

# REPRODUCIBILITY CRISIS AND OPEN SCIENCE

---

Arnaud Legrand



Séminaire RR, Montpellier, June 2021



# 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



Science has lost its way, at a big cost to humanity

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.

The screenshot shows the Science journal website. At the top, there are links for LOCAL, U.S., WORLD, BUSINESS, SPORTS, ENTERTAINMENT, HEALTH, STYLE, TRAVEL, and more. Below this, a large headline reads "Science has lost its way, at a big cost to humanity". Underneath, a sub-headline says "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." The main content area features an article by Marcia McNutt titled "Reproducibility". The article summary states: "Science advances on a foundation of trusted data. But a recent study found that the approach that scientists used to gain confidence in their results was failing. The scientific community was shaken by reports that a troubling number of studies could not be reproduced." The article includes a "Read Full Text to Comment (8)" link. On the left sidebar, there are sections for Article Views, Summary, Full Text, Full Text (PDF), Article Tools, and Related Content. The right sidebar includes links for Save to My Folders, Download Citation, Alert Me When Article Is Cited, Post to CiteULike, E-mail This Page, Rights & Permissions, Commercial Reprints and E-Prints, and View publication Citation.

The screenshot shows the Nature journal website. At the top, there is a banner for "Announcement: Reducing our irreproducibility". Below this, the main navigation menu includes Home, News & Comment, Research, Careers & Jobs, Current Issue, Archive, and more. The main content area features an article titled "Announcement: Reducing our irreproducibility" dated 24 April 2013. The article discusses the journal's commitment to improving reproducibility in scientific research.

A large graphic on the right side of the page. It features the word "nature" in its signature font, followed by "international weekly journal of science". Above the "nature" logo, there is a small image of a brain. To the left of the "nature" logo, the text "Announcement: Reducing our irreproducibility" is visible. To the right, there is a large, stylized title "HOW SCIENCE GOES WRONG" with a colorful, abstract background. Below the main title, there is smaller text: "Jon's lawyer scruples", "a nuclear deal with Jon", "Investment tips from Nobel economists", "Junk bonds are back", and "The meaning of Sezin Tendilur".

The screenshot shows the Nature journal website. At the top, there is a banner for "Announcement: Reducing our irreproducibility". Below this, the main navigation menu includes Menu, Advanced search, and Search. The main content area features an article titled "Must try harder" dated 28 March 2012. The article discusses the challenges of reproducibility in science and the need for better practices. The right sidebar includes links for PDF, Citation, Reprints, Rights & permissions, and Article metrics.

The screenshot shows The Scientist magazine website. The main headline is "NIH Tackles Irreproducibility" with the sub-headline "The federal agency speaks out about how to improve the quality of scientific research." Below the headline, it says "By Jef Akst | January 28, 2014". The right sidebar includes links for PDF, Citation, Reprints, Rights & permissions, and Article metrics.

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. \$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  $\approx$  \$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. \$7.2 million!

**Diederik Stapel** Professor, Social Psychology, Univ. Tilburg, 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  $\approx$  \$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. \$7.2 million!

**Diederik Stapel** Professor, Social Psychology, Univ. Tilburg, 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

*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. [...] 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

# SCIENTIFIC MISCONDUCT? WHAT ARE THE CONSEQUENCES ?

**Reinhart and Rogoff** Professors of Economics at Harvard

*gross debt [...] exceeding 90 percent of the economy has a significant negative effect on economic growth* – Growth in a Time of Debt (2010)

*While using RR's working spreadsheet, we identified coding errors, selective exclusion of available data, and unconventional weighting of summary statistics.* – 2013: Herndon, Ash and Pollin

*For 3 years, austerity was not presented as an option but as a necessity.*

– 2013: Paul Krugman

At least, a scientific debate has been possible.

# SCIENTIFIC MISCONDUCT? WHAT ARE THE CONSEQUENCES ?

**Reinhart and Rogoff** Professors of Economics at Harvard

*gross debt [...] exceeding 90 percent of the economy has a significant negative effect on economic growth* – Growth in a Time of Debt (2010)

*While using RR's working spreadsheet, we identified coding errors, selective exclusion of available data, and unconventional weighting of summary statistics.* – 2013: Herndon, Ash and Pollin

*For 3 years, austerity was not presented as an option but as a necessity.*

– 2013: Paul Krugman

At least, a scientific debate has been possible.

## Bad science is deleterious

- It is used to backup stupid politics, it affects people's life, ...
- It blurs the frontier between scientists and crooks

## Media attention inflates conspiracy opinions 😞

- *Scientific result are worthless.*
- *Scientists can't even agree with each others on economy/climate/vaccine/5G/...*
- *Stop the scientific dictatorship/lobby!*

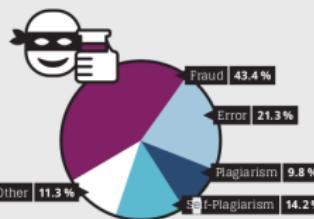
# A CREDIBILITY CRISIS?

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

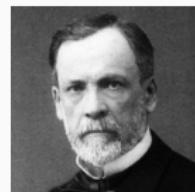
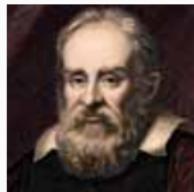
## *The Battle against Scientific Fraud in the CNRS International Magazine*

### Biomedical fraud in figures

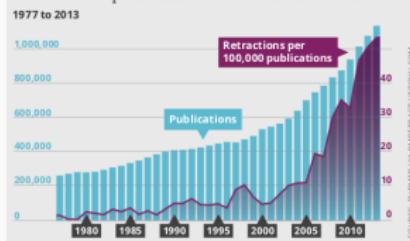
Cause of retraction 1977 to 2012



SOURCE: FAUCI ET AL. (2012) PNAS



Number of publications and retractions



Galileo (data fabrication), Ptolemy (plagiarism), Mendel (data enhancement), Pasteur (rigorous but hid failures), ...

