

Reproducible Research

Vinícius Garcia Pinto
Arnaud Legrand



Porto Alegre - March 23, 2017

Outline

- ① A Few Motivating Examples
- ② What about Computer Science ?
- ③ Reproducible Research/Open Science: Illustrating Nice Ideas Through Different Tools

Inconsistencies

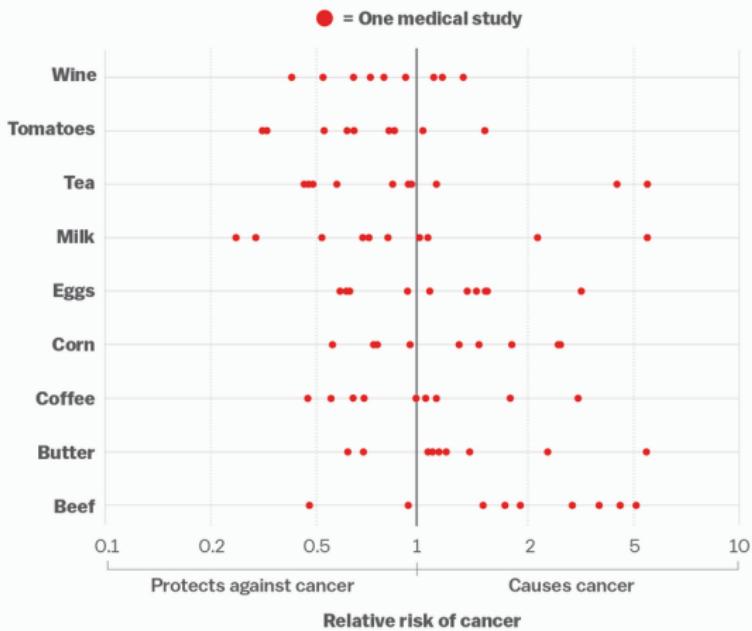
What should we eat then?



Is everything we eat associated with cancer? A systematic cookbook review, Schoenfeld and Ioannidis, *Amer. Jour. of Clinical Nutrition*, 2013.

Inconsistencies

What should we eat? Everything we eat both causes and prevents cancer



SOURCE: Schoenfeld and Ioannidis, *American Journal of Clinical Nutrition*

Vox

Is everything we eat associated with cancer? A systematic cookbook review, Schoenfeld and Ioannidis, *Amer. Jour. of Clinical Nutrition*, 2013.

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

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 All Science Journals Editor Search View

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

Science 17 January 2014 • Vol. 343 no. 6158 p. 229 DOI: 10.1126/science.1250475

Article Views Summary Full Text Full Text (PDF) Article Tools

Editorial Marcia McNutt

Science advances on a foundation of trusted data approach that scientists used to gain confidence community was shaken by reports that a result not reproducible. Because confidence in result community, we are announcing new initiatives Science. For preclinical studies one of the standards commercialization of U.S. results. Increasing transparency. * Authors will indicate handling (such as how to deal with outliers), whether ensure a sufficient signal-to-noise ratio, whether experimenter was blind to the conduct of the guidelines.

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

Announcement: Reducing our irreproducibility : Nature News & Comment

www.nature.com/nature-reducing-irreproducibility/

nature International weekly journal of science

Home News & Comment Research Careers & Jobs Current Issue Archive

Audio & Video For Authors

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 reliability and reproducibility of published research (collected and free

The Economist

How SCIENCE GOES WRONG

nature International weekly journal of science

Menu 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

PDF Citation Reprints Rights & permissions Article metrics

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

Courtesy V. Stodden, SC, 2015

Public evidence for a Lack of Reproducibility - Brazil

- Os alquimistas (Piauí), Sept. 2011; Pela 1a vez, Fapesp ... (O Estado de S. Paulo), Oct. 2014; Reincidencia (Piauí), Sept. 2015.
- Facebook group ("Bolsistas *****")

piauí NESTE MÊS | FESTIVAL | BLOGS | THE PIAUÍ HERALD | PERFIS | REPORTAGENS | ENSAIOS | HISTÓRIA PESSOAL | HUMOR | FICÇÃO E POESIA | ESQUINAS | ASSINE | C

EDIÇÃO 80 | SETEMBRO DE 2011 _anais da ciência

OS ALQUIMISTAS

Químicos acusados de forjar resultados de onze estudos colocam o Brasil no mapa da fraude científica mundial

= ESTADÃO CIÊNCIA Pela 1ª vez, Fapesp torna públicas fraudes científicas

Pela 1ª vez, Fapesp torna públicas fraudes científicas

Cinco pesquisadores que receberam financiamento são acusados de má conduta e teriam os nomes divulgados no site da fundação

piauí
NESTE MÊS | FESTIVAL | BLOGS | THE PIAUÍ HERALD | PERFIS |

Questões da Ciência

REINCIDÊNCIA

Quatro anos depois de revelado o maior caso de fraude já documentado envolvendo cientistas brasileiros, vêm à tona novas acusações de má conduta dirigidas a químicos do país, inclusive um professor que

Public evidence for a Lack of Reproducibility - Brazil

- Os alquimistas (Piauí), Sept. 2011; Pela 1a vez, Fapesp ... (O Estado de S. Paulo), Oct. 2014; Reincidencia (Piauí), Sept. 2015.
- Facebook group ("Bolsistas *****")

The screenshot shows a Facebook group conversation with several posts and replies. The posts are in Portuguese and discuss various issues related to data handling and reproducibility.

