

Reproducible Research: Where Do We Stand?

Arnaud Legrand
CNRS, Inria, University of Grenoble

May 17, 2017 – PRECIS Summer School, Fréjus

Outline

- ① A Few Motivating Examples
- ② Science crisis ?
- ③ Reproducible Research/Open Science in a Nutshell
- ④ Illustrating Nice Ideas Through Different Tools
- ⑤ What can Computer Scientists do ?

Frustration



As an Author

- Advisor: "Did you take care of setting this?" Me: "Uh?"
- I thought I used the same parameters but I'm getting different results! I swear it **worked yesterday!**
- A new student wants to compare with the method I proposed last year
- The damned fourth reviewer asked for a major revision and wants me to **change figure 3**. 😞 Which code and which data set did I use to generate this figure?
- 6 months later: **Why** did I do that?

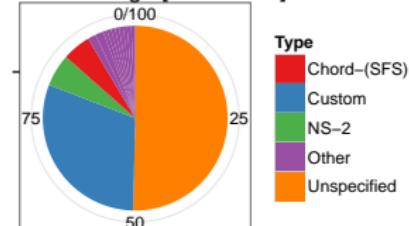
As a Reviewer This may be an interesting contribution but:

- There is no label/legend/... What is the **meaning of this graph**? If only I could access the generation script and get rid of the **logscale**
- This **average value** must hide something. As usual, no **confidence interval**... I wonder whether the difference is **significant** at all
- That can't be true, I'm sure they **removed some points** or decided to show only a **subset of the data**. I wonder what the rest looks like
- Is this improvement **solely the result of this naive idea**?

A Few Edifying Examples in my Domain

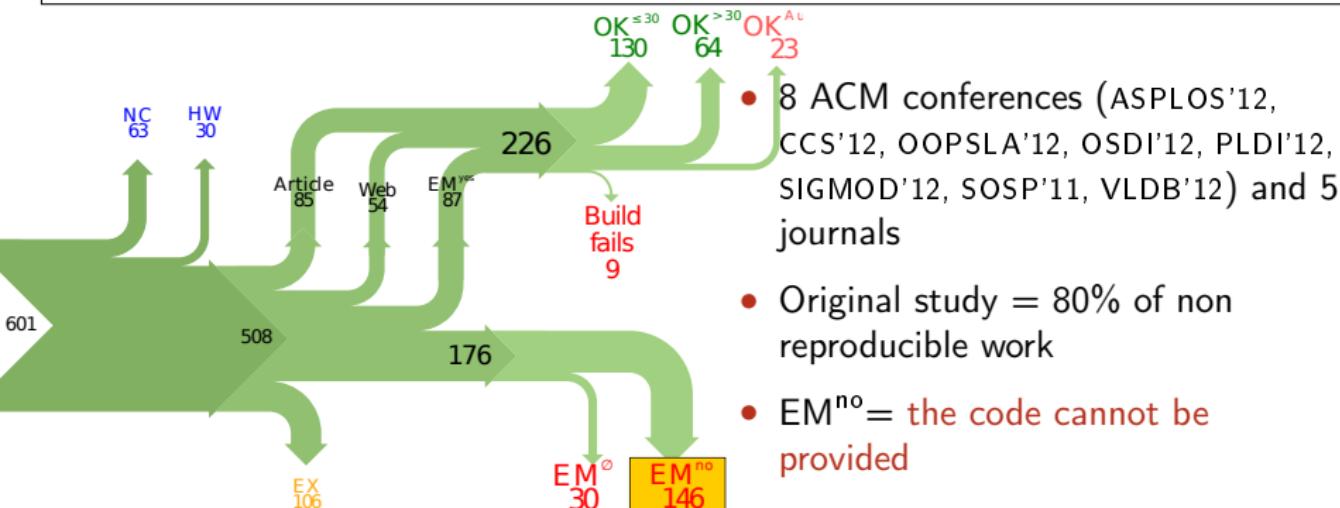
Naicken, Stephen et Al., *Towards Yet Another Peer-to-Peer Simulator*, HET-NETs'06.

Simulator usage [Naicken06]



From 141 P2P sim.papers, 30% use a custom tool,
50% don't report used tool

Collberg, Christian et Al., *Measuring Reproducibility in Computer Systems Research*,
<http://reproducibility.cs.arizona.edu/> 2014,2015



The Dog Ate my Homework !!!

- Versioning Problems

Thanks for your interest in the implementation of our paper. The good news is that I was able to find some code. I am just hoping that it is a stable working version of the code, and matches the implementation we finally used for the paper. Unfortunately, I have lost some data when my laptop was stolen last year. The bad news is that the code is not commented and/or clean.

Attached is the <system> source code of our algorithm. I'm not very sure whether it is the final version of the code used in our paper, but it should be at least 99% close. Hope it will help.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices

Unfortunately, the server in which my implementation was stored had a disk crash in April and three disks crashed simultaneously. While the help desk made significant effort to save the data, my entire implementation for this paper was not found.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon

*Unfortunately the current system is **not mature enough at the moment**, so it's not yet publicly available. We are actively working on a number of extensions and **things are somewhat volatile**. However, once things stabilize we plan to release it to outside users. At that point, we would be happy to send you a copy.*

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release

I am afraid that the source code was never released. The code was never intended to be released so is not in any shape for general use.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release
- Programmer Left

⟨STUDENT⟩ was a graduate student in our program but he left a while back so I am responding instead. For the paper we used a prototype that included many moving pieces that only ⟨STUDENT⟩ knew how to operate and we did not have the time to integrate them in a ready-to-share implementation before he left. Still, I hope you can build on the ideas/technique of the paper.

Unfortunately, the author who has done most of the coding for this paper has passed away and the code is no longer maintained.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release
- Programmer Left
- Commercial Code

Since this work has been done at <COMPANY> we don't open-source code unless there is a compelling business reason to do so. So unfortunately I don't think we'll be able to share it with you.

The code owned by <COMPANY>, and AFAIK the code is not open-source. Your best bet is to reimplement :(Sorry.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release
- Programmer Left
- Commercial Code
- Proprietary Academic Code

Unfortunately, the `<SYSTEM>` sources are not meant to be opensource (the code is partially property of `(UNIVERSITY 1)`, `(UNIVERSITY 2)` and `(UNIVERSITY 3)`.)

If this will change I will let you know, albeit I do not think there is an intention to make the `<SYSTEM>` sources opensource in the near future.

If you're interested in obtaining the code, we only ask for a description of the research project that the code will be used in (which may lead to some joint research), and we also have a software license agreement that the University would need to sign.

The Dog Ate my Homework !!!

- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release
- Programmer Left
- Commercial Code
- Proprietary Academic Code
- Research vs. Sharing
- ...
- ...

In the past when we attempted to share it, we found ourselves spending more time getting outsiders up to speed than on our own research. So I finally had to establish the policy that we will not provide the source code outside the group.

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

Los Angeles Times | BUSINESS

LOCAL U.S. WORLD BUSINESS SPORTS ENTERTAINMENT HEALTH STYLE TRAVEL

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.

Science AAAS.ORG FEEDBACK HELP LIBRARIANS GUEST ALERTS ACCESS RISER

MAAAS NEWS SCIENCE JOURNALS CAREERS MULTIMEDIA COLLECTIONS

Science The World's Leading Journal of Original Scientific Research, Global News, and Commentary

Science Home Current Issue Previous Issues Science Express Science Products My Science About the Journal

Home > Science Magazine > 17 January 2014 > McNutt, 343 (5168): 229

Article Views Article Summary Full Text PDF Article Tools

Science 17 January 2014: Vol. 343 no. 6168 p. 229 DOI: 10.1126/science.1250475

EDITORIAL Reproducibility Marcia McNutt

Marcia McNutt is Editor-in-Chief of Science.

Science advances on a foundation of trusted data. But the approach that scientists use to gain confidence in their results can lead to irreproducibility. As the science community was shaken by reports that a not all results are reproducible, because confidence in results is based on a lack of transparency, we are announcing new initiatives. Science, for practical purposes, is one of the last remaining bastions of the U.S. scientific enterprise. Increasing transparency. * Authors will indicate handling (such as how to deal with outliers), whether a sufficient signal-to-noise ratio, whether the experimenter was blind to the conduct of the experiments, and other guidelines.