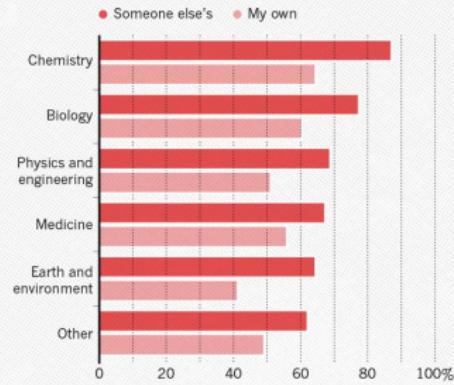
Scientific misconduct is obviously wrong but it's **not new!**

- Every domain has its black sheep
- The publish or perish pressure is a pain 4/100

# A REPRODUCIBILITY CRISIS?

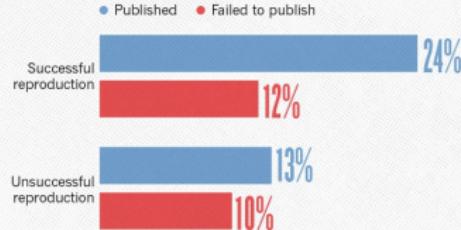
## HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



## HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.



1,500 scientists lift the lid on reproducibility,

Nature, May 2016

## Social causes

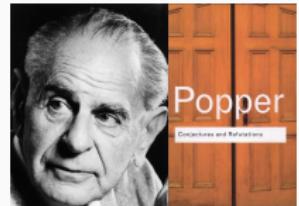
- Fraud, conflict of interest (pharmaceutic, ...)
- No incentive to reproduce/check our own work (afap), nor the work of others (big results!), nor to allow others to check (competition)
- Peer review does not scale: 1+ million articles per year!

## Methodological or technical causes

- The many biases (apophenia, confirmation, hindsight, experimenter, ...): bad designs
- Selective reporting, weak analysis (statistics, data manipulation mistakes, computational errors)
- Lack of information, code/raw data unavailable

# REPRODUCIBILITY OF EXPERIMENTAL RESULTS: THE HALLMARK OF SCIENCE

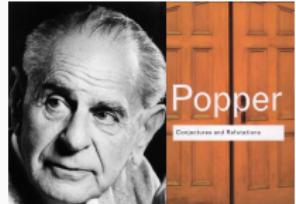
1934: Karl Popper puts the notions of **falsifiability** and **crucial experiment** as the **hallmark of science**



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

# REPRODUCIBILITY: A CORE VALUE OF SCIENCE

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

Reproducibility acts as a Universality/Robustness 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/Robustness 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

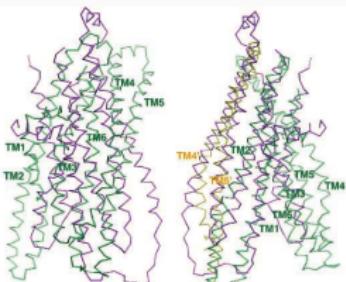
But, scientific practices have greatly evolved, in particular since we rely on computers



How computers broke science – and what we can do about it

– Ben Marwick, The conversation, 2015

# How COMPUTERS BROKE SCIENCE



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

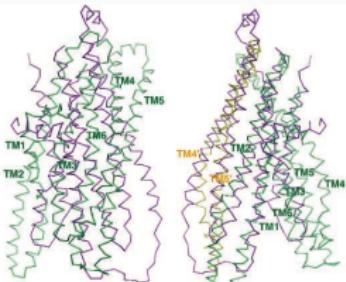
He specialized in structures of **multidrug resistant transporter proteins in bacteria**: MsbA de Escherichia Choli (Science, 2001), Vibrio cholera (Mol. Biology, 2003), Salmonella typhimurium (Science, 2005)

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



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** (**COVID tracing**)
- Relying on **black box** statistical methods is infinitely easier than understanding them  
(Learning and Data Analytics frameworks = nuke)
- Numerical errors and software environment unawareness

# DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

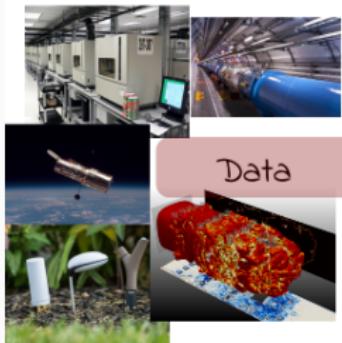
**Social Sciences, Oncology, ...** methodology, statistics, pre-registration