Post 1: Gente... como faz quando o teu orientador cobra uns dados e você não acha ? que desculpa vcs usam? me ajuda.

Post 2: "Meu notebook não tá ligando..."

Post 3: Meu cachorro comeu tá muito old né?

Post 4: sim... ele sabe que não tenho cachorro aqui 😊

Post 5: Pensai logo no cachorro... rsrs

Post 6: Que tipo de dados? Dado brutos? Dados analisados?

Post 7: os dois... coloquei os dois no mesmo arquivo 😊

Post 8: Vc não anotou num caderninho?

Post 9: Se ele quer os dados ele q de o método

Post 10: Diz q o HD queimou, tá mais moderno!!

Post 11: Se você não tiver como recuperar os dados, você precisa falar a verdade, porque quanto mais tempo você perder tentando enrolar seu orientador, menos tempo você vai ter pra replanejar o trabalho e refazer o experimento. Merdas acontecem. 😊 Boa sorte!

Reply 1: Pergunta se ele não tem referências para os dados que você precisa!

Reply 2: os dados são do meu experimento 😊

Reply 3: Quem hoje em dia não tem backup? Acho que se tivesse perdido mesmo não precisaria de desculpa. Tô achando que isso é calote rs

Reply 4: Interpola ou extrapola...

Reply 5: Usa os dados que vc tem para treinar uma rede neural para gerar os dados no ponto que vc precisa...Mas acho que daria outro mestrado fazer isso...

Reply 6: O loko meu!! Mas gostei da idéia!

Reply 7: Qual a metodologia para chegar nesses dados ????

Image: A photograph of a man sitting at a desk, looking stressed with his hand to his forehead while writing in a notebook.

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 ?



- ✓ read the formulas
- ✓ rely on results
- ✗ check the results
- ✗ reproduce the results
- ✗ see images in details
- ✗ see graphics in details

(Translated) Courtesy of Enric Meinhardt-Llopis, CANUM 2016

Quick poll

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



- ✓ read the formulas
- ✓ rely on results
- ✗ check the results
- ✗ reproduce the results
- ✗ see images in details
- ✗ see graphics in details

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

(Translated) Courtesy of Enric Meinhardt-Llopis, CANUM 2016

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?

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?

Outline

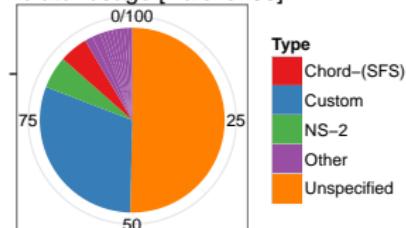
- ① A Few Motivating Examples
- ② What about Computer Science ?
- ③ Reproducible Research/Open Science: Illustrating Nice Ideas Through Different Tools

A Few Edifying Examples

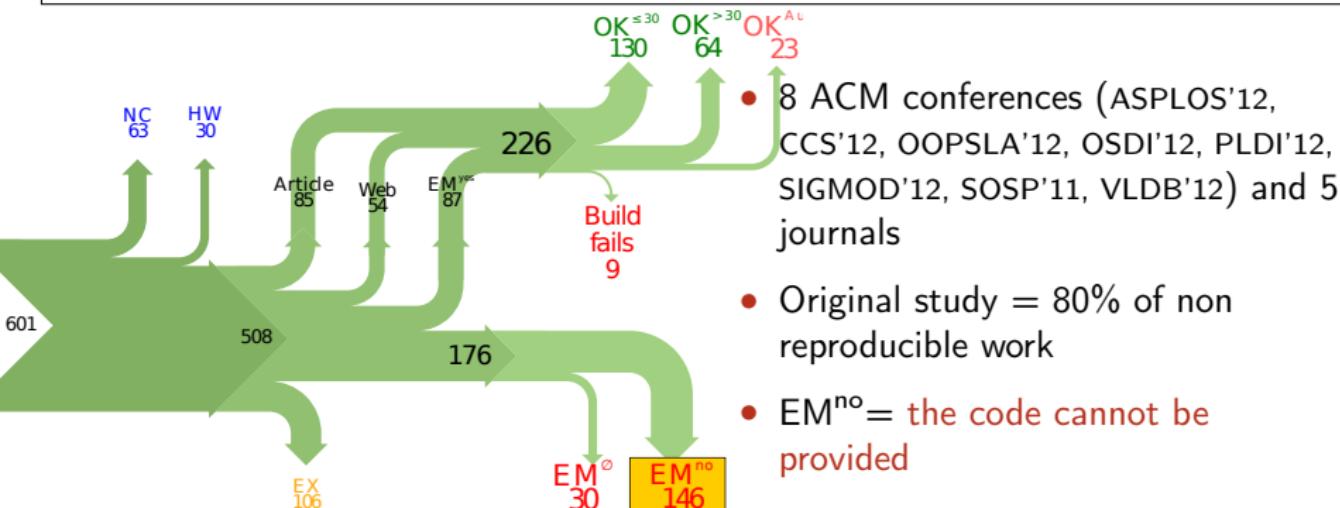
Naicken, Stephen et Al., *Towards Yet Another Peer-to-Peer Simulator*, HET-NETs'06.