« Prev | Table of Contents | Next »

Read Full Text to Comment (8)

TheScientist EXPLORING LIFE. INSPIRING INNOVATION

NIH Tackles Irreproducibility

The federal agency speaks out about how to improve the quality of scientific research

By Jef Akst | January 28, 2014

nature.com Sitemap Login Register

nature international weekly journal of science Search Advanced search

Home News & Comment Research Careers & Jobs Current Issue Archive

Archive Volume 496 Issue 7446 Editorial Article

NATURE | EDITORIAL

Announcement: Reducing our irreproducibility

24 April 2013

PDF Rights & Permissions

Over the past year, *Nature* has published a string of articles that highlight the reliability and reproducibility of published research (collected at [nature.com/reproducibility](#)).

nature international weekly journal of science

Menu Advanced search Search

archive > volume 483 > issue 7391 > editorials > article

NATURE | EDITORIAL

Must try harder

Nature 483, 509 (29 March 2012) | doi:10.1038/483509a

Published online 28 March 2012

Courtesy V. Stodden, SC, 2015

Too many sloppy mistakes are creeping into scientific papers. Lab heads must look more rigorously at the data — and at themselves.

The Economist Washington's lawyer surplus How to do a nuclear deal with Iran tips from Nobel economists are back King of Sachin Tendulkar

HOW SCIENCE GOES WRONG.

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



Last Week Tonight with John Oliver:
Scientific Studies (HBO), May 2016

Courtesy V. Stodden, SC, 2015

Quick poll

- ① Have you ever tried to reproduce some research results ?

Quick poll

- ① Have you ever tried to reproduce some research results ?
- ② Have you ever failed ?

Quick poll

- ① Have you ever tried to reproduce some research results ?
 - ② Have you ever failed ?

Article typique en traitement d'image

fichier PDF

ce qu'on peut faire avec :



- ✓ lire les formules
 - ✓ croire les résultats
 - ✗ vérifier les résultats
 - ✗ reproduire les résultats
 - ✗ voir les images en détail
 - ✗ voir les graphes en détail

Courtesy of Enric Meinhardt-Llopis, CANUM 2016

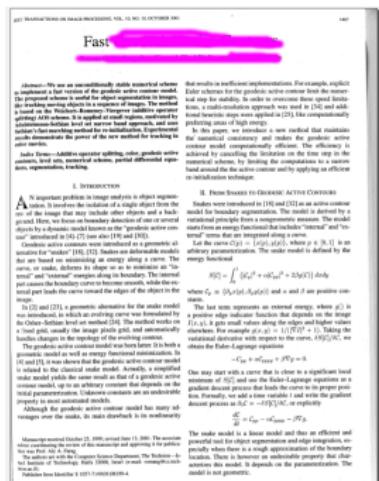
Quick poll

- ① Have you ever tried to reproduce some research results ?
- ② Have you ever failed ?

Article typique en traitement d'image

fichier PDF

ce qu'on peut faire avec :



- ✓ lire les formules
- ✓ croire les résultats
- ✗ vérifier les résultats
- ✗ reproduire les résultats
- ✗ voir les images en détail
- ✗ voir les graphes en détail

Courtesy of Enric Meinhardt-Llopis, CANUM 2016

- ③ Have you ever had trouble reproducing the work of one of your student (or even your own work 😊)?

Reproducibility of experimental results is the hallmark of science



What Descartes did was a good step. You have added much several ways [...] If I have seen further it is by standing on the shoulders of Giants.

– Isaac Newton, February 1676



Science allows to discover truth by building on previous discoveries.

Reproducibility of experimental results is the hallmark of science



What Descartes did was a good step. You have added much several ways [...] If I have seen further it is by standing on the shoulders of Giants.

– **Isaac Newton**, February 1676



In a letter to his rival Robert Hooke

Science allows to discover truth by building on previous discoveries.

1662: **Robert Hooke**, Curator of Experiments for the Royal Society, coins the term *experimentum crucis*.

Only good experiments allow to build sound theories and refute bad ones

Reproducibility of experimental results is the hallmark of science



What Descartes did was a good step. You have added much several ways [...] If I have seen further it is by standing on the shoulders of Giants.
– **Isaac Newton**, February 1676



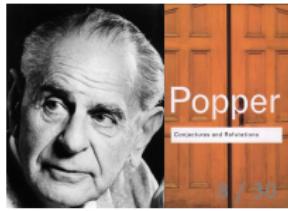
In a letter to his rival Robert Hooke

Science allows to discover truth by building on previous discoveries.

1662: **Robert Hooke**, Curator of Experiments for the Royal Society, coins the term *experimentum crucis*.

Only good experiments allow to build sound theories and refute bad ones

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



A Reproducibility Crisis? (Scientific Misconduct)

Many scandals in social psychology, stem cells, organic transistors (**The Economics of Replication**) leading to retraction/resignation.

- At CNRS, we recently had the Olivier Viannet case and it appears we'll have the sequel with Susana Rivas at INRA 😞

A recent scandal In 2013, **Dong-Pyou Han**, a former assistant professor of biomedical sciences at Iowa State University was disgraced:

- Falsified blood results to make it appear as though a vaccine he was working on had exhibited anti-HIV activity
- Han and his team received $\approx \$19$ million from NIH
- Retraction and resignation of university
- Han was sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials. He was also fined US \$7.2 million!

A Reproducibility Crisis? (Scientific Misconduct)

Many scandals in social psychology, stem cells, organic transistors (**The Economics of Replication**) leading to retraction/resignation.

- At CNRS, we recently had the Olivier Viannet case and it appears we'll have the sequel with Susana Rivas at INRA 😞

A recent scandal In 2013, **Dong-Pyou Han**, a former assistant professor of biomedical sciences at Iowa State University was disgraced:

- Falsified blood results to make it appear as though a vaccine he was working on had exhibited anti-HIV activity
- Han and his team received $\approx \$19$ million from NIH
- Retraction and resignation of university
- Han was sentenced in 2015 to 57 months imprisonment for fabricating and falsifying data in HIV vaccine trials. He was also fined US \$7.2 million!

Scientific fraud is bad but we should avoid witch-hunt

- August 5, 2014, Yoshiki Sasai (stem cell, considered for Nobel Prize) hanged in his laboratory at the RIKEN (Japan). Fraud suspicion...
- More and more alerts but also false alarms...

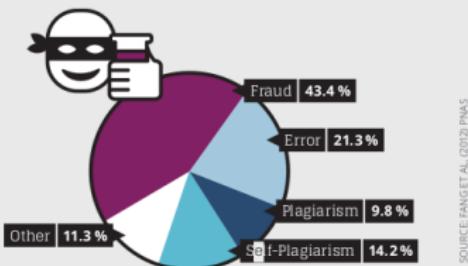
Loosing trust What's the difference between science and charlatanism in people's mind if they get worthless/unreliable information?

Is Fraud a new phenomenon?

The Battle against Scientific Fraud in the CNRS International Magazine

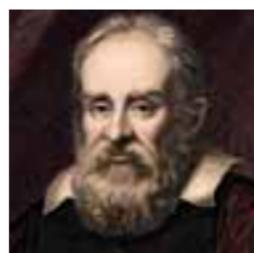
Biomedical fraud in figures

Cause of retraction 1977 to 2012



Number of publications and retractions

1977 to 2013



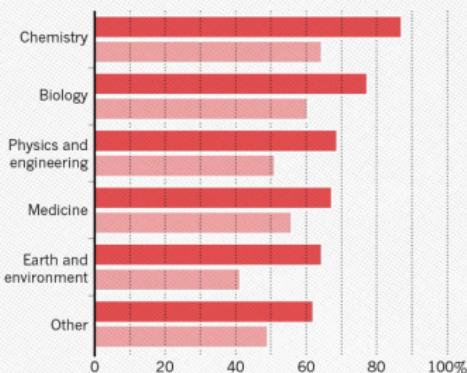
Galileo (data fabrication), **Ptolemy** (plagiarism), **Mendel** (data enhancement), **Pasteur** (rigorous but hided failures), ...

Why are scientific studies so difficult to reproduce?

HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.

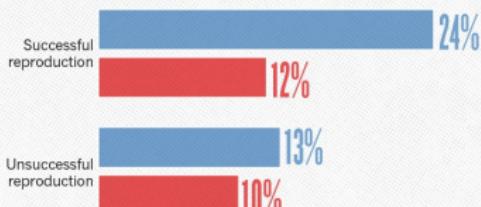
● Someone else's ● My own



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.

● Published ● Failed to publish



Number of respondents from each discipline:
Biology 703, Chemistry 106, Earth and environmental 95.