**Genomics** software engineering, computational reproducibility, provenance

**Computational fluid dynamics** numerical issues

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

## Authors



Data

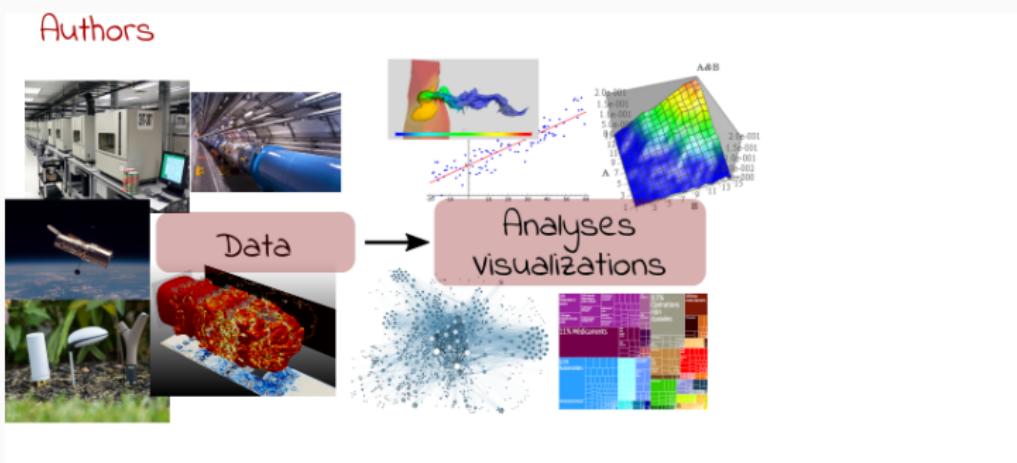
# DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

**Social Sciences, Oncology, ...** methodology, statistics, pre-registration

**Genomics** software engineering, computational reproducibility, provenance

**Computational fluid dynamics** numerical issues

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



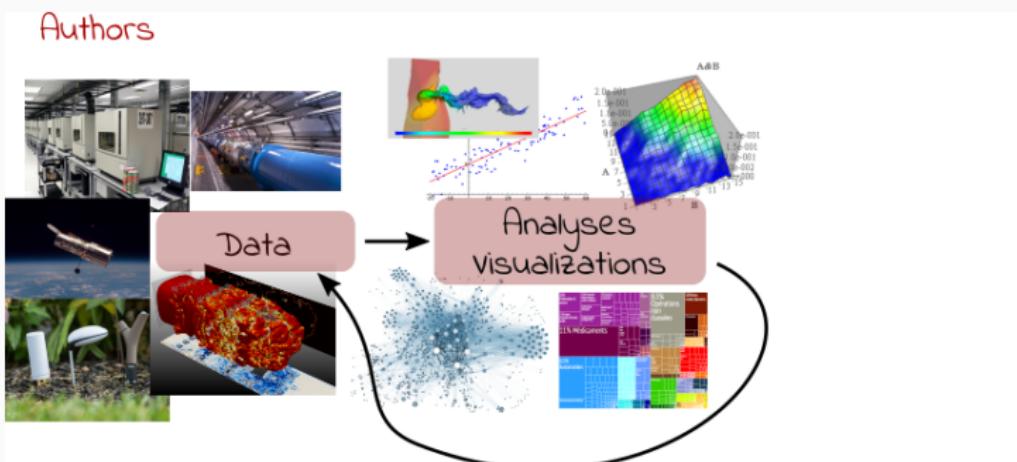
## Different Reproducibility Concerns in Modern Science

Social Sciences, Oncology, ... methodology, statistics, pre-registration

**Genomics** software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

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



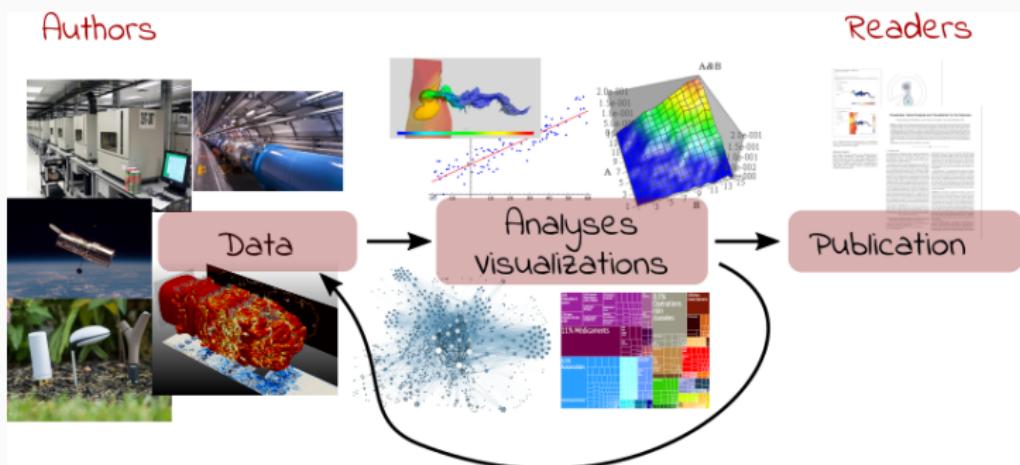
# DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics, pre-registration

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

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



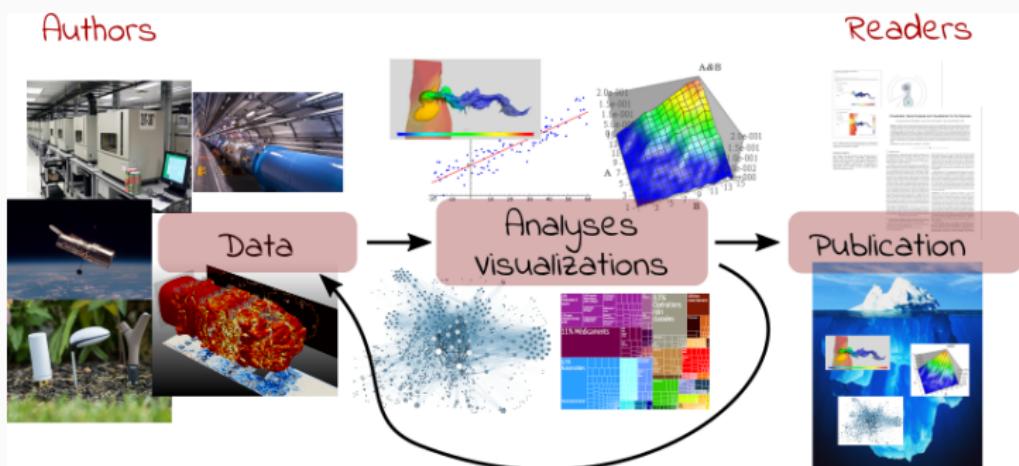
# DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics, pre-registration

