

REPRODUCIBILITY CRISIS, OPEN SCIENCE, AND SOFTWARE FACTORIES

Arnaud Legrand



ACKNOWLEDGMENTS

Konrad Hinsen, Christophe Pouzat, Alexandre Hocquet

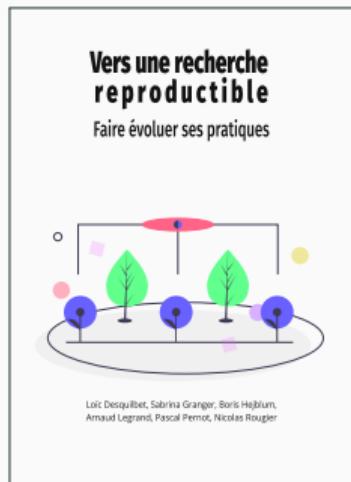
- MOOC Reproducible Research: Methodological principles for a transparent science
- 3rd Edition: ≈ March 2020
- MOOC RR "Advanced" planned for 2021



ACKNOWLEDGMENTS

Konrad Hinsen, Christophe Pouzat, Alexandre Hocquet

- MOOC Reproducible Research: Methodological principles for a transparent science
- 3rd Edition: ≈ March 2020
- MOOC RR "Advanced" planned for 2021



Sabrina Granger, Loïc Desquillet, Sabrina Granger, Boris Hejblum, Pascal Pernot, Nicolas Rougier

- A non-technical introduction to reproducibility issues (in French)

Slides (with hyperlinks) available at:

https://github.com/alegrand/SMPE/raw/master/lectures/talk_20_07_07_JDEV.pdf

PUBLIC EVIDENCE FOR A LACK OF REPRODUCIBILITY

- J.P. Ioannidis. *Why Most Published Research Findings Are False* PLoS Med. 2005.
- *Lies, Damned Lies, and Medical Science*, The Atlantic. Nov, 2010
- *Reproducibility: A tragedy of errors*, Nature, Feb 2016.
- Steen RG, *Retractions in the scientific literature: is the incidence of research fraud increasing?*, J. Med. Ethics 37, 2011

The collage includes the following sources:

- Los Angeles Times BUSINESS**: Headline: "Science has lost its way, at a big cost to humanity". Subtext: "Researchers are rewarded for splashy findings, not for double-checking accuracy. So many scientists looking for cures to diseases have been building on ideas that aren't even true."
- Science AAAS NEWS SCIENCE JOURNALS CAREERS MULTIMEDIA COLLECTIONS**: Headline: "Science 17 January 2014: Vol. 343 no. 6168 p. 229 DOI: 10.1126/science.1250475". Article title: "Reproducibility".
- nature International weekly journal of science**: Headline: "Announcement: Reducing our irreproducibility". Subtext: "Over the past year, Nature has published a string of articles that highlight the reliability and reproducibility of published research (collected in a special issue of the journal)."
- The Economist**: Headline: "How SCIENCE GOEs WRONG".
- TheScientist EXPLORING LIFE. INSPIRING INNOVATION NIH Tackles Irreproducibility**: Headline: "The federal agency speaks out about how to improve the quality of scientific research."

Courtesy V. Stodden,
SC, 2015

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received \approx \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received ≈ \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

Dieterik Stapel Professor, Social Psychology, Univ. Amsterdam, 2011

I failed as a scientist. I adapted research data and fabricated research. Not once, but several times, not for a short period, but over a longer period of time. [...] I am aware of the suffering and sorrow that I caused to my colleagues... I did not withstand the pressure to score, to publish, the pressure to get better in time. I wanted too much, too fast. In a system where there are few checks and balances, where people work alone, I took the wrong turn.