1,500 scientists lift the lid on reproducibility,
Nature, May 2016

Social causes

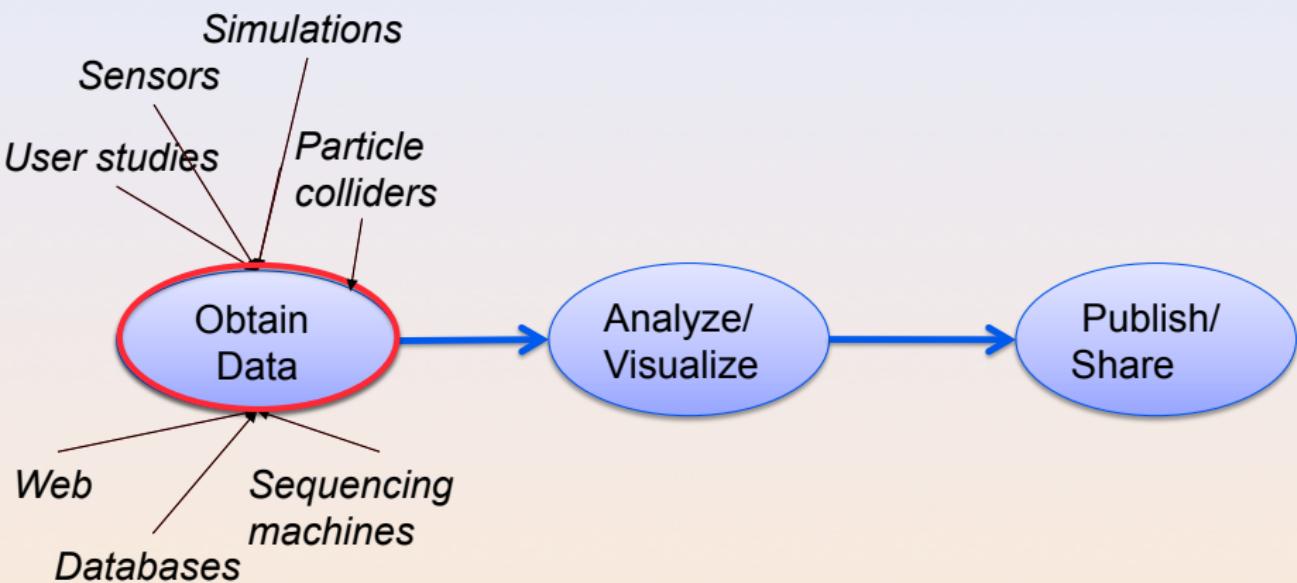
- Fraud, conflict of interest (pharmaceutic, ...)
- No incentive to reproduce/check our own work (afap), nor the work of others (big results!), nor to allow others to check (competition)
- Peer review does not scale: 1+ million articles per year!

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

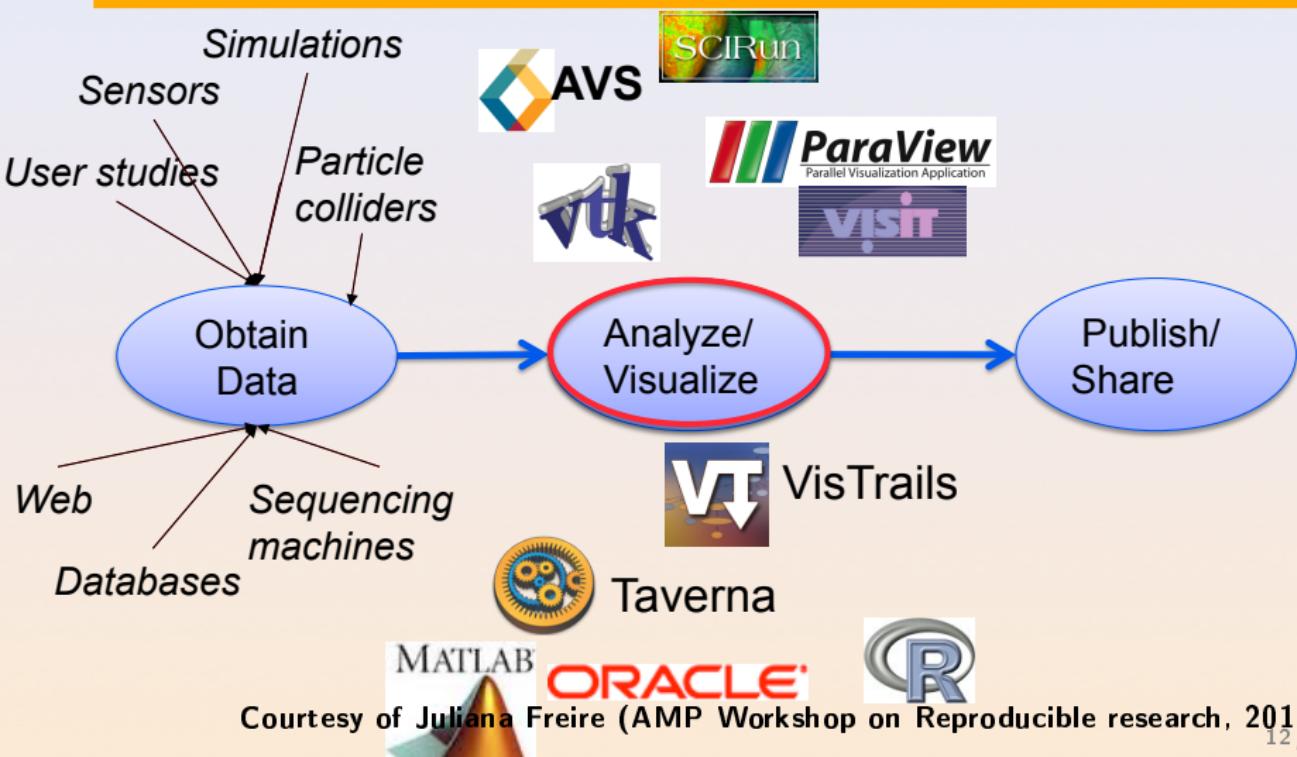
What does Science look like today ?

Science Today: Data Intensive



What does Science look like today ?

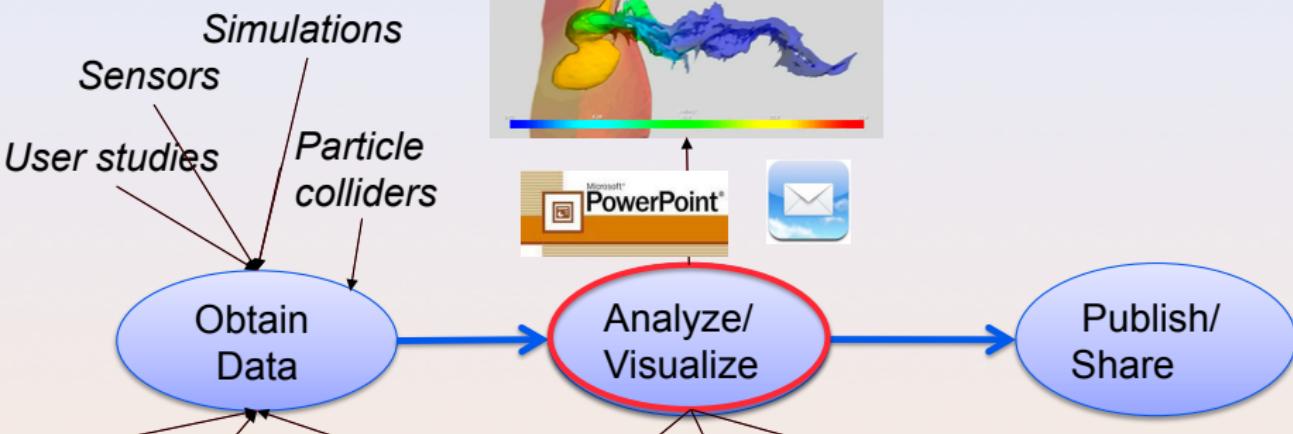
Science Today: Data + Computing Intensive



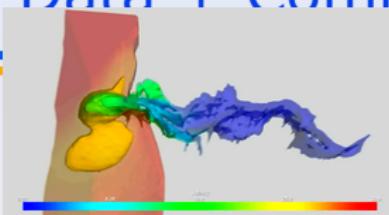
Courtesy of Juliana Freire (AMP Workshop on Reproducible research, 2011)

What does Science look like today ?