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

Simulator usage [Naicken06]



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 !!!

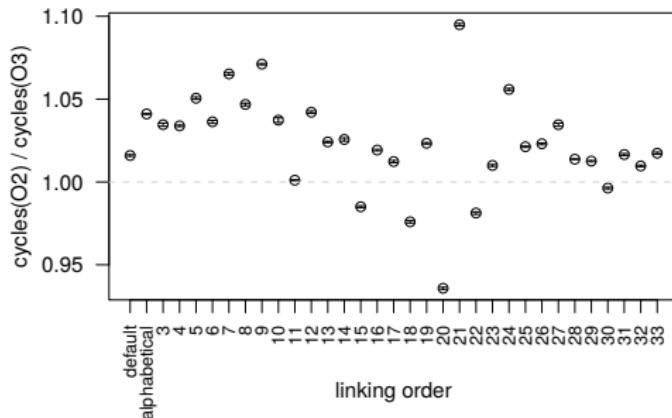
- Versioning Problems
- Bad Backup Practices
- Code Will be Available Soon
- No Intention to Release
- Programmer Left
- Commercial Code
- Proprietary Academic Code
- Research vs. Sharing
- ...
- ...

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.

But do we **really** have to care in CS?

Yes. Model \neq Reality. Although designed and built by human beings, computers are **so complex** that mistakes are easy to do...

- Experiments: Mytkowicz, Diwan, Hauswirth, Sweeney. **Producing wrong data without doing anything obviously wrong!**. SIGPLAN Not. 44(3), March 2009



But do we **really** have to care in CS?

Yes. Model \neq Reality. Although designed and built by human beings, computers are **so complex** that mistakes are easy to do...

- **Experiments:** Mytkowicz, Diwan, Hauswirth, Sweeney. **Producing wrong data without doing anything obviously wrong!**. SIGPLAN Not. 44(3), March 2009

What your research supposedly looks like:

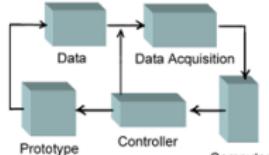


Figure 1. Experimental Diagram

What your research *actually* looks like:



Figure 2. Experimental Mess

C.S. suffers from the same difficulties as natural sciences

- Rely on large, distributed, **evolving**, prototype hard/software
- Validation on a few datasets/scenarios? 😞

But do we **really** have to care in CS?

Yes. Model \neq Reality. Although designed and built by human beings, computers are **so complex** that mistakes are easy to do...

- **Experiments:** Mytkowicz, Diwan, Hauswirth, Sweeney. **Producing wrong data without doing anything obviously wrong!**. SIGPLAN Not. 44(3), March 2009

What your research supposedly looks like:

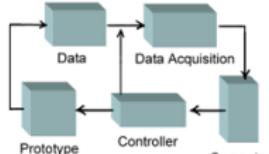


Figure 1. Experimental Diagram

What your research *actually* looks like:



Figure 2. Experimental Mess

C.S. suffers from the same difficulties as natural sciences

- Rely on large, distributed, **evolving**, prototype hard/software
- Validation on a few datasets/scenarios? 😞

- **Statistics:** **Trouble at the lab**, The Economist 2013

According to some estimates, three-quarters of published scientific papers in the field of machine learning are bunk because of this "overfitting".

Sandy Pentland, MIT

- **Numerical reproducibility:** simulated results are often non reproducible when moving from a platform to another or exploiting parallel architectures

All I care about is the algorithm output

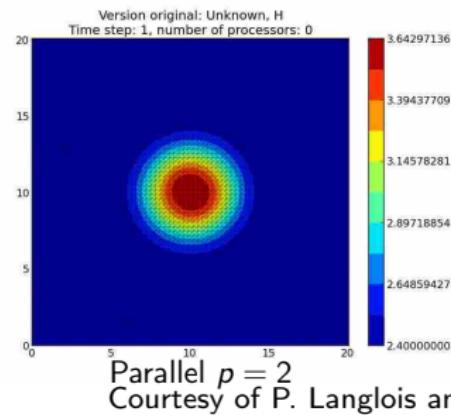
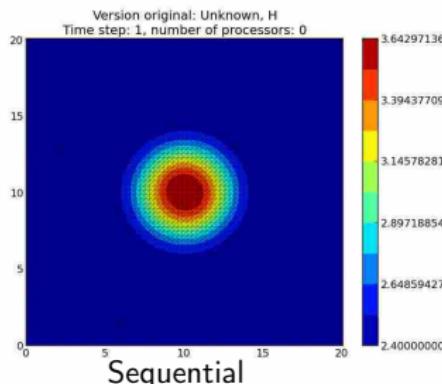
Did I mention we have parallel machines nowadays? 😊