58 retracted publications

NEWSWORTHY STORIES ABOUT SCIENTIFIC MISCONDUCT

Dong-Pyou Han Assistant professor, Biomedical sciences, Iowa State University, 2013

Falsified blood results to make it appear as though a vaccine exhibited anti-HIV activity

- Han and his team received ≈ \$19 million from NIH
- 1 retracted publication and resignation of university. Sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials He was also fined US \$7.2 million!

Dieterik Stapel Professor, Social Psychology, Univ. Amsterdam, 2011

I failed as a scientist. I adapted research data and fabricated research. Not once, but several times, not for a short period, but over a longer period of time. [...] I am aware of the suffering and sorrow that I caused to my colleagues... I did not withstand the pressure to score, to publish, the pressure to get better in time. I wanted too much, too fast. In a system where there are few checks and balances, where people work alone, I took the wrong turn.

58 retracted publications

Brian Wansink Professor, Psychological Nutrition, Cornell, 2016

When she arrived, I gave her a data set of a self-funded, failed study which had null results. I said "This cost us a lot of time and our own money to collect. There's got to be something here we can salvage because it's a cool (rich & unique) data set." I told her what the analyses should be and what the tables should look like. [...] Every day she came back with puzzling new results, and every day we would scratch our heads, ask "Why," and come up with another way to reanalyze the data with yet another set of plausible hypotheses

17 retracted publications

A CREDIBILITY CRISIS?

Scientific misconduct is obviously wrong but it's **not new!**

- Every domain has its black sheep
- The publish or perish pressure is a huge pain

Media attention **inflates conspiracy opinions** 😞

Scientific results are worthless. Stop the scientific dictatorship/lobby!

The Battle against Scientific Fraud

CNRS International Magazine



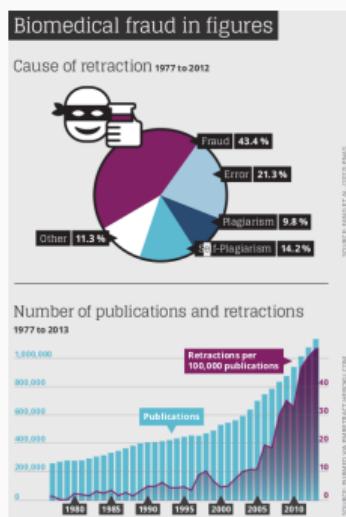
A CREDIBILITY CRISIS?

Scientific misconduct is obviously wrong but it's **not new!**

- Every domain has its black sheep
- The publish or perish pressure is a huge pain

Media attention inflates **conspiracy opinions** 😞

Scientific results are worthless. Stop the scientific dictatorship/lobby!



Fraud is the (**uninteresting**) visible part of the iceberg

- **Failing to reproduce the results of others is common**

1,500 scientists lift the lid on reproducibility,

Nature, May 2016

- How so? **Why now? Why is this important?** What can we do about it?

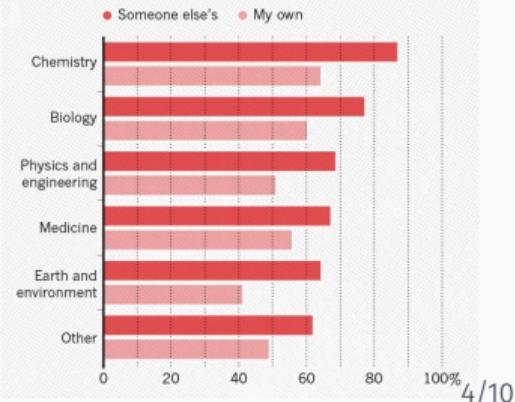
The Battle against Scientific Fraud

CNRS International Magazine



HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

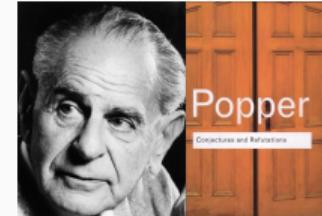
Most scientists have experienced failure to reproduce results.