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

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



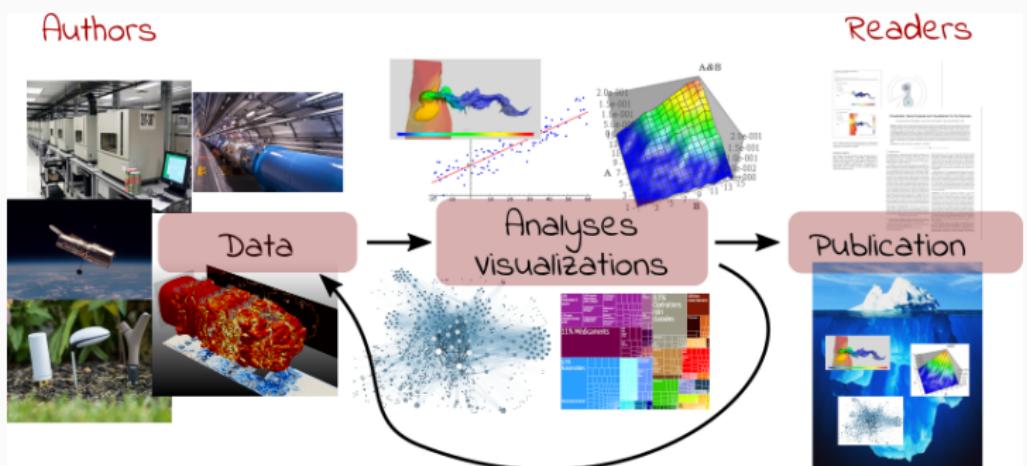
# DIFFERENT REPRODUCIBILITY CONCERN IN MODERN SCIENCE

Social Sciences, Oncology, ... methodology, statistics, pre-registration

Genomics software engineering, computational reproducibility, provenance

Computational fluid dynamics numerical issues

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



Reproducible Research = Bridging the Gap by working Transparently 9/100

# REPRODUCIBLE RESEARCH PRACTICES

---

# "REPRODUCIBLE RESEARCH": FIRST APPEARANCE

Claerbout & Karrenbach, meeting of the Society of Exploration Geophysics, 1992

## Electronic Documents Give Reproducible Research a New Meaning

RE1.3

Jon F. Claerbout and Martin Karrenbach, Stanford Univ.

### SUMMARY

A revolution in education and technology transfer follows from the marriage of word processing and software command scripts. In this marriage an author attaches to every figure caption a pushbutton or a name tag usable to recalculate the figure from all its data, parameters, and programs. This provides a new meaning of reproducibility in computer documents.

In 1990, we set this sequence of goals:

- Learn how to merge a publication with its underlying computational analysis.
- Teach researchers how to prepare a document in a form where they themselves can reproduce their own research results a year or more later by "pressing a single button".
- Learn how to leave finished work in a condition where coworkers can reproduce the calculation including the final illustration by pressing a button in its caption.
- Prepare a complete copy of our local software environment so that graduating students can take their work away with them to other sites, press a button, and reproduce their Stanford work.
- Merge electronic documents written by multiple authors (SEP reports).

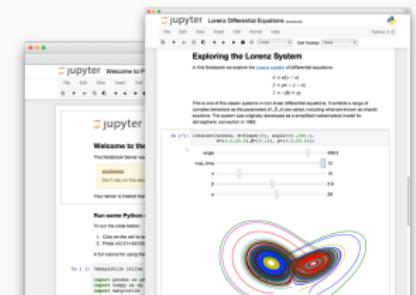
- make incremental improvements in electronic-document software
- seek partners for broadening standards (and making incremental improvements).

Our basic goal is reproducible research. The electronic document is our means to this end. In principle, reproducibility in research can be achieved without electronic documents and that is how we started. Our first nonelectronic reproducible document was a textbook in which the paper document contained the name of a program script in every figure caption. The program scripts were organized by book chapter and section so they could be correlated to an accompanying magnetic tape dump of the file system. The magnetic tape also contained all the necessary data to feed the program script.

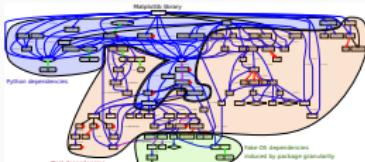
Now that we have begun using CD-ROM publication, we can go much further. Every figure caption contains a pushbutton that jumps to the appropriate science directory (folder) and initiates a figure rebuild command and then displays the figure, possibly as a movie or interactive program. We normally display seismic images of the earth's interior, but to reach wider audiences, Figure 1 shows a satellite weather picture which the pushbutton will animate as seen on commercial television. We include all our plot software as well as freely available software from many sources, including compilers and the L<sup>A</sup>T<sub>E</sub>X word processing systems. Naturally we cannot include licensed software, but with the exception

# EXISTING TOOLS, EMERGING STANDARDS

## Notebooks and workflows



## Software environments



## Sharing platforms



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

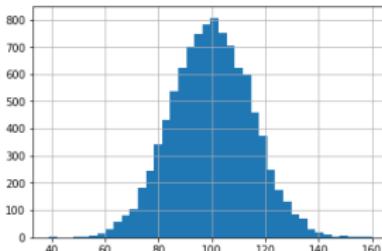
3.141592653589793

Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1)/N)
```

3.1437198694098765

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook window titled "example\_ip" (modified). It contains three code cells:

- In [1]:** A cell containing Python code to print pi.