Science Today: Data + Computing Intensive



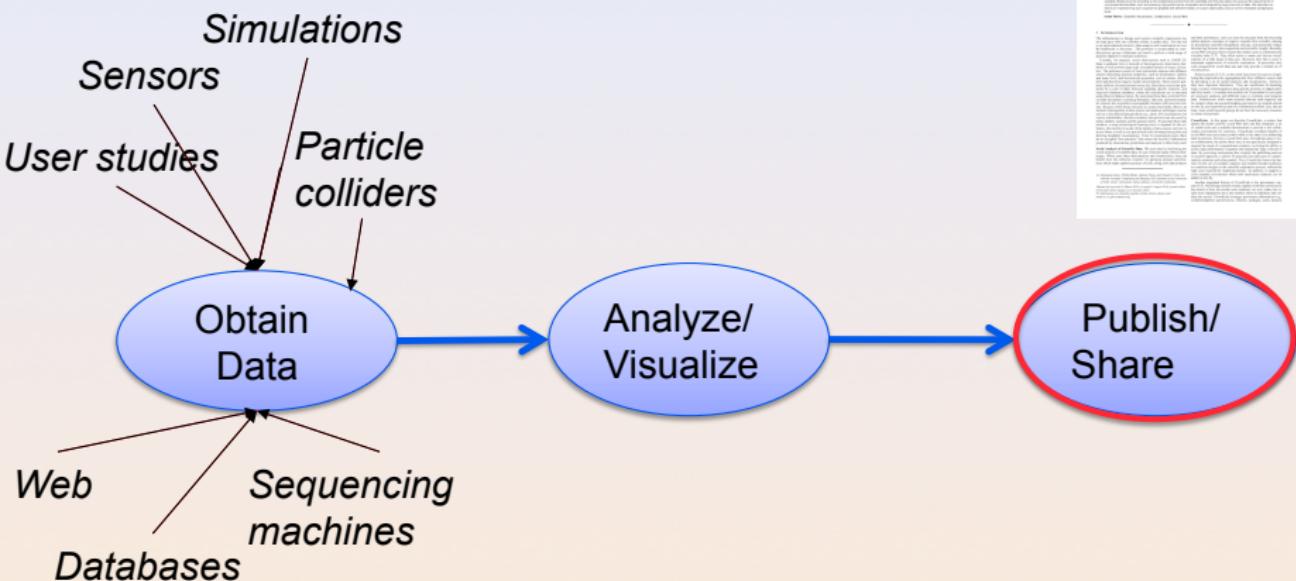
Courtesy of Juliana Freire (ANL) | Reproducible research, 2011 | 12 / 30



CANADA	CHILE	CHINA	COLOMBIA
CZECHOSLOVAKIA	DENMARK	FRANCE	GERMANY
FINLAND	FRANCE	GERMANY	GERMANY, WEST
INDIA	IRAN	JAPAN	KOREA
MEXICO	NETHERLANDS	NEW ZEALAND	NIGERIA
HUNGARY			

What does Science look like today ?

Science Today: Data + Computing Inte



What does Science look like today ?

Science Today: Incomplete Publications

- ◆ Publications are just the tip of the iceberg
 - Scientific record is incomplete---to large to fit in a paper
 - Large volumes of data
 - Complex processes
- ◆ Can't (easily) reproduce results

Software is an essential component of modern scientific research Top 100 papers (Nature, 2014)

[...] the vast majority describe experimental methods or software that have become essential in their fields.

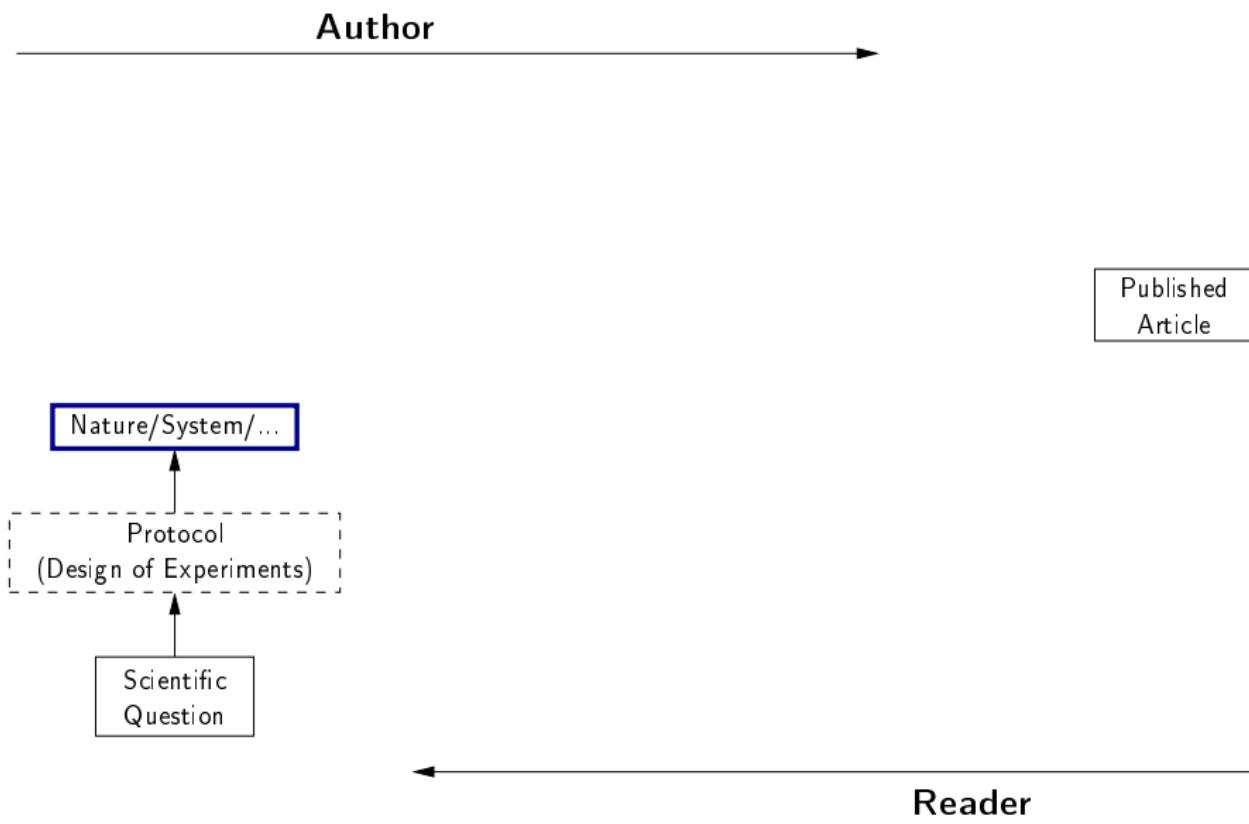


Courtesy of Juliana Freire (AMP Workshop on Reproducible research, 2011)

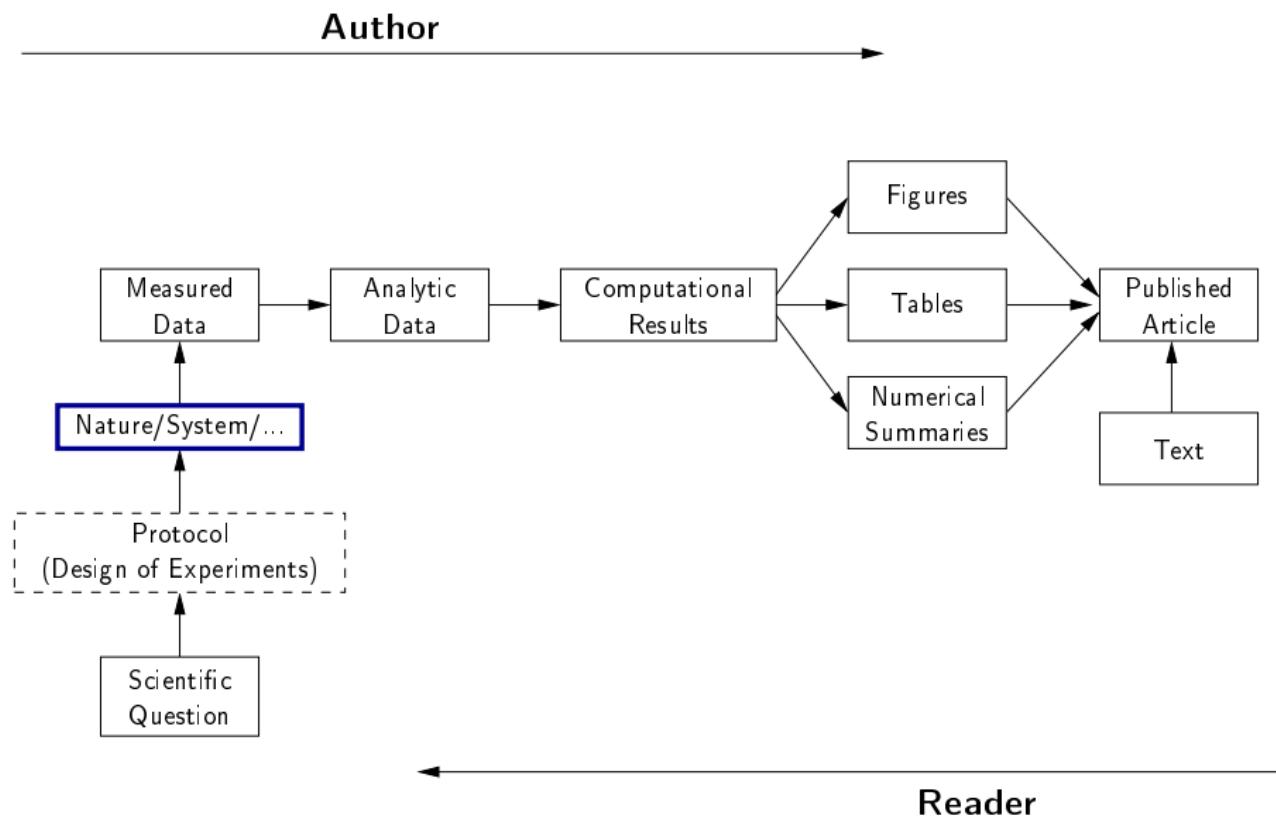
Outline

- ① A Few Motivating Examples
- ② Science crisis ?
- ③ Reproducible Research/Open Science in a Nutshell
- ④ Illustrating Nice Ideas Through Different Tools
- ⑤ What can Computer Scientists do ?