Telemac2D: the simplest gouttedo simulation

The gouttedo test case

- 2D-simulation of a water drop fall in a square bassin
- Unknown: water depth for a 0.2 sec time step
- Triangular mesh: 8978 elements and 4624 nodes

Expected numerical reproducibility (time step = 1, 2, ...)



Courtesy of P. Langlois and R. Nheili

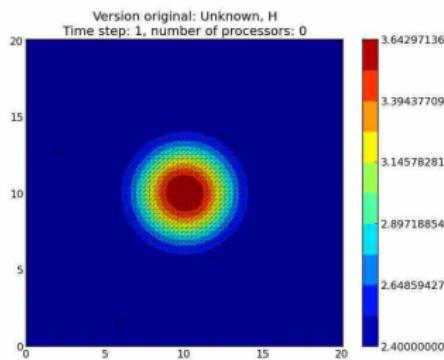
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

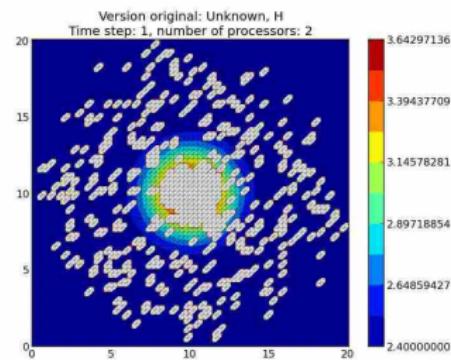
A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 1



Sequential



Parallel $p = 2$

Courtesy of P. Langlois and R. Nheili

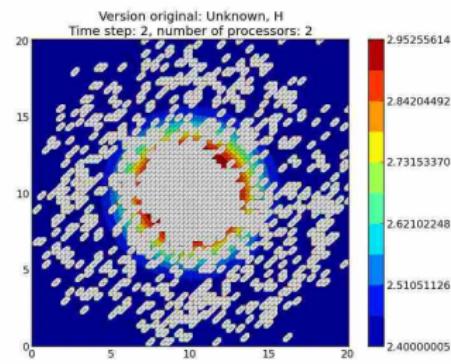
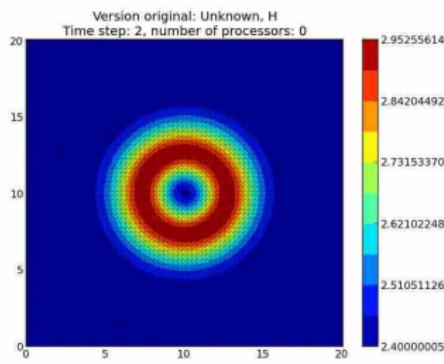
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 2



Courtesy of P. Langlois and R. Nheili

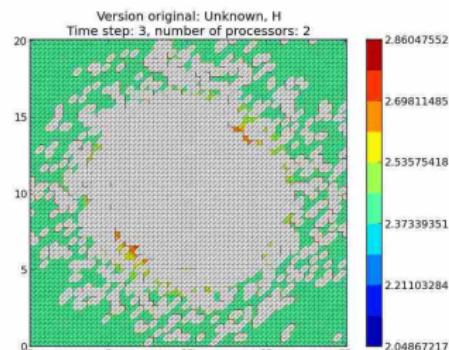
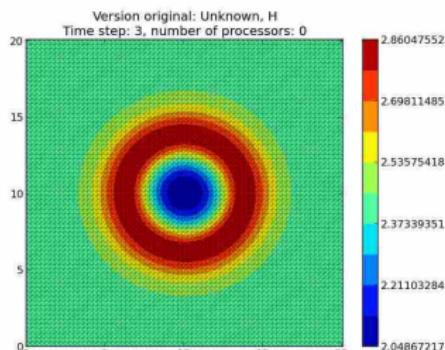
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 3