```
from math import *  
print(pi)  
3.141592653589793
```

Output: Mais calculé avec la [méthode des aiguilles de Buffon](#) ([https://fr.wikipedia.org/wiki/Aiguille\\_de\\_Buffon](https://fr.wikipedia.org/wiki/Aiguille_de_Buffon)), on obtient d'abord comme approximation :
- In [2]:** A cell containing Python code to calculate the value of pi using theBuffon's needle method.

```
import numpy as np  
N = 1000000  
x = np.random.uniform(size=N, low=0, high=1)  
theta = np.random.uniform(size=N, low=0, high=np.pi/2)  
2/(sum((x+np.sin(theta))>1))/N
```

Output: 3.14371986944998765
- In [3]:** A cell containing Python code to generate a histogram of random numbers between 0 and 150, centered at 100.

```
%matplotlib inline  
import matplotlib.pyplot as plt  
  
mu, sigma = 100, 15  
x = mu + sigma*np.random.randn(10000)  
  
plt.hist(x,40)  
plt.grid(True)  
plt.show()
```

Output: A histogram showing a bell-shaped distribution of values between 40 and 160, centered around 100.

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

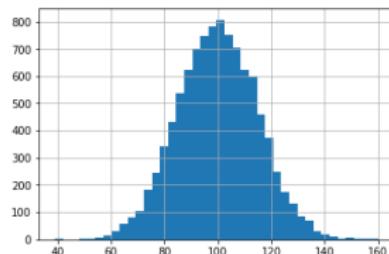
3.141592653589793

Mais calculé avec la [méthode des aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np  
N = 1000000  
x = np.random.uniform(size=N, low=0, high=1)  
theta = np.random.uniform(size=N, low=0, high=np.pi/2)  
2/(sum((x+np.sin(theta))>1))/N
```

3.14371986944998765

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

A screenshot of a Jupyter Notebook interface. The title cell contains the text '# Un document computationnel'. Below it, a text cell says 'Mon ordinateur m'indique que  $\pi$  vaut "approximativement"'. An In [1] cell shows Python code to print pi, which outputs 3.141592653589793. An Out [1] cell contains text about calculating pi using theBuffon's needle method, mentioning the URL [https://fr.wikipedia.org/wiki/Aiguille\\_de\\_Buffon](https://fr.wikipedia.org/wiki/Aiguille_de_Buffon). An In [2] cell shows code to calculate pi using the needle method, outputting 3.1437198694998765. An Out [2] cell contains mathematical formulas for the Buffon's needle method and a note about normalization. An In [3] cell shows code to generate a histogram, outputting a plot of a normal distribution centered at 100.

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

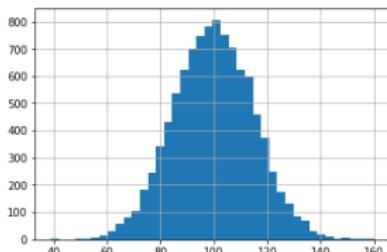
3.141592653589793

Mais calculé avec la [méthode des aiguilles de Buffon](#) ([https://fr.wikipedia.org/wiki/Aiguille\\_de\\_Buffon](https://fr.wikipedia.org/wiki/Aiguille_de_Buffon)), on obtient une approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=np.pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694998765

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

```
# Un document computationnel

In [1]: from math import * print(pi) 3,141592653589793

Mais calculé avec la __methodes__ des __émpilements__ de __Buffon__ (https://fr.wikipedia.org/wiki/Alg\u00fclie\_de\_Buffon), on obtiendrait comme __approximation__ : 3.141592653589793

In [2]: import numpy as np N = 1000000 x = np.random.uniform(size=N, low=0, high=1) theta = np.random.uniform(size=N, low=0, high=pi/2) 2*(sum((x+np.sin(theta))>1))/N Out[2]: 3,1437198694098765

On peut inclure des formules math\u00e9matiques comme $\\frac{1}{\\sigma\\sqrt{2\\pi}} \\exp\\left(-\\frac{(x-\\mu)^2}{2\\sigma^2}\\right)$ et des dessins qui n'ont rien \u00e0 voir avec $\\pi$ (si ce n'est une constante de normalisation... ☺).

In [3]: %matplotlib inline import matplotlib.pyplot as plt mu, sigma = 100, 15 x = mu + sigma*np.random.randn(10000) plt.hist(x,60) plt.grid(True) plt.show()
```

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

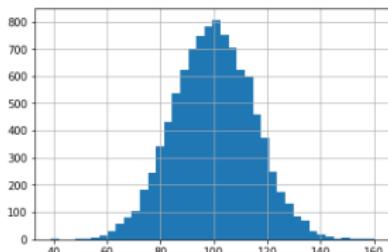
3.141592653589793

Mais calculé avec la **méthode des aiguilles de Buffon**, on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2*(sum((x+np.sin(theta))>1))/N
```

3.1437198694098765

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

A screenshot of a Jupyter Notebook interface. The title bar says "jupyter example\_pi". The notebook contains three code cells:

- In [1]:** Prints the value of pi as 3.141592653589793.
- In [2]:** Calculates pi using theBuffon's needle method. It imports numpy, generates random numbers for x and theta, and calculates the ratio of points where x+np.sin(theta) > 1 to N. The output is 3.1437198694098765.
- In [3]:** Plots a histogram of x values from 40 to 160. The distribution is roughly bell-shaped, centered around 100.

Annotations in red highlight the output of In [2] and the histogram in In [3]. A large red arrow labeled "Résultats" points from the notebook towards the final document on the right.

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

3.141592653589793

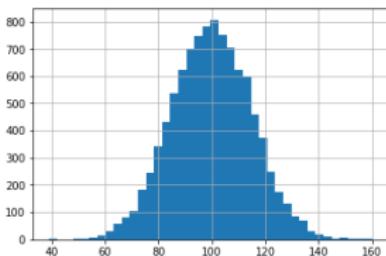
Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2*(sum((x+np.sin(theta))>1))/N
```