REPRODUCIBILITY OF EXPERIMENTAL RESULTS IS THE HALLMARK OF SCIENCE

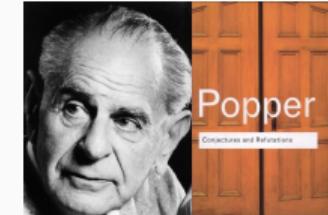
1934: Karl Popper puts the notions of **falsifiability** and **crucial experiment** as the **hallmark of science**

- If no experiment can be set up to **disprove** your theory, it is not science
- Good experiments **discriminate good theories from bad ones**
- **Non-reproducible** single occurrences are of no significance to science



REPRODUCIBILITY OF EXPERIMENTAL RESULTS IS THE HALLMARK OF SCIENCE

1934: Karl Popper puts the notions of **falsifiability** and **crucial experiment** as the **hallmark of science**



- If no experiment can be set up to **disprove** your theory, it is not science
- Good experiments **discriminate good theories from bad ones**
- **Non-reproducible** single occurrences are of no significance to science

An ideal rather than the norm

Popper's proposal works well for Physics from the 18th century but is not so simple for many other domains:

- Theory of evolution
- Spotting a SuperNova
- Particle Physics (a single LHC)
- Biology (every animal does not behave in the same way)
- Anthropology (impact on people from a remote culture)

REPRODUCIBILITY: A CORE VALUE OF SCIENCE

1. Universality: Science aims for **objective findings**, accessible to anyone
Reproducibility acts as a **Universality control**

2. Incremental: We build on each others work but everybody makes mistakes
Methods, biases, ... How to discriminate sound **theories** experiments from bad ones? 😊
Reproducibility acts as a **Quality control**

REPRODUCIBILITY: A CORE VALUE OF SCIENCE

1. Universality: Science aims for **objective findings**, accessible to anyone

Reproducibility acts as a **Universality control**

2. Incremental: We build on each others work but everybody makes mistakes

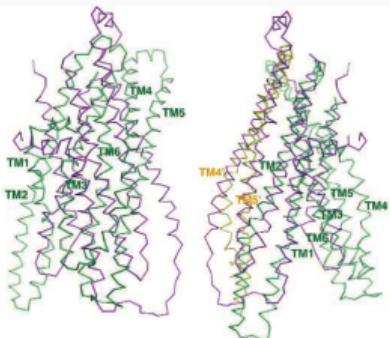
Methods, biases, ... How to discriminate sound **theories** experiments from bad ones? 😊

Reproducibility acts as a **Quality control**

But, **scientific practices have greatly evolved**, in particular since we rely on **computers**



HOW COMPUTERS BROKE SCIENCE – AND WHAT WE CAN DO ABOUT IT



Geoffrey Chang (Scripps, UCSD) works on crystallography and studies the structure of cell membrane proteins.

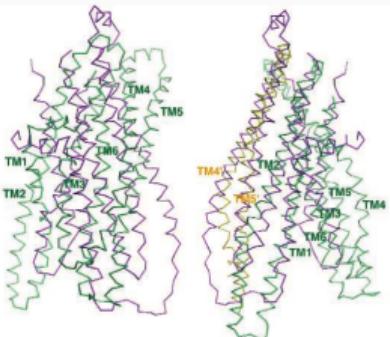
He specialized in structures of **multidrug resistant transporter proteins in bacteria**: MsbA de Escherichia Choli (Science, 2001), Vibrio cholera (Mol. Biology, 2003), Salmonella typhimurium (Science, 2005)

2006: Inconsistencies reveal a programming mistake

A homemade data-analysis program had flipped two columns of data, inverting the electron-density map from which his team had derived the protein structure.

5 retractions that motivate improved software engineering practices in comp. biology

HOW COMPUTERS BROKE SCIENCE – AND WHAT WE CAN DO ABOUT IT



Geoffrey Chang (Scripps, UCSD) works on crystallography and studies the structure of cell membrane proteins.