Courtesy of P. Langlois and R. Nheili

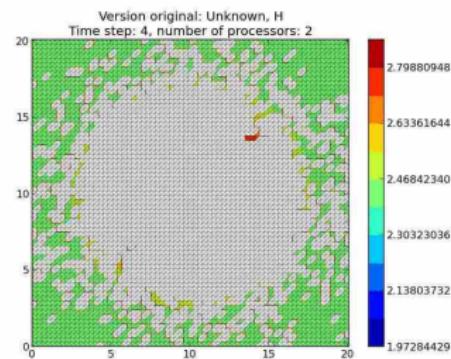
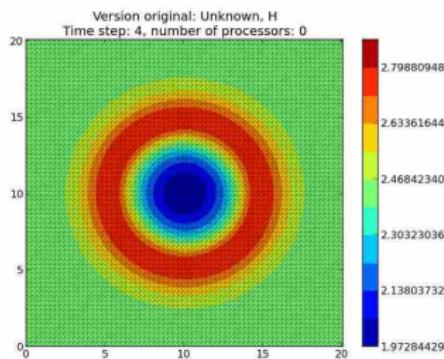
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 4



Courtesy of P. Langlois and R. Nheili

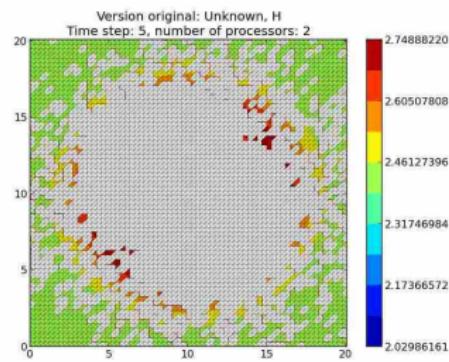
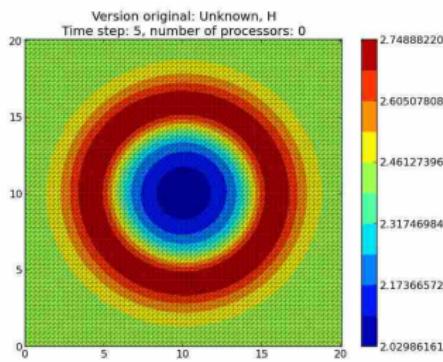
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 5



Courtesy of P. Langlois and R. Nheili

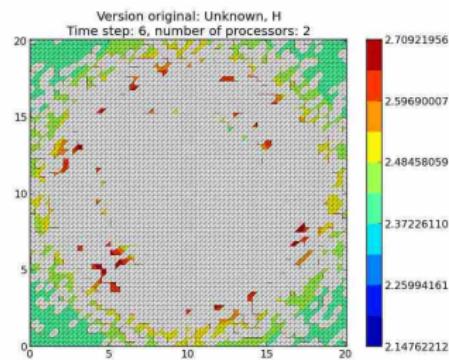
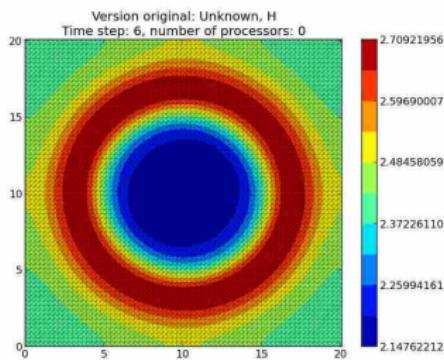
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 6



Courtesy of P. Langlois and R. Nheili

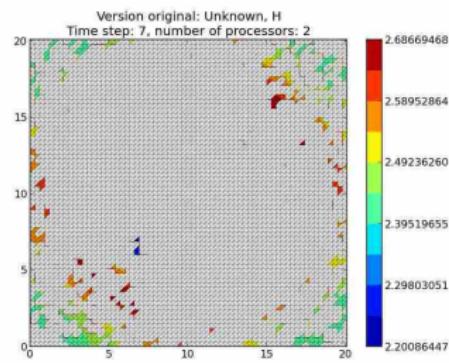
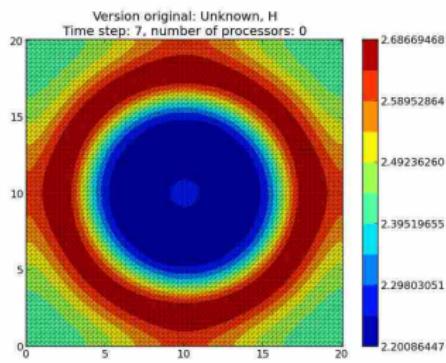
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 7



Courtesy of P. Langlois and R. Nheili

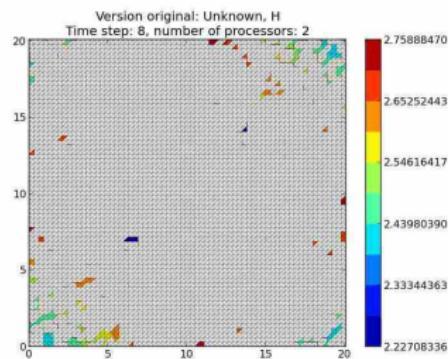
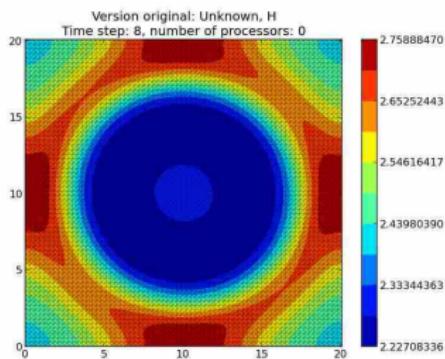
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 8