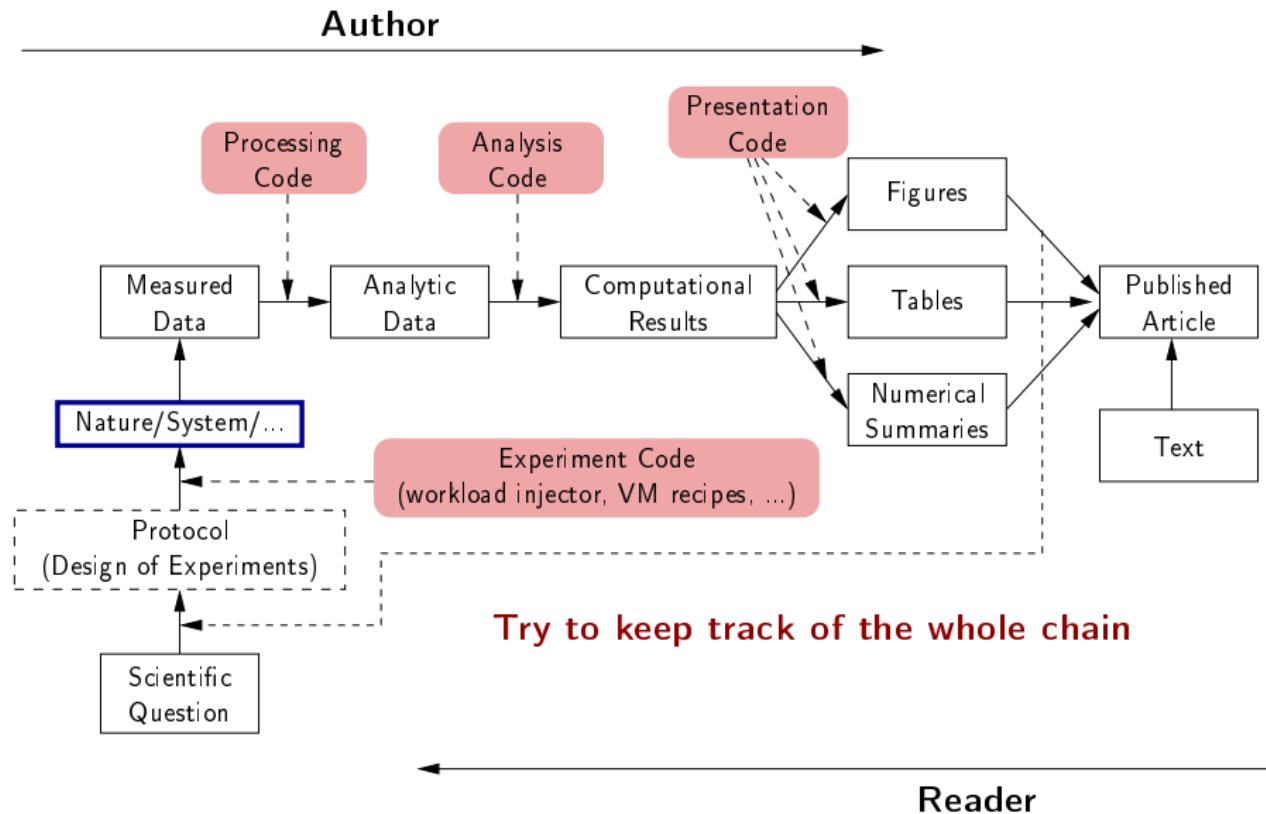
Reproducible Research: Trying to Bridge the Gap



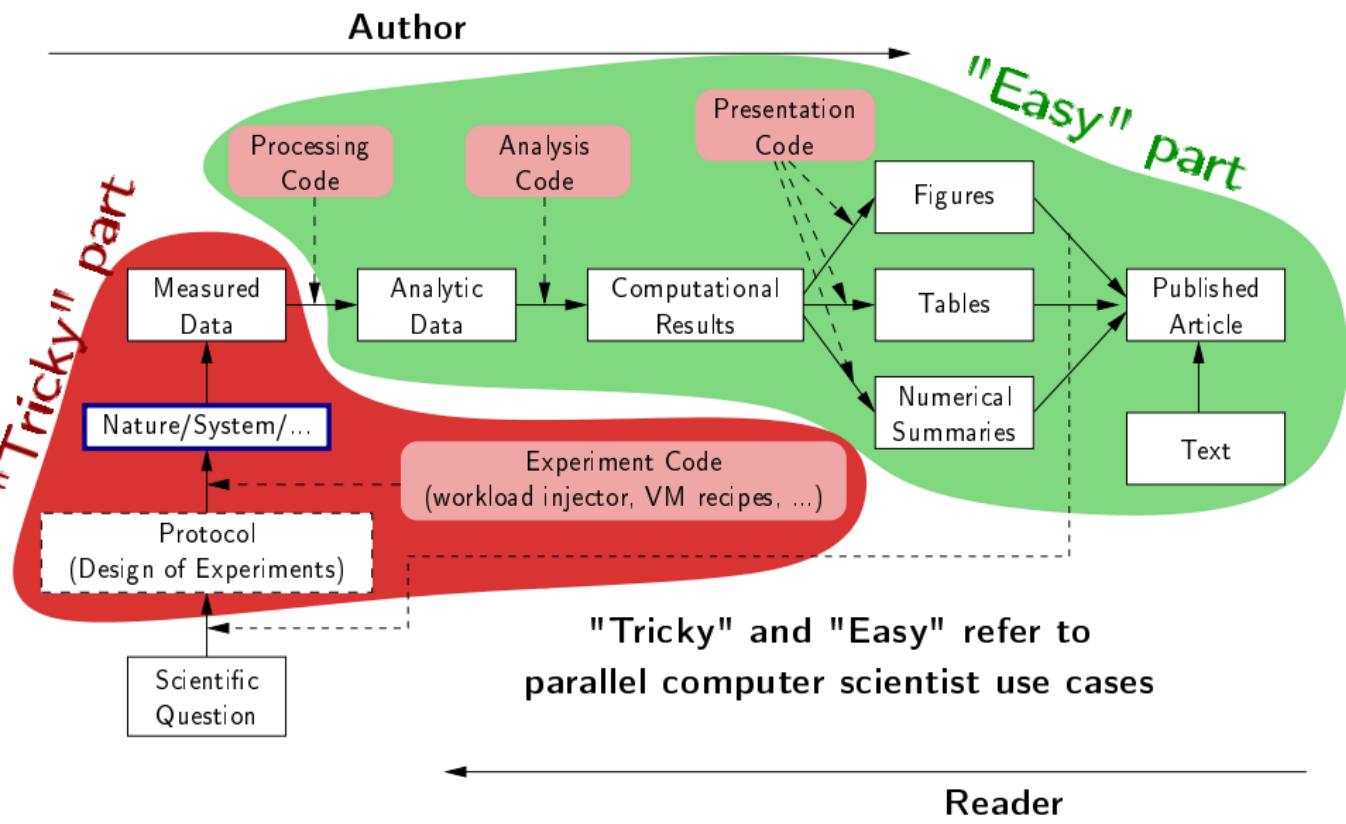
Reproducible Research: Trying to Bridge the Gap