He specialized in structures of **multidrug resistant transporter proteins in bacteria**: MsbA de Escherichia Choli (Science, 2001), Vibrio cholera (Mol. Biology, 2003), Salmonella typhimurium (Science, 2005)

2006: Inconsistencies reveal a programming mistake

A homemade data-analysis program had flipped two columns of data, inverting the electron-density map from which his team had derived the protein structure.

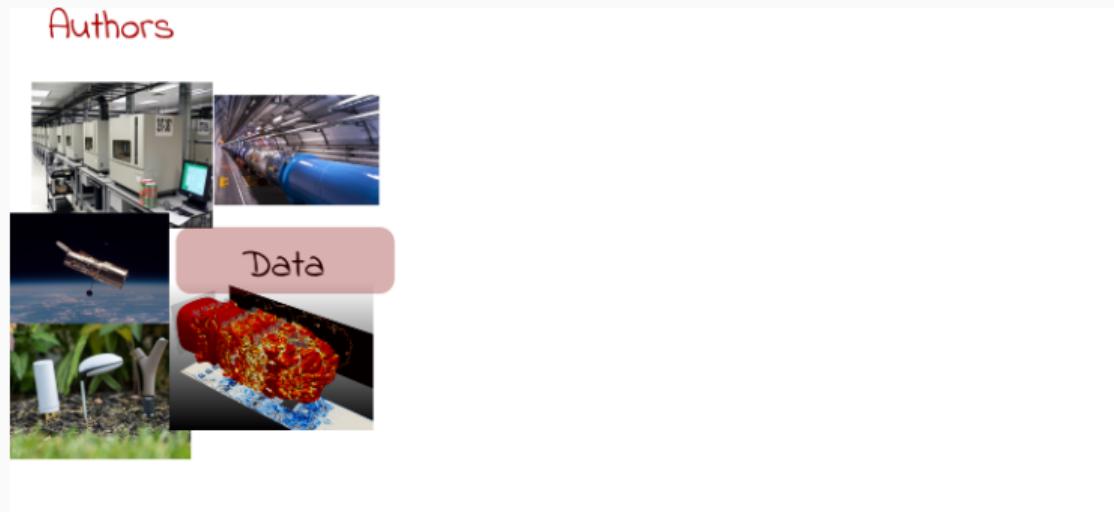
5 retractions that motivate **improved software engineering practices** in comp. biology

There is **worse**!

- The generalized and intensive use of **spreadsheets**
- Relying on **black box** statistical methods is infinitely easier than understanding them
 - Learning and Data Analytics frameworks are nuclear weapons
- Numerical errors and **software environment** unawareness

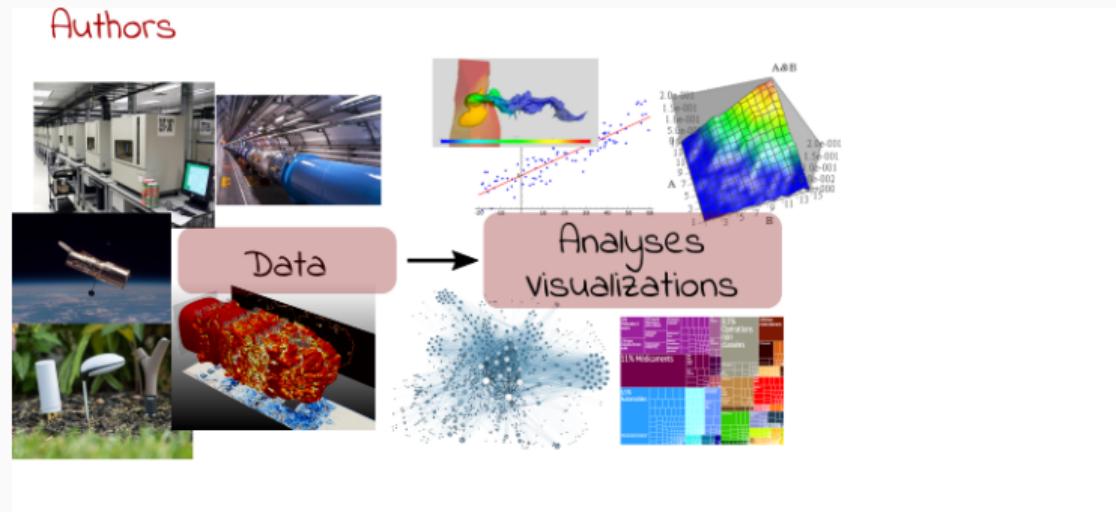
MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



