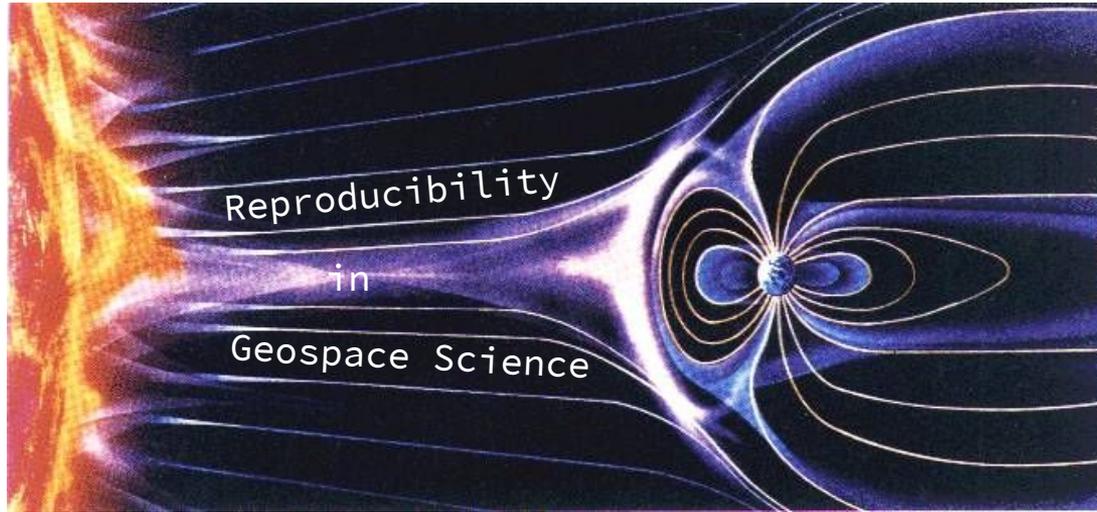


REPRODUCIBILITY OF RESEARCH RESULTS

[IN GEOSPACE SCIENCE]

Asti Bhatt, Todd Valentic, Ashton Reimer,
Leslie Lamarche & Pablo Reyes
SRI International

[FORMAT OF THIS WEBINAR]



Introduction to reproducibility

Expanding on 'Computational Reproducibility'

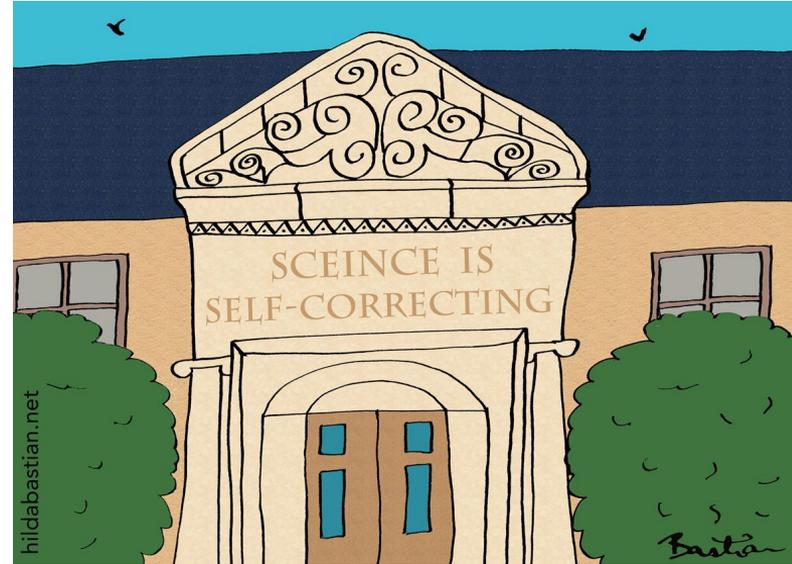
Some best practices towards creating reproducible results

Open discussion forum on challenges for geospace science

WHAT IS REPRODUCIBILITY?

One of the core tenets of scientific process is that other scientists are able to repeat your method and either confirm or refute your results.

This is known as reproducibility.