3.1437198694098765

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et

des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

A screenshot of a Jupyter Notebook interface. The top bar shows 'jupyter example\_pi' and 'Python 3'. Below is a toolbar with various icons. The notebook has three cells:

- In [1]:** Displays the code `from math import \*` and `print(pi)`, followed by the output `3.141592653589793`.
- In [2]:** Displays the code for calculating pi using Buffon's needle method, including imports for numpy and random, and a formula for approximation.
- In [3]:** Displays the code for generating a histogram of random numbers, with a plot showing a bell-shaped curve centered around 100.

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

3.141592653589793

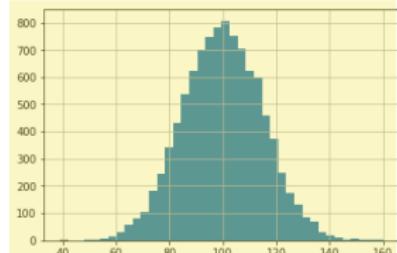
Mais calculé avec la méthode des [aiguilles de Buffon](#), on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1))/N
```

3.1437198694098765

Export

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



# TOOL 1: COMPUTATIONAL NOTEBOOKS/LITTERATE PROGRAMMING

Document initial dans son environnement

The screenshot shows a Jupyter Notebook interface with three code cells:

- In [1]:** Prints the value of pi (3.141592653589793) and includes a note about calculating pi with Buffon's needle method.
- In [2]:** Generates random points (x, theta) and calculates an approximation of pi based on the ratio of points where sin(theta) >= 0.
- In [3]:** Plots a histogram of x values, showing a bell-shaped distribution centered around 100.

At the bottom, there are logos for Jupyter and R Studio.

Document final

## Un document computationnel

Mon ordinateur m'indique que  $\pi$  vaut approximativement

3.141592653589793

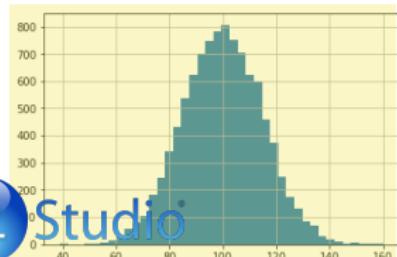
Mais calculé avec la **méthode des aiguilles de Buffon**, on obtiendrait comme approximation :

```
import numpy as np
N = 1000000
x = np.random.uniform(size=N, low=0, high=1)
theta = np.random.uniform(size=N, low=0, high=pi/2)
2/(sum((x+np.sin(theta))>1)/N)
```

3.1437198694098765

Export

On peut inclure des formules mathématiques comme  $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  et des dessins qui n'ont rien à voir avec  $\pi$  (si ce n'est une constante de normalisation... ☺).



<https://jupyterhub.u-ga.fr/>

## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

What is hiding behind a simple

```
import matplotlib
```

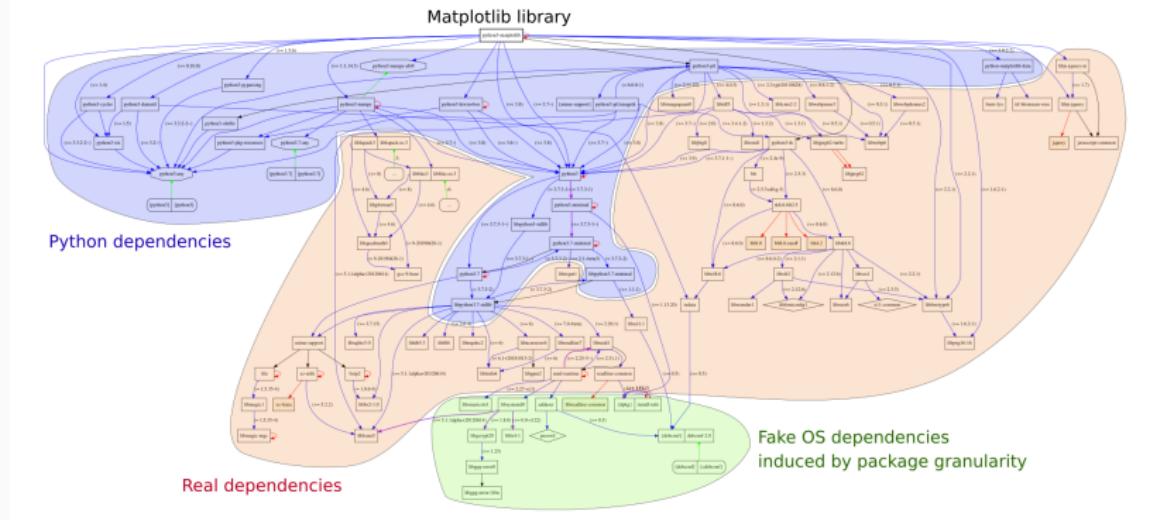
Package: python3-matplotlib  
Version: 2.1.1-2  
Depends: python3-dateutil, python-matplotlib-data (>= 2.1.1-2),  
python3-pyparsing (>= 1.5.6), python3-six (>= 1.10), python3-tz,  
libjs-jquery, libjs-jquery-ui, python3-numpy (>= 1:1.13.1),  
python3-numpy-abi9, python3 (<< 3.7), python3 (>= 3.6~),  
python3-cycler (>= 0.10.0), python3:any (>= 3.3.2-2~), libc6 (>= 2.14), libfreetype6 (>= 2.2.1), libgcc1 (>= 1:3.0), libpng16-16 (>= 1.6.2-1), libstdc++6 (>= 5.2), zlib1g (>= 1:1.1.4)

## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

## What is hiding behind a simple