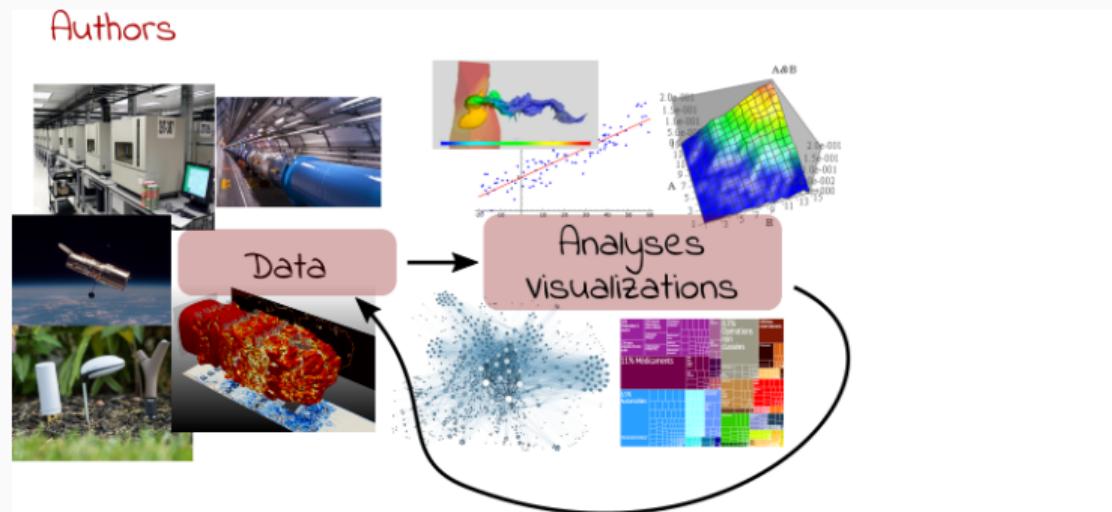
MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



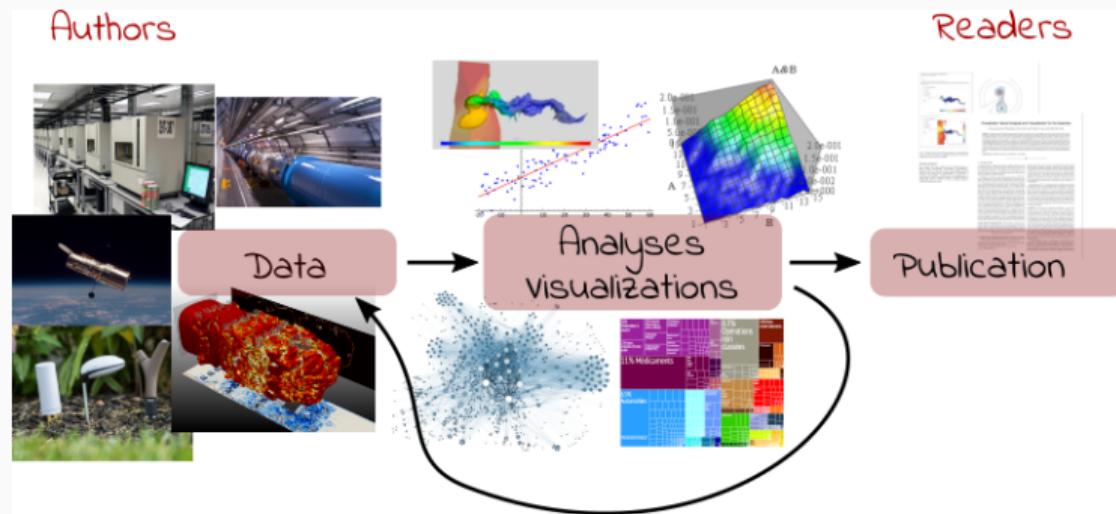
MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