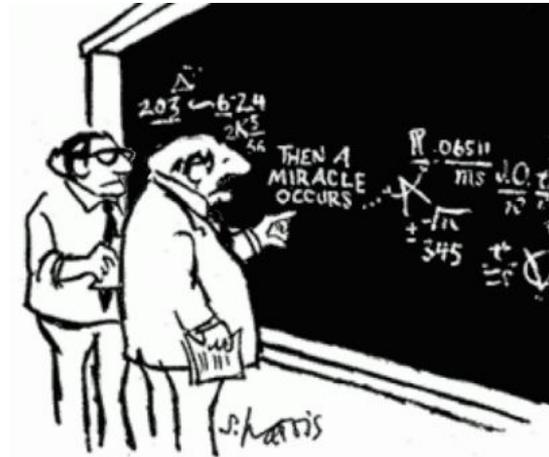
SURE, BUT WHAT DOES IT *REALLY* MEAN?

The definition of reproducibility varies:
Empirical, Statistical, Computational

Reproduction can be: Analytical, Direct,
Systematic, Conceptual

So, reproducibility means different things
to different people: Repeatability,
Robustness, Reliability, Generalizability.

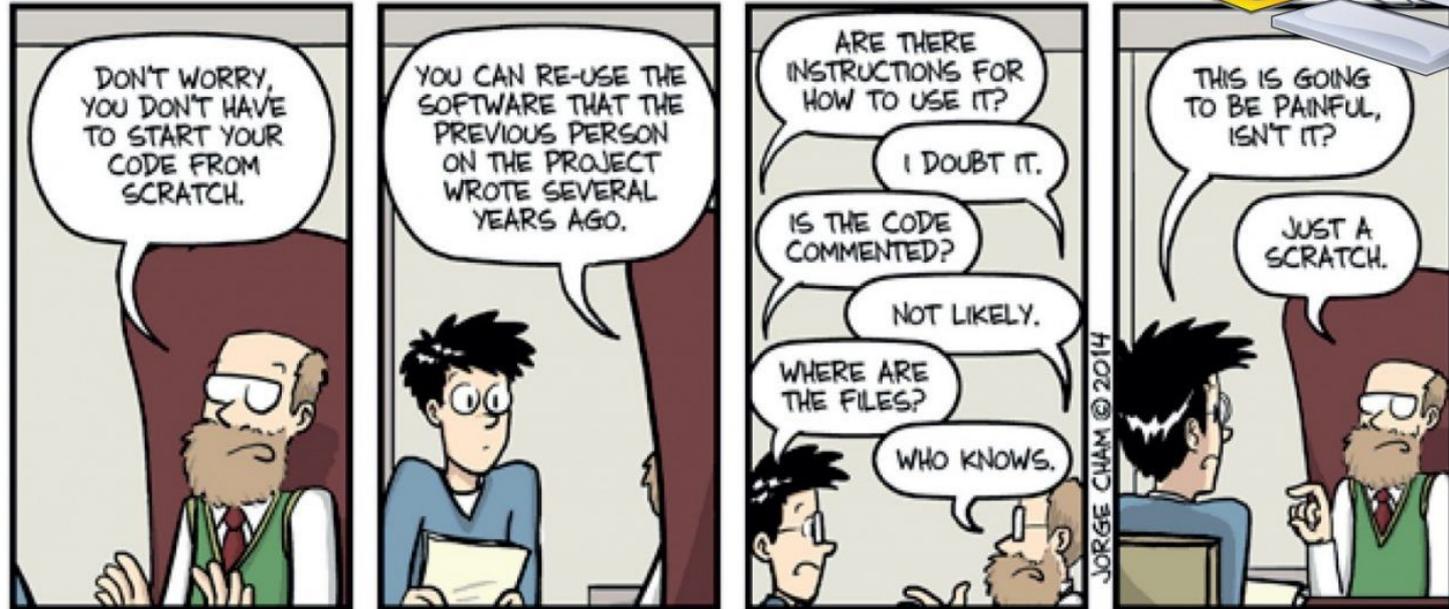
[What does reproducibility mean to you?]



"I THINK YOU SHOULD BE MORE
EXPLICIT HERE IN STEP TWO."