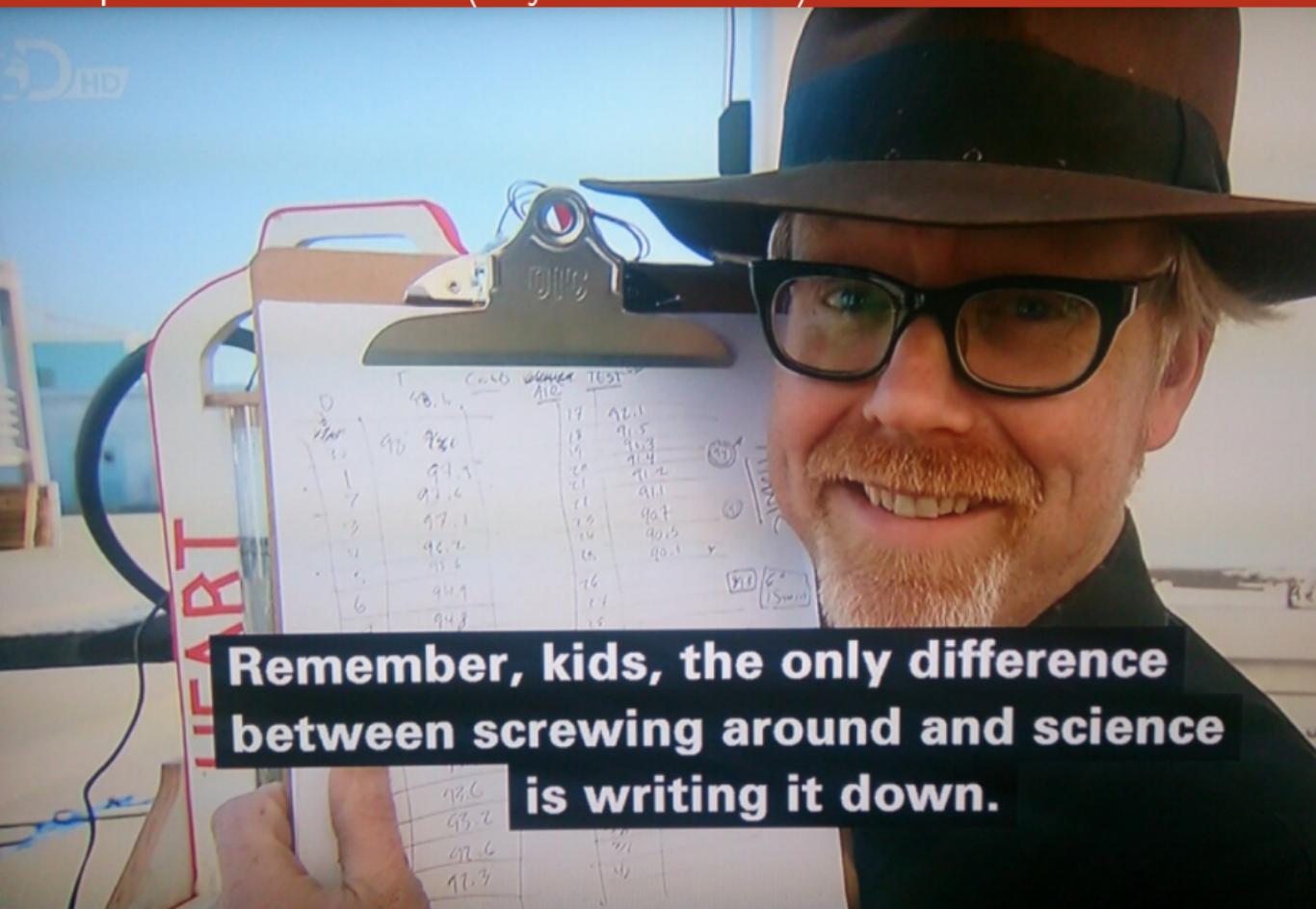
Reproducible Research: Trying to Bridge the Gap



Reproducible Research: Trying to Bridge the Gap



Step 1: Take notes! (Mythbusters 😊)



**Remember, kids, the only difference
between screwing around and science
is writing it down.**

Step 2: Control what's happening

Both the **scientific process** and the **software environment** have become extremely complex

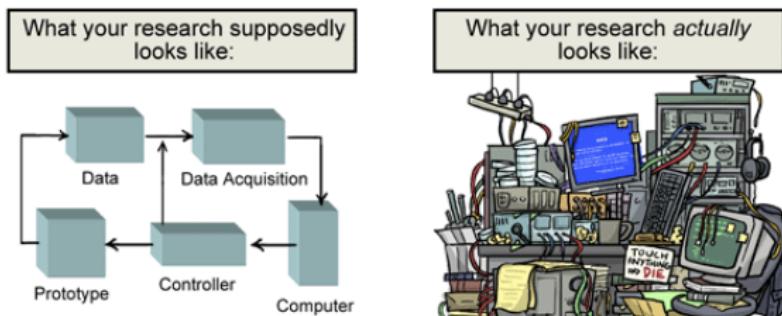


Figure 1. Experimental Diagram

Figure 2. Experimental Mess

Computational biologists and physicist have pushed forward with the development of many tools:

- Software and environment preservation/reconstruction
- Workflow engines, data provenance
- Dissemination platforms, data/workload archive
- Laboratory notebooks, literate programming, replicable articles

And this landscape is rapidly evolving

Step 3: Reproducibility? What Are We Talking About Exactly?

Replicability

Reproducibility

Reproduction of the original results using the same tools

by the original author on the same machine

by someone in the same lab/using a different machine

by someone in a different lab

Reproduction using different software, but with access to the original code

Completely independent reproduction based only on text description, without access to the original code

Note that the **terminology varies** (repetition, replication, reproduction, reuse, corroboration, ...)

And in practice: remember the value of data

Ten Simple Rules for Reproducible Computational Research, PlosOne

- ① For Every Result, Keep Track of How It Was Produced
- ② Avoid Manual Data Manipulation Steps (and Avoid Spreadsheets! 😊)
- ③ Archive the Exact Versions of All External (Open-Source! 😊) Programs Used
- ④ Version Control All Custom Scripts
- ⑤ Record All Intermediate Results, When Possible in Standardized Formats
- ⑥ For Analyses That Include Randomness, Note Underlying Random Seeds
- ⑦ Always Store Raw Data behind Plots
- ⑧ Generate Hierarchical Analysis Output, Allowing Layers of Increasing Detail to Be Inspected
- ⑨ Connect Textual Statements to Underlying Results
- ⑩ Provide Public Access to Scripts, Runs, and Results

But **several ways to follow them**

- **Preserve the Mess**

Automatically keeping track of everything

- **Encourage Cleanliness**

Ensure others can understand/adapt what was done

Outline

- ① A Few Motivating Examples
- ② Science crisis ?
- ③ Reproducible Research/Open Science in a Nutshell
- ④ Illustrating Nice Ideas Through Different Tools
- ⑤ What can Computer Scientists do ?

Vistrails: a Workflow Engine for Provenance Tracking

An *Provenance-Rich* Paper: ALPS2.0

The ALPS project release 2.0:
Open source software for strongly correlated
systems

B. Bauer¹ L. D. Carr² H.G. Evertz³ A. Feiguin⁴ J. Freire⁵
S. Fuchs⁶ L. Gamper¹ J. Gukelberger¹ E. Gulf⁷ S. Guertler⁸
A. Hehn¹ R. Igarashi^{9,10} S.V. Isakov¹ D. Koop⁵ P.N. Ma¹
P. Mates^{1,5} H. Matsuno¹¹ O. Parcollet¹² G. Pawłowski¹³
J.D. Picon¹⁴ L. Pollet^{1,12} E. Santos⁹ V.W. Scarola¹⁵
U. Schollwöck¹⁷ C. Silva⁵ B. Sturer¹ S. Todo^{10,11} S. Trebst¹⁸
M. Troyer¹ M. L. Wall¹⁹ P. Werner¹ S. Wessel^{19,20}

¹Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland

²Department of Physics, Colorado School of Mines, Golden, CO 80401, USA

³Institut für Theoretische Physik, Technische Universität Graz, A-8010 Graz, Austria

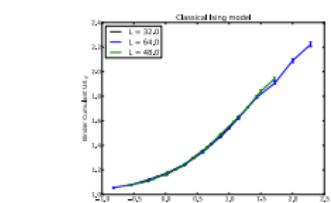
⁴Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA

⁵Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah 84112, USA

⁶Institut für Theoretische Physik, Georg-August-Universität Göttingen, Göttingen, Germany

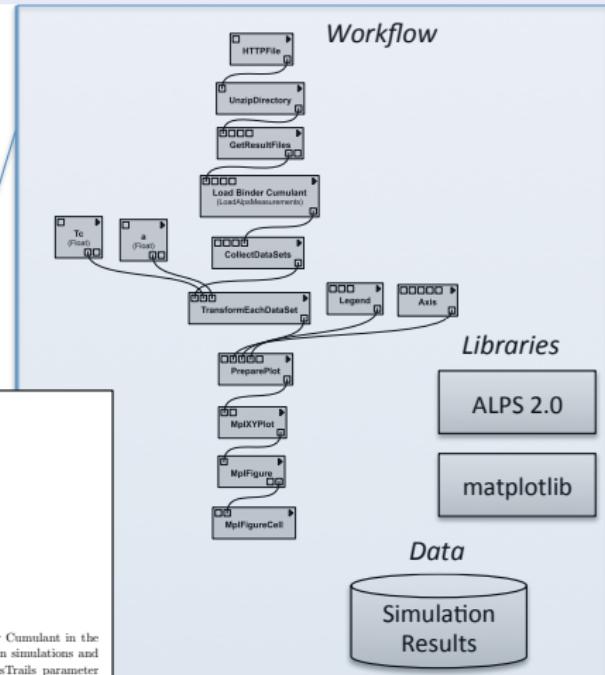
⁷Columbia University, New York, NY 10027, USA

⁸Bethe Center for Theoretical Physics, Universität Bonn, Nussallee 12, 53115 Bonn, Germany



¹ Correspondence

Figure 3. In this example we show a data collapse of the Binder Cumulant in the classical Ising model. The data has been produced by remotely run simulations and the critical exponent has been obtained with the help of the VisTrails parameter exploration functionality.