Courtesy of P. Langlois and R. Nheili

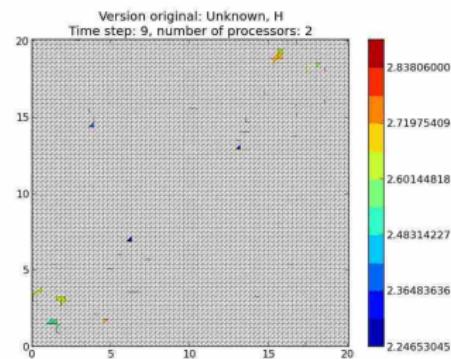
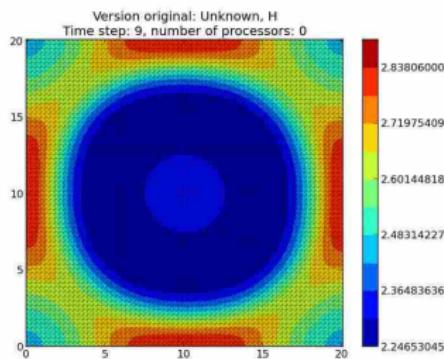
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 9



Courtesy of P. Langlois and R. Nheili

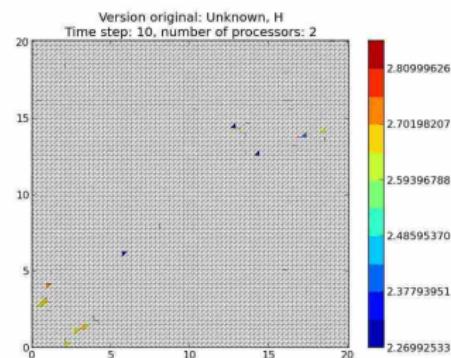
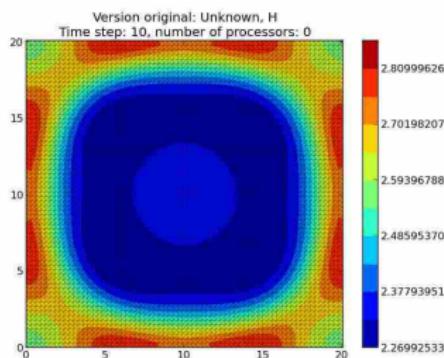
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 10



Courtesy of P. Langlois and R. Nheili

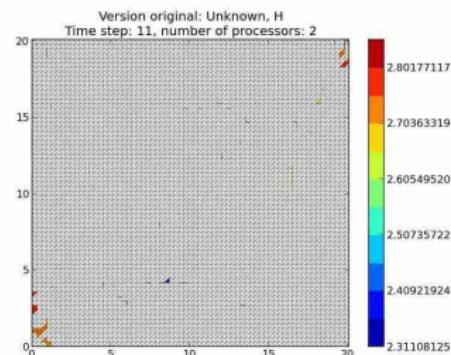
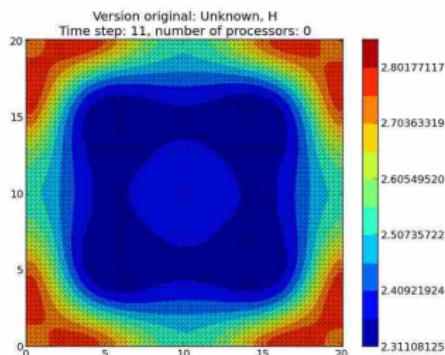
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 11



Courtesy of P. Langlois and R. Nheili

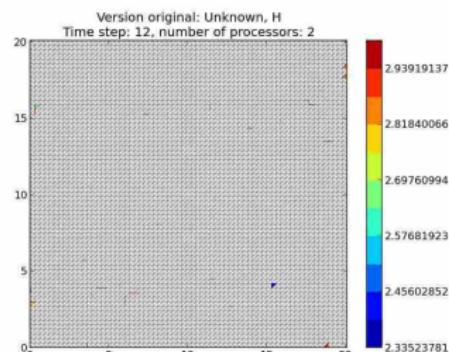
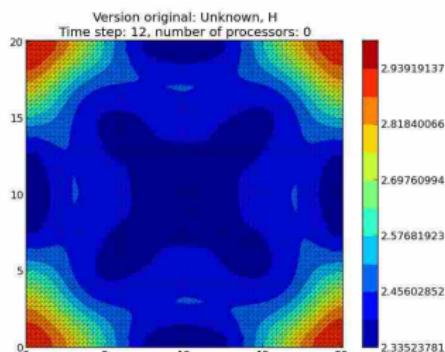
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 12