<https://scientificbsides.wordpress.com/2015/07/15/five-selfish-reasons-for-working-reproducibly/>

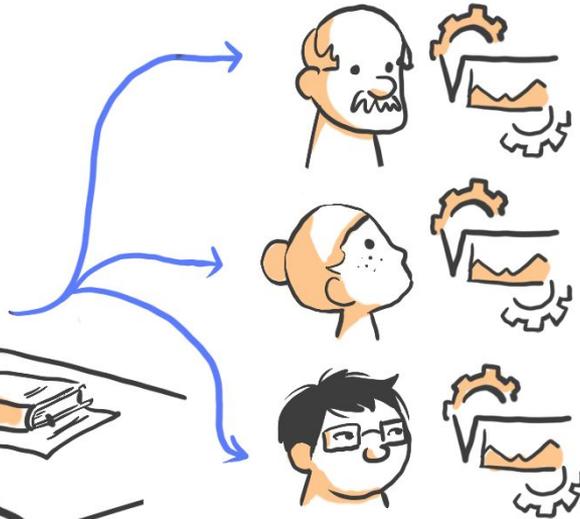
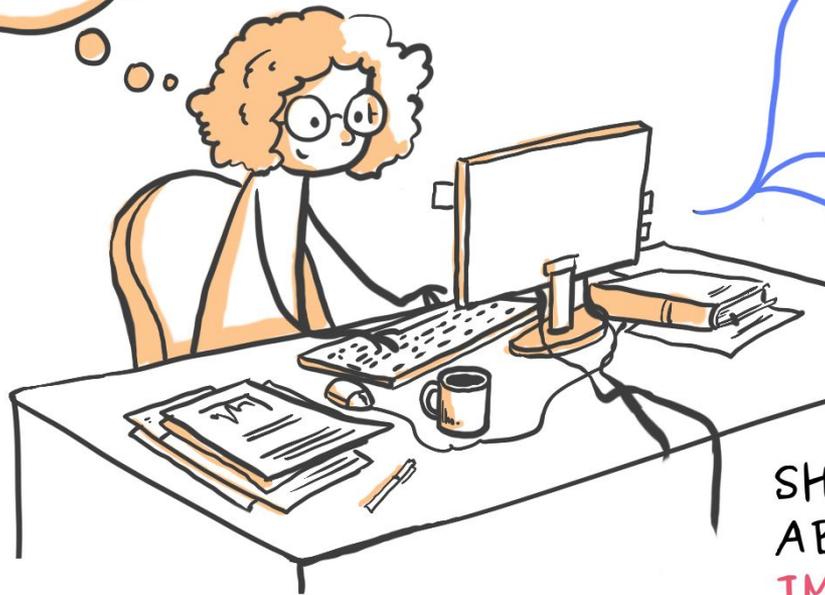
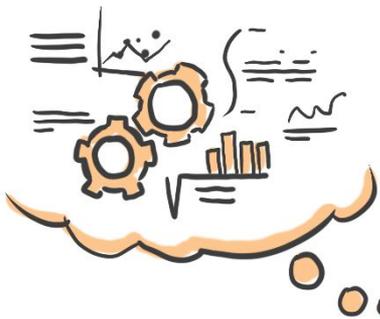
OK, BUT HOW IMPORTANT IS IT?



JORGE CHAM © 2014

'COMPUTATIONAL REPRODUCIBILITY'