```
import matplotlib
```

Package: python3-matplotlib



## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```

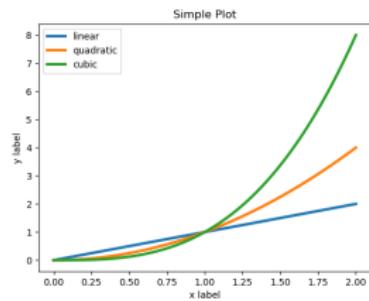
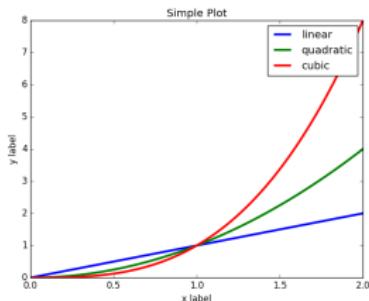
3

3.3333333333333335

## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

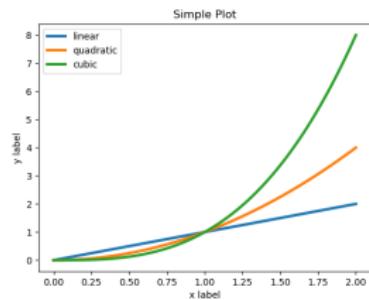
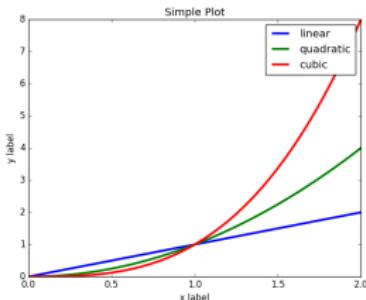
```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```



## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```

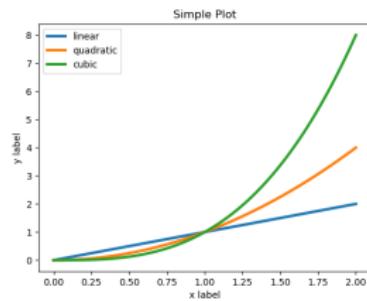
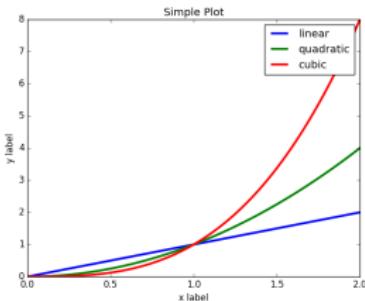


Cortical Thickness Measurements (PLOS ONE, June 2012) FreeSurfer:  
differences were found between the Mac and HP workstations and between  
Mac OSX 10.5 and OSX 10.6.

## TOOL 2: FIGHTING SOFTWARE ENVIRONMENTS NIGHTMARE

Python and its rapidly evolving environment

```
python2 -c "print(10/3)"  
python3 -c "print(10/3)"
```



Cortical Thickness Measurements (PLOS ONE, June 2012) FreeSurfer:  
differences were found between the Mac and HP workstations and between  
Mac OSX 10.5 and OSX 10.6.



## TOOL 3: FIGHTING INFORMATION LOSS WITH ARCHIVES

- D. Spinellis. *The Decay and Failures of URL References*. CACM, 46(1), 2003  
*The half-life of a referenced URL is approximately 4 years from its publication date.*
- P. Habibzadeh. *Decay of References to Web sites in Articles Published in General Medical Journals: Mainstream vs Small Journals*. Applied Clinical Informatics. 4 (4), 2013  
*half life ranged from 2.2 years in EMHJ to 5.3 years in BMJ*

## TOOL 3: FIGHTING INFORMATION LOSS WITH ARCHIVES

- D. Spinellis. *The Decay and Failures of URL References*. CACM, 46(1), 2003  
*The half-life of a referenced URL is approximately 4 years from its publication date.*
- P. Habibzadeh. *Decay of References to Web sites in Articles Published in General Medical Journals: Mainstream vs Small Journals*. Applied Clinical Informatics. 4 (4), 2013  
*half life ranged from 2.2 years in EMHJ to 5.3 years in BMJ*

Article archives



Data archives



figshare



Software Archive



Software Heritage



or

= awesome collaborations ( $\neq$  archive)

WHAT WILL IT TAKE ?