Courtesy of Juliana Freire (AMP Workshop on Reproducible research) 10

VCR: A Universal Identifier for Computational Results

Chronicling computations in real-time

VCR computation platform Plugin = Computation recorder

Regular program code

```
figure1 = plot(x)
save(figure1,'figure1.eps')
```

```
> file /home/figure1.eps saved
>
```

VCR: A Universal Identifier for Computational Results

Chronicling computations in real-time

VCR computation platform Plugin = Computation recorder

Program code with VCR plugin

```
repository vcr.nature.com  
verifiable figure1 = plot(x)
```

```
> vcr.nature.com approved:
```

```
> access figure1 at https://vcr.nature.com/ffaaffb148d7
```

VCR: A Universal Identifier for Computational Results

Word-processor plugin App

LaTeX source

```
\includegraphics{figure1.eps}
```

LaTeX source with VCR package

```
\includeresult{vcr.thelancet.com/ffaaffb148d7}
```

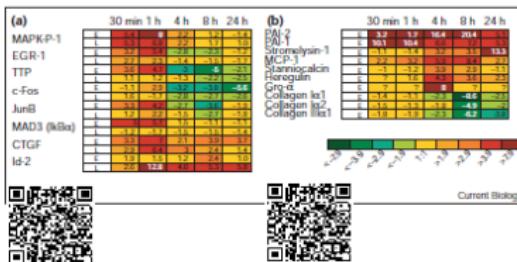
Permanently bind printed graphics to underlying result content

VCR: A Universal Identifier for Computational Results

Research Paper Analysis of replicative senescence Shelton et al. 943

Figure 3

Time course of serum stimulation. (a) Early passage (E; PD30) or late passage (L; PD89) BJ cultures were held in 0.5% serum for 2 days, then stimulated with 10% FBS. RNA levels from cultures at the indicated time points (Cy5 channel) were compared with the uninduced starting culture (Cy3 channel). Positive values indicate higher expression in induced cells; negative values indicate lower expression in induced cells. Question marks indicate that there was insufficient signal for detection. A complete listing of serum-responsive genes from this analysis is provided in Supplementary material. (b) The serum-responsiveness of select senescence-regulated genes in early passage (PD30) BJ fibroblasts.



senescence response appears to overlap substantially with gene expression patterns observed in activated fibroblasts during wound healing [24–26]. MCP-1, Gro- α , IL-1 β and IL-15 are strong effectors of macrophage and neutrophil recruitment and activation [27,28]. The upregulation of Toll (Tlr-4) in senescent fibroblasts confirms the overall immune response behavior at senescence. Tlr-4 is an IL-1 receptor homolog and is implicated in the activation of the gene regulatory protein NF- κ B, a function proposed to be part of the innate immune response [29]. The induction of IL-15 at senescence is also consistent with an innate immune response, as IL-15 can be induced by NF- κ B-dependent transcription [30] and also participates in inflammatory disease processes [28].

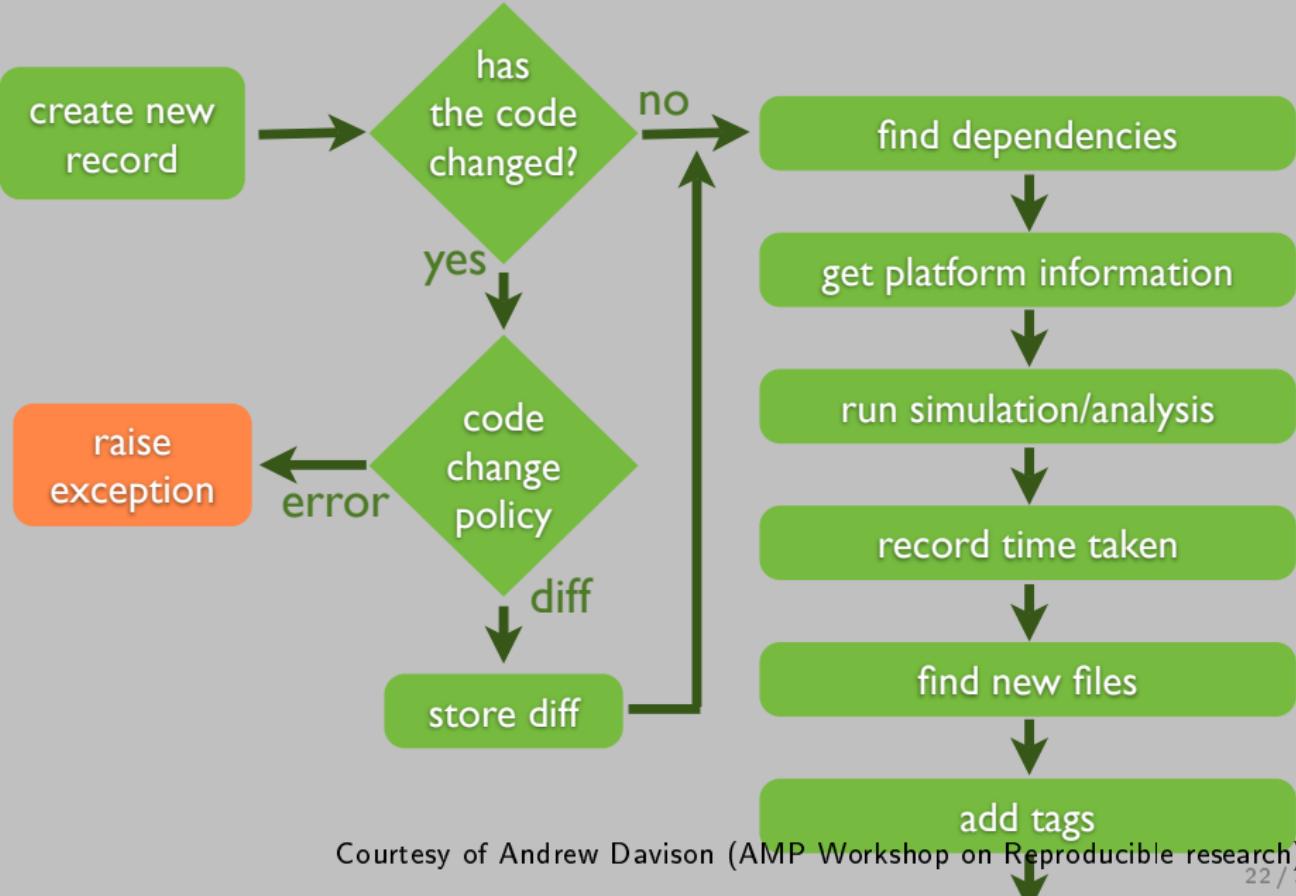
Deficiencies in the response of senescent cells to serum stimulation have been reported, and include an inability to induce the expression of *c-fos* mRNA [31] and markers of late G1 and S phase [32]. In response to serum, expression of inflammatory chemokines, matrix-degrading proteases and their modulators is induced in early-passage dermal fibroblasts, and expression of matrix collagens is reduced. This transient burst of activity may represent a natural progression of events in early-passage cells. Transcripts were hyper-induced in serum-stimulated senescent cells, and included markers of the immune system, such as TLR-4, IL-15, and cathepsin-1, and markers of

states overlap substantially with those in telomere-induced senescence (W.F., D.N.S., R. Allsopp, S. Lowe, and G. Ferbeyre, unpublished observations) and thus are likely to use many of the same activation processes.