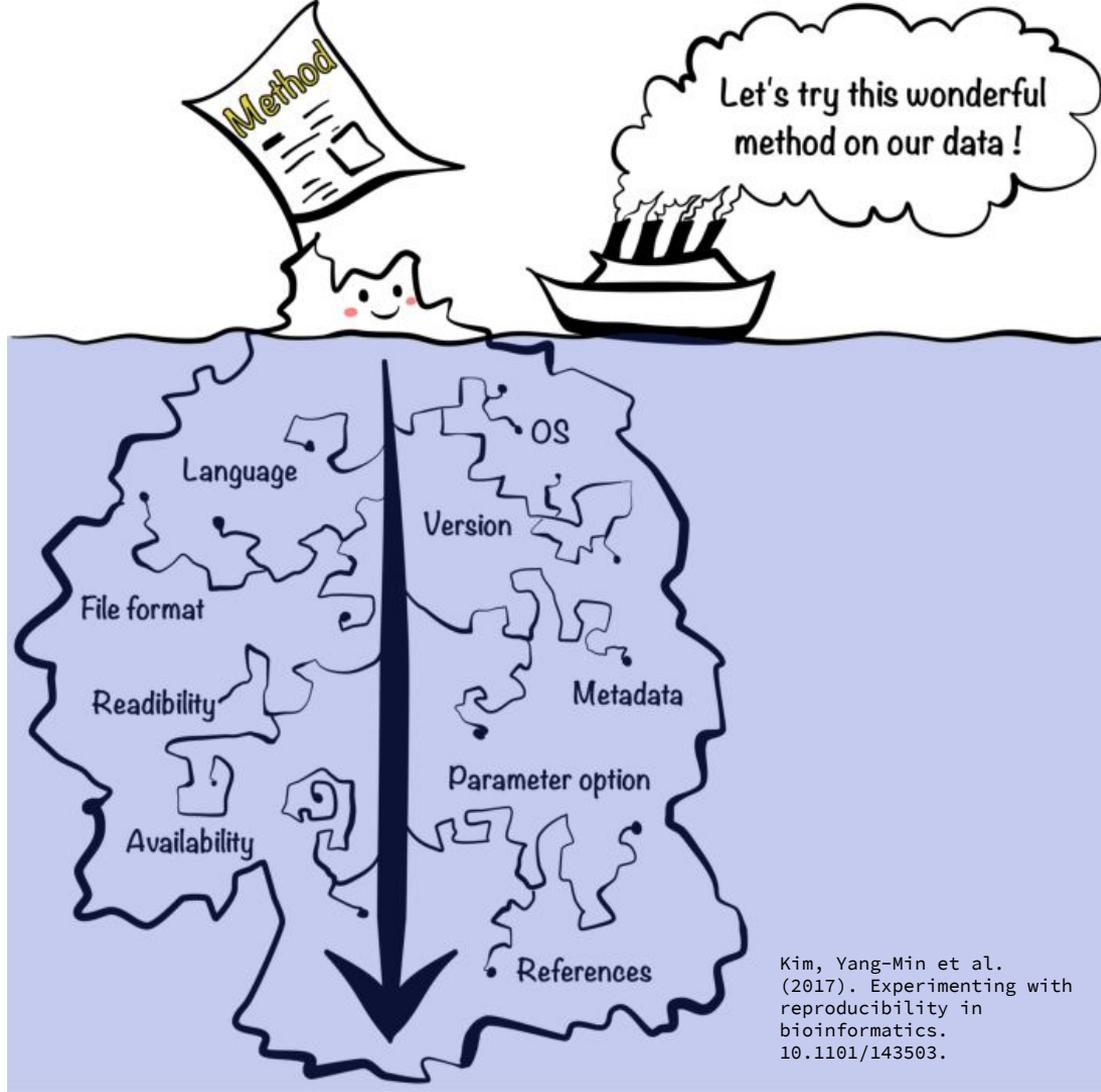
JANE HAS WRITTEN A PAPER
BASED ON HER EXPERIMENTS.



SHE WOULD LIKE ANYONE TO BE
ABLE TO REPRODUCE, CHECK, AND
IMPROVE HER CALCULATIONS

WHAT SHOULD JANE DO?

In the digital era, computational reproducibility is not only achievable, but is significantly more desirable.



Kim, Yang-Min et al.
(2017). Experimenting with
reproducibility in
bioinformatics.
10.1101/143503.

SOME COMMUNITY-ADOPTED BEST PRACTICES

Data from various instruments available in community data repositories, which offer -

- Persistent citations for data
- Version control for data
- Standardized metadata
- Standardized data formats



[What challenges have you encountered with accessing and acknowledging data?]

SOME INDIVIDUALLY-ADOPTED BEST PRACTICES

Ensuring data citations

Using Open-source tools

Software sharing with citations

Writing and publishing self-contained documents (e.g. notebooks)

Version control and licensing of code

Containerizing computational environment (e.g., Resen)

Reproducibility Spectrum



Non-reproducible

Gold standard

REPRODUCIBILITY ROADBLOCKS

Proprietary software (e.g. Matlab, IDL)

License managers

Hardware dependencies (e.g. GPUs)

Obsolescence (e.g. operating system)

Network APIs

Large data sets

Organizational restrictions (e.g. IP rights)



CHALLENGES IN GEOSPACE SCIENCES

What does reproducibility mean in geospace sciences?

What challenges have you faced in trying to create reproducible results?

What specific practices do you use when journals ask for source material?

What challenges have you faced as a data provider or a modeler?

How has FAIR data policy affected the way you publish?