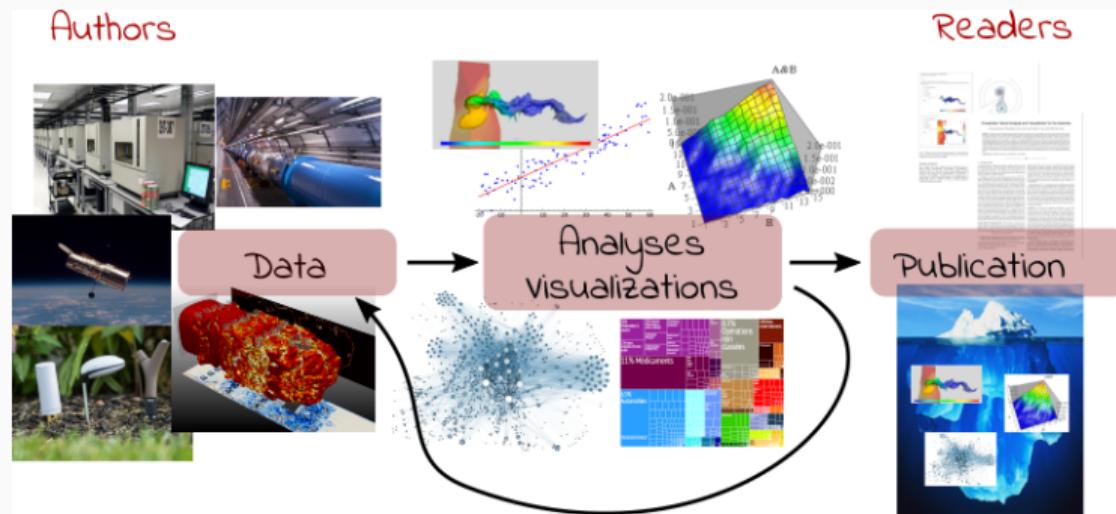
MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



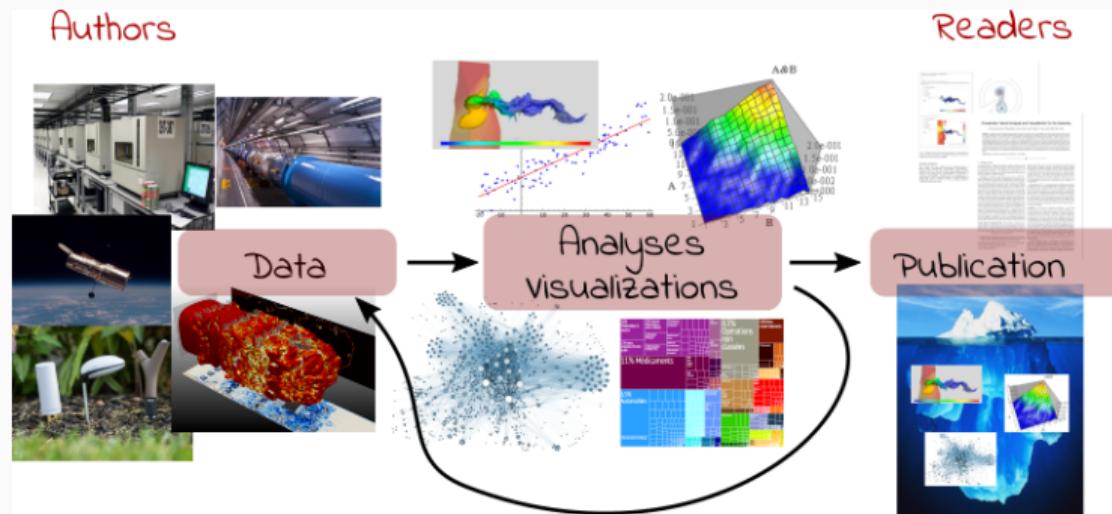
MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