---

# CHANGING RESEARCH PRACTICES

Soft. Engineering, Statistics, and Reproducible Research in the curricula

**Manifesto:** "*I solemnly pledge*" (**WSSSPE, Lorena Barba, FAIR**)

1. I will teach my graduate students about reproducibility
2. All our research code (and writing) is under version control
3. We will always carry out verification and validation
4. We will share data, plotting script & figure under CC-BY
5. We will upload the preprint to arXiv at the time of submission of a paper
6. We will release code at the time of submission of a paper
7. We will add a "Reproducibility" declaration at the end of each paper
8. I will keep an up-to-date web presence



## Artifact evaluation and ACM badges



## Major conferences

- Supercomputing: Artifact Description (AD) mandatory, Artifact Evaluation (AE) still optional, Double blind vs. RR
- NeurIPS, ICLR: open reviews, reproducibility challenge



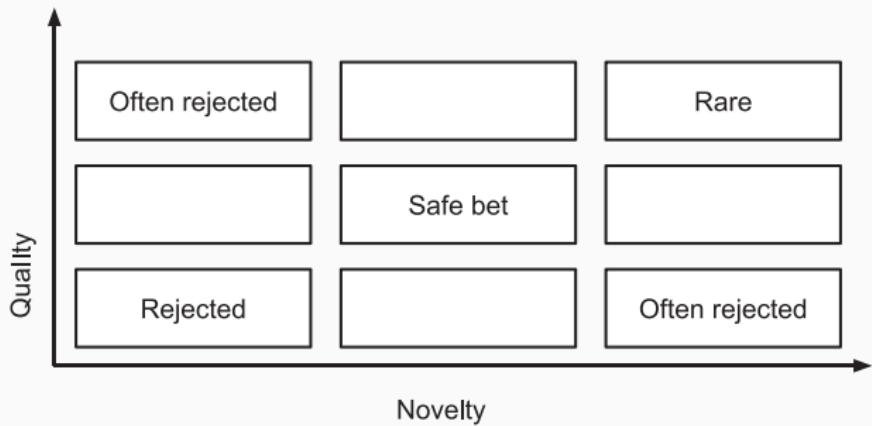
Joelle Pineau @ NeurIPS'18

- ACM SIGMOD 2015-2019, Most Reproducible Paper Award...

Mentalities are evolving people care, make stuff available, errors are found and fixed

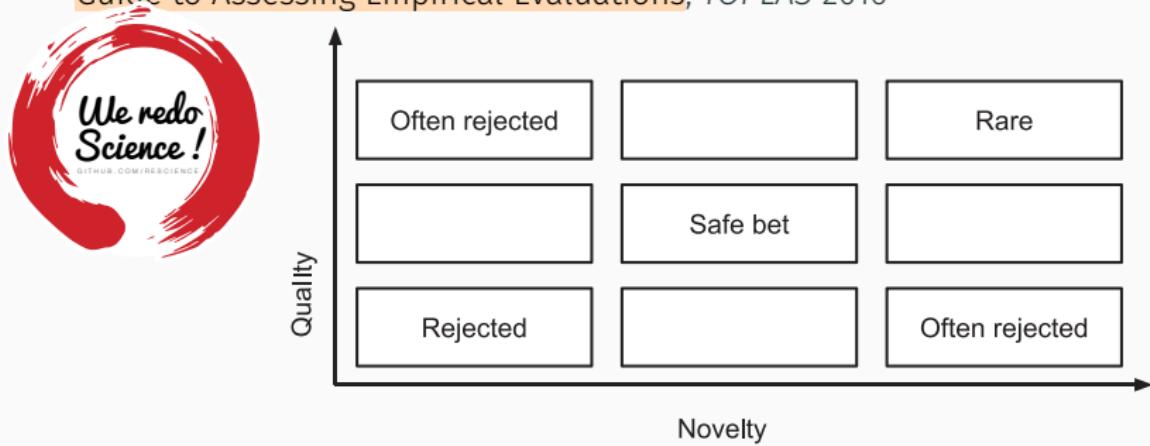
# CHANGING ACADEMIC PRACTICES (PUBLISH OR PERISH)

- Goodhart's Law: Are Academic Metrics Being Gamed?, M. Fire 2019
  - AI: over 1,000 ranked journals ( $\times 10$  in 15 years)
  - Shorter papers with increasing self references
  - More and more papers without any citation
  - Sharp increase in the number of new authors publishing at a much faster rate given their career age
- The Truth, The Whole Truth, and Nothing But the Truth: A Pragmatic, Guide to Assessing Empirical Evaluations, TOPLAS 2016



# CHANGING ACADEMIC PRACTICES (PUBLISH OR PERISH)

- Goodhart's Law: Are Academic Metrics Being Gamed?, M. Fire 2019
  - AI: over 1,000 ranked journals ( $\times 10$  in 15 years)
  - Shorter papers with increasing self references
  - More and more papers without any citation
  - Sharp increase in the number of new authors publishing at a much faster rate given their career age
- The Truth, The Whole Truth, and Nothing But the Truth: A Pragmatic, Guide to Assessing Empirical Evaluations, TOPLAS 2016



# WHAT ABOUT OPEN SCIENCE ?

## Plan National pour la Science Ouverte (BSN ~ CoSO)

- CNRS, Inria, INRAE, ...
- Many flavors: *Citizen Science*

### Main pillars:

1. Open access
2. Open data
3. Open source
  - Open hardware
4. Open methodology (**Reproducible Research**)
  - Open-notebook science
  - Open science infrastructures
5. Open peer review



**NO TRANSPARENCY  
NO CONSENSUS**



6. Open educational resources



# RESOURCES AND ACKNOWLEDGMENTS



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

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

# RESOURCES AND ACKNOWLEDGMENTS



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

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

MOOC Reproducible Research: Methodological principles for a transparent science, Learning Lab Inria

- Konrad Hinsen, Christophe Pouzat
- 3rd Edition: March 2020 – March 2022
- MOOC RR "Advanced" planned for 2021 2022
  - Software environment control
  - Scientific workflow
  - Managing data