Courtesy of P. Langlois and R. Nheili

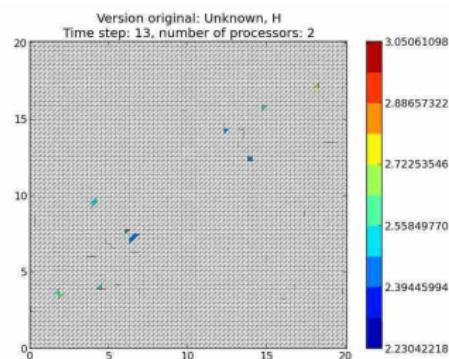
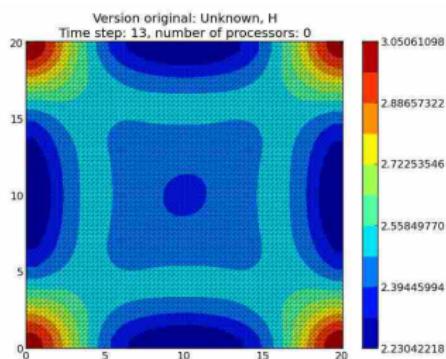
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 13



Courtesy of P. Langlois and R. Nheili

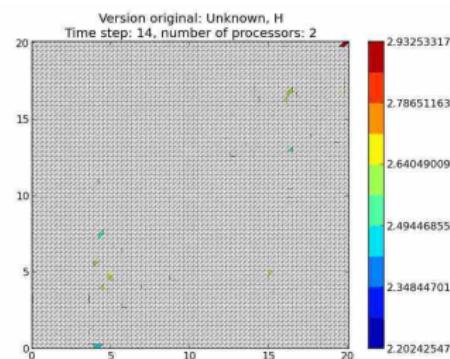
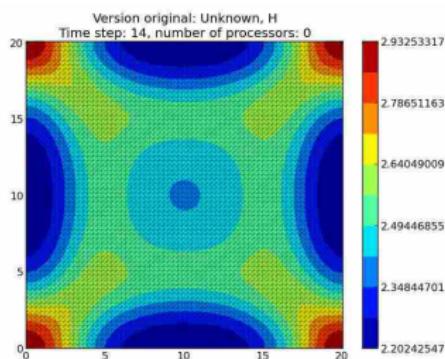
All I care about is the algorithm output

Did I mention we have parallel machines nowadays? 😊

A white plot displays a non-reproducible value

Numerical reproducibility?

time step = 14



Courtesy of P. Langlois and R. Nheili

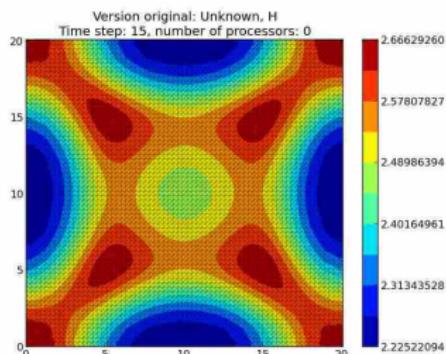
All I care about is the algorithm output

Did I mention we have **parallel machines** nowadays? 😊

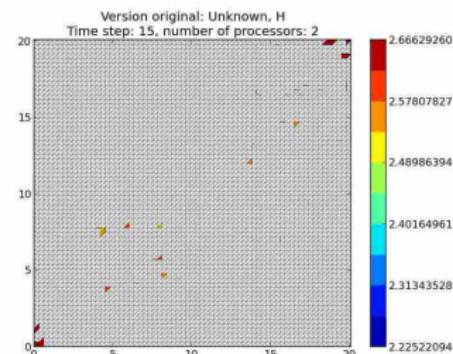
A white plot displays a non-reproducible value

NO numerical reproducibility!

time step = 15



Sequential



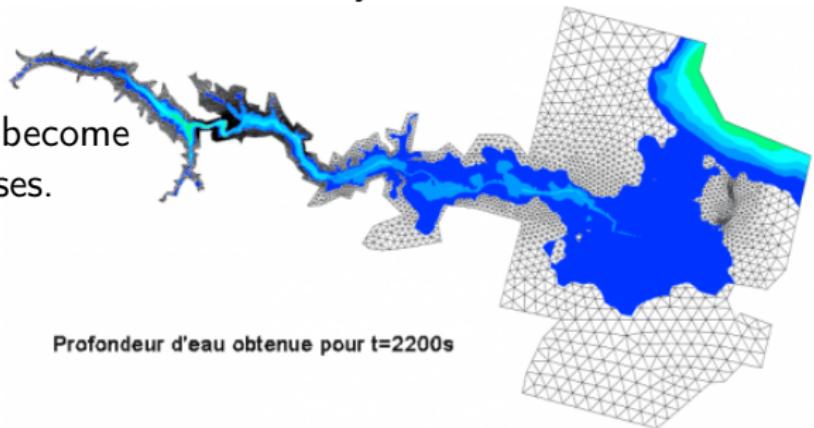
Parallel $p = 2$

Courtesy of P. Langlois and R. Nheili

All I care about is the algorithm output

Did I mention we have **parallel machines** nowadays? 😊

These numerical issues can become quite harmful in real use cases.