MODERN SCIENCE

The processing steps between raw observations and findings have gotten increasingly numerous and complex.



Reproducible Research = Bridging the Gap by working Transparently

DIFFERENT (BUT CONVERGING) REPRODUCIBILITY CONCERNS

Bad computer/statistic/publication practices and "broke science" 😊

- Ensure articles and data are **available**
 - Open Access, Open Data, FAIR and DMP

DIFFERENT (BUT CONVERGING) REPRODUCIBILITY CONCERNS

Bad computer/statistic/publication practices and "broke science" 😊

- Ensure articles and data are **available**
 - Open Access, Open Data, FAIR and DMP
- *Ensure the experiment design and analysis is not broken*
 - ~~data torture and hacking, pre-registration, education, ...~~

DIFFERENT (BUT CONVERGING) REPRODUCIBILITY CONCERNS

Bad computer/statistic/publication practices and "broke science" 😊

- Ensure articles and data are **available**
 - Open Access, Open Data, FAIR and DMP
- *Ensure the experiment design and analysis is not broken*
 - ~~data torture and hacking, pre-registration, education, ...~~
- Explain **how** the computations unfold and **why**
 - Notebooks (that can quickly get out of control)
 - Simple computational workflows

DIFFERENT (BUT CONVERGING) REPRODUCIBILITY CONCERNS

Bad computer/statistic/publication practices and "broke science" 😊

- Ensure articles and data are **available**
 - Open Access, Open Data, FAIR and DMP
- *Ensure the experiment design and analysis is not broken*
 - ~~data torture and hacking, pre-registration, education, ...~~
- Explain **how** the computations unfold and **why**
 - Notebooks (that can quickly get out of control)
 - Simple computational workflows
- Allow to inspect, rerun, make variations on computations
 - Access to code, data, ...

DIFFERENT (BUT CONVERGING) REPRODUCIBILITY CONCERNS

Bad computer/statistic/publication practices and "broke science" 😊

- Ensure articles and data are **available**
 - Open Access, Open Data, FAIR and DMP
- *Ensure the experiment design and analysis is not broken*
 - ~~data torture and hacking, pre-registration, education, ...~~
- Explain **how** the computations unfold and **why**
 - Notebooks (that can quickly get out of control)
 - Simple computational workflows
- Allow to inspect, rerun, make variations on computations
 - Access to code, data, ... **options/parameters, environment, resources?**

This requires **first class software engineering practices** instead of building on prototypes

Software factories, Archives, and Provenance Tracking tools

WHAT'S NEXT?

Raising awareness on reproducibility issues, covering the "basics" of Reproducible Research?

- Note taking, version control, computational documents

Covered in the Reproducible Research MOOC



WHAT'S NEXT?

Raising awareness on reproducibility issues, covering the "basics" of Reproducible Research?

- Note taking, version control, computational documents

Covered in the Reproducible Research MOOC



Friday at the JDEV, we'll cover technical questions

- Software Environment Control (Docker, Guix / Nix, ...)
- Continuous Integration / deployment and DevOps
- Pipelines and Workflows
- Perennial software Archives and References (Software Heritage)
- Numerical Reproducibility and Parallel Codes
- Dealing with all this at once, in an industrial way, whereas researchers are crafts[wo]men