The pattern of gene expression at senescence varies substantially in different cell types. Although the expression matrix and structural proteins, such as the collagens, keratins and auxiliary factors, is repressed in RPE cells, inflammatory regulators are not induced, in contrast to dermal fibroblasts. Physiologically, this would make sense, as an acute inflammatory response in a tissue critical for normal vision would be likely to have deleterious consequences. However, as the RPE layer has a central role in the deposition and maintenance of extracellular matrix in the retina, decrements in the ability of senescent RPE cells to maintain appropriate expression patterns, as evidenced by decreased expression of collagens, keratins, aggrecan, transglutaminase and so on, would be predicted to have adverse effects on retinal architecture. Dysfunction of the RPE cell layer is considered to be a substantial factor in the development of age-related macular degeneration [36].

surprisingly, early-passage cells also overexpress many of the markers associated with senescence in dermal fibroblasts, including TLR-4, IL-15, cathepsin-1, and

Sumatra: an "experiment engine" that helps taking notes



Courtesy of Andrew Davison (AMP Workshop on Reproducible research)

Sumatra: an "experiment engine" that helps taking notes

```
$ smt comment 20110713-174949 "Eureka! Nobel prize  
here we come."
```

Sumatra: an "experiment engine" that helps taking notes

```
$ smt tag "Figure 6"
```

Sumatra: an "experiment engine" that helps taking notes

Sumatra: TestProject: List of records

TestProject: List of records

Delete Include data	Label	Reason	Outcome	Duration	Processes	Simulator		Script			Date	Time	Tags
						Name	Version	Repository	Main file	Version			
<input type="checkbox"/>	20100709-154255		'Eureka! Nobel prize here we come.'	0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:42:55	
<input type="checkbox"/>	20100709-154309			0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:09	
<input type="checkbox"/>	haggling	'determine whether the gourd is worth 3 or 4 shekels'	'apparently, it is worth NaN shekels.'	0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:20	foobar
<input type="checkbox"/>	20100709-154338	'test effect of a smaller time constant'		0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:38	
<input type="checkbox"/>	haggling_repeat	Repeat experiment haggling	The new record exactly matches the original.	0.58 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:47	

Ipython/Jupyter Notebook

Web app: create and share documents that contain live code, equations, visualizations, and explanatory text

The image shows two side-by-side screenshots of the Jupyter Notebook interface. The left screenshot shows the main dashboard with sections for 'Welcome to the Notebook', 'Run some Python code', and a code cell starting with 'In []: %matplotlib inline'. The right screenshot shows a specific notebook titled 'Lorenz Differential Equations' with the following content:

Exploring the Lorenz System

In this Notebook we explore the [Lorenz system](#) of differential equations:

$$\begin{aligned}\dot{x} &= \sigma(y - x) \\ \dot{y} &= \rho x - y - xz \\ \dot{z} &= -\beta z + xy\end{aligned}$$

This is one of the classic systems in non-linear differential equations. It exhibits a range of complex behaviors as the parameters (σ , β , ρ) are varied, including what are known as chaotic solutions. The system was originally developed as a simplified mathematical model for atmospheric convection in 1963.

In [7]: `interact(Lorenz, N=fixed(10), angle=(0.,360.),
σ=(0.0,50.0),β=(0.,5), ρ=(0.0,50.0));`

Below this is an interactive slider interface with four sliders:

- angle: 308.2
- max_time: 12
- σ: 10
- β: 2.6
- ρ: 28

At the bottom is a colorful 3D plot of the Lorenz attractor, showing its characteristic butterfly shape.

Reprozip

Automagically pack your experiment to fight **dependency hell**

ON THE ORIGINAL MACHINE

```
$ pip install reprozip
$ reprozip trace ./myexperiment --options inputs/somefile.csv other_file_here.bin
experiment: 0%... 25%... 50%... 75%... 100%
result: 42.137
Configuration file written in .reprozip/config.yml
Edit that file then run the packer -- use 'reprozip pack -h' for help
$ reprozip pack my_experiment.rpz
[REPROZIP] 17:26:42.588 INFO: Creating pack my_experiment.rpz...
[REPROZIP] 17:26:42.589 INFO: Adding files from package coreutils...
[REPROZIP] 17:26:42.601 INFO: Adding files from package libc6...
[REPROZIP] 17:26:42.906 INFO: Adding other files...
[REPROZIP] 17:26:43.450 INFO: Adding metadata...
```

ON ANOTHER MACHINE

```
$ pip install reprounzip[all]
$ reprounzip vagrant setup my_experiment.rpz mydirectory
Bringing machine 'default' up with 'virtualbox' provider...
==> default: Importing base box 'remram/debian-7-amd64'...
==> default: Booting VM...
==> default: Machine booted and ready!
==> default: Running provisioner: shell...
$ reprounzip vagrant run mydirectory
experiment: 0%... 25%... 50%... 75%... 100%
result: 42.137
$ reprounzip vagrant upload /tmp/new_config:global-config
$ reprounzip vagrant run mydirectory --cmdline ./myexperiment --other --options
inputs/somefile.csv
experiment: 0%... 25%... 50%... 75%... 100%
result: -17.814
```

So many new tools

New Tools for Computational Reproducibility

- Dissemination Platforms:

[ResearchCompendia.org](#)

[IPOL](#)

[Madagascar](#)

[MLOSS.org](#)

[thedatahub.org](#)

[nanoHUB.org](#)

[Open Science Framework](#)

[The DataVerse Network](#)

[RunMyCode.org](#)

- Workflow Tracking and Research Environments:

[VisTrails](#)

[Kepler](#)

[CDE](#)

[Galaxy](#)

[GenePattern](#)

[Synapse](#)

[Sumatra](#)

[Taverna](#)

[Pegasus](#)

- Embedded Publishing: Courtesy of Victoria Stodden (UC Davis, Feb 13, 2014)

[Verifiable Computational Research](#) [Sweave](#) [knitR](#)

[Collage Authoring Environment](#) [SHARE](#)

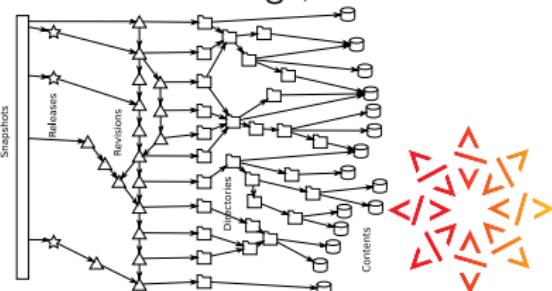
And also: Org-Mode 😊, ActivePapers, Figshare, Zenodo, Elsevier executable paper 😞, Code Ocean, ...

Outline

- ① A Few Motivating Examples
- ② Science crisis ?
- ③ Reproducible Research/Open Science in a Nutshell
- ④ Illustrating Nice Ideas Through Different Tools
- ⑤ What can Computer Scientists do ?

On the "technical" side (1/2)

- Better documenting what we do: **Laboratory notebooks**
 - Literate programming is great for analysis, and reproducible articles but does not go well yet with conducting experiments and workflows
 - A real adoption of such practice requires more storage and the ability to navigate in such information
- Better software engineering practice: Public releases, **devops approach**, reproducible builds, numerical stability
 - Moving/evolving technology. Preservation ? Adoption ?
 - Should not slow down research
- Fighting against software/data degradation: **Software Heritage**, zenodo
 - Challenges: multiple! curation, access/privacy, exploitation, navigation, storage, ...



Software Heritage
THE GREAT LIBRARY OF SOURCE CODE

On the "technical" side (2/2)

- Better experimental practice and platforms: FIT IoT-lab, G5K are world leading experimental infrastructures; rely on standard simulators (**SimGrid, NS3**)
 - Maintenance cost, keeping in pace with technology, practices for prototype platforms, control, sharing of experimental conditions with others, experimental engines



- Need for convergence in term of software infrastructure and practice (e.g., security, account management, access, isolation, experiment management, etc.) ?

On the "social" side

Slight **cultural changes** in our relation to publication and daily practice

- **Changing our social model to favor adoption of better practice**
 - Artifact evaluation, open reviews, ... (e.g., IPOL, ReScience)
 - Promote a different model
- **Learning** is the essence of our work. \rightsquigarrow Train our researchers and students
 - Better teaching/understanding of statistics, experimental practice, design of experiments

It's up to us. We should care and take the lead

Main benefit:

- **Higher confidence** in our work \rightsquigarrow definite **competitive advantage**
- Our research becomes **sound, deeper, auditable, more visible, reusable,**
...

Webinars: Learning by Doing

Many different tools/approaches developed in various communities

- 1 Replicable article
- 2 Logging your activity
- 3 Logging and backing up your data
- 4 Organizing your data
- 5 Mastering your environment
- 6 Controlling your experiments
- 7 Making your data/code/article available
- 8 Publication modes
- 9 Artifact Evaluation

Reproducible Research, Open Science
Motivation, Challenges, Approaches, ...

Amaud Legrand
CNRS, Inria, University of Grenoble

March 7, 2016 – Reproducible Research Webinar

People involved in preparing this talk:

- Michael Mercier (Inria/Atom)
- Cristian Ruiz (Inria)
- Grid5000, Kameleon, Expa, ...

Thanks for the feedback of:

- Pierre Neyron (CNRS)
- Arnaud Legrand (CNRS)
- Olivier Richard (UGA)
- Lucas Nussbaum (Loria)

Here is the pad for interactions: <http://tinyurl.com/RRW-pad2>

Material (demo, slides) available on [github](#):

- A Docker Demo
- A complete use case

Literate programming

Controlling your environment



Reproducible Research, Open Science
Logging and backing up your work
Git Tips and Tricks, a Scientist Perspective

V. Danjean, A. Legrand, L. Stanisic
University of Grenoble, CNRS, Inria Bordeaux

June 7, 2016 – Reproducible Research Webinar (Episode IV)

Numerical reproducibility

Logging and backing up

https://github.com/alegrand/RR_webinars