Numerical reproducibility?: Approximations in the model, in the algorithm, in its implementation, in its execution.

The whole chain needs to be revisited.

Why Are Scientific Studies so Difficult to Reproduce?

Human error:

- Experimenter bias
- Programming errors or data manipulation mistakes
- Poorly selected statistical test

There is just no real incentive in doing so:

- Legal barriers, copyright
- Competition issue
- Publication bias (only the idea matters, not the details)
- Rewards for positive results, not for consolidating results

Technical difficulty:

- Hardware and software evolve too quickly. It's not worth it
- No resources for storing so much data/information
- Lack of easy-to-use tools

Reproducibility: What Are We Talking About?

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

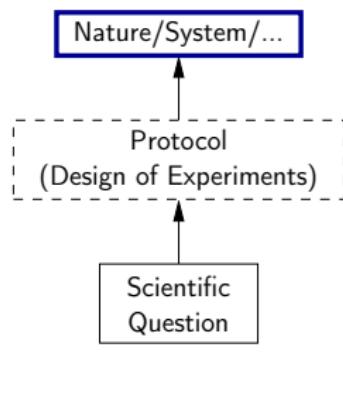
Completely independent reproduction based only on text description, without access to the original code

Reproducible Research: Trying to Bridge the Gap

Author

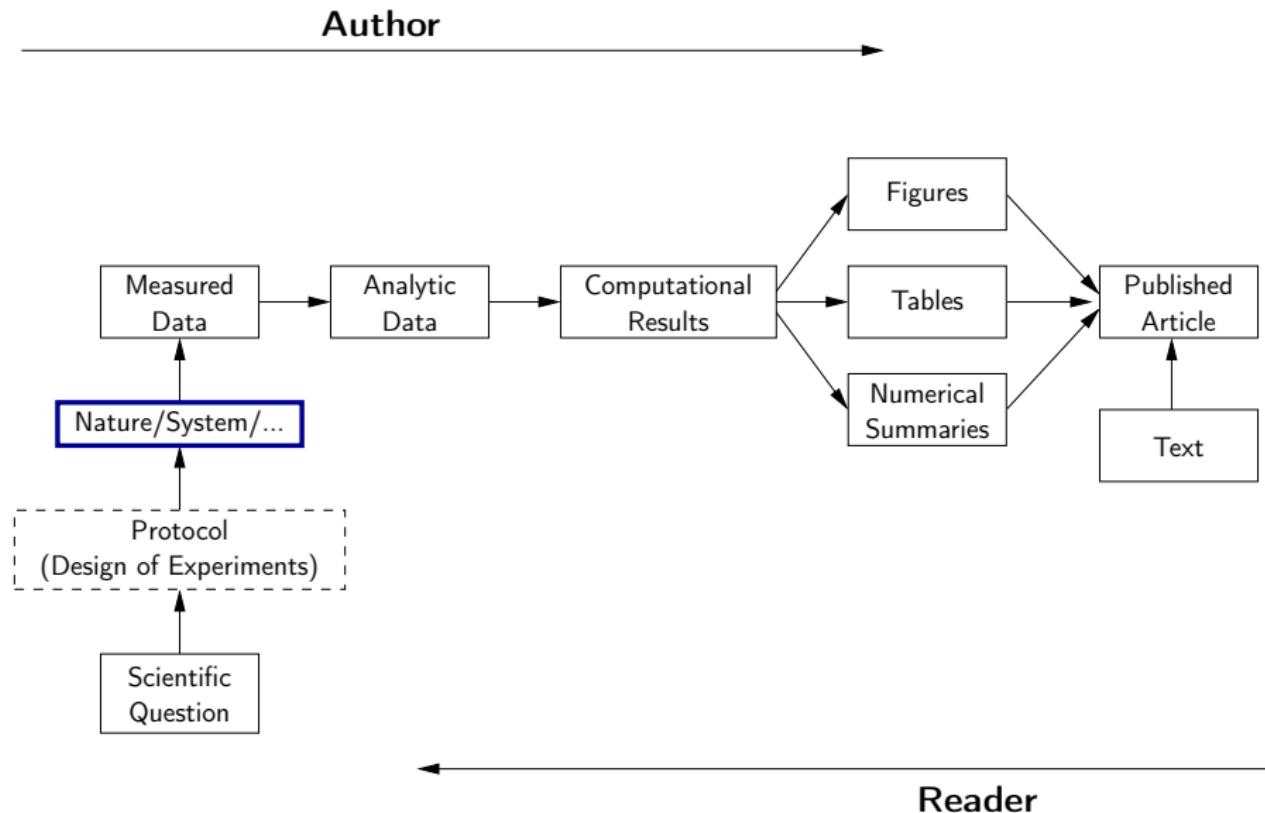


Published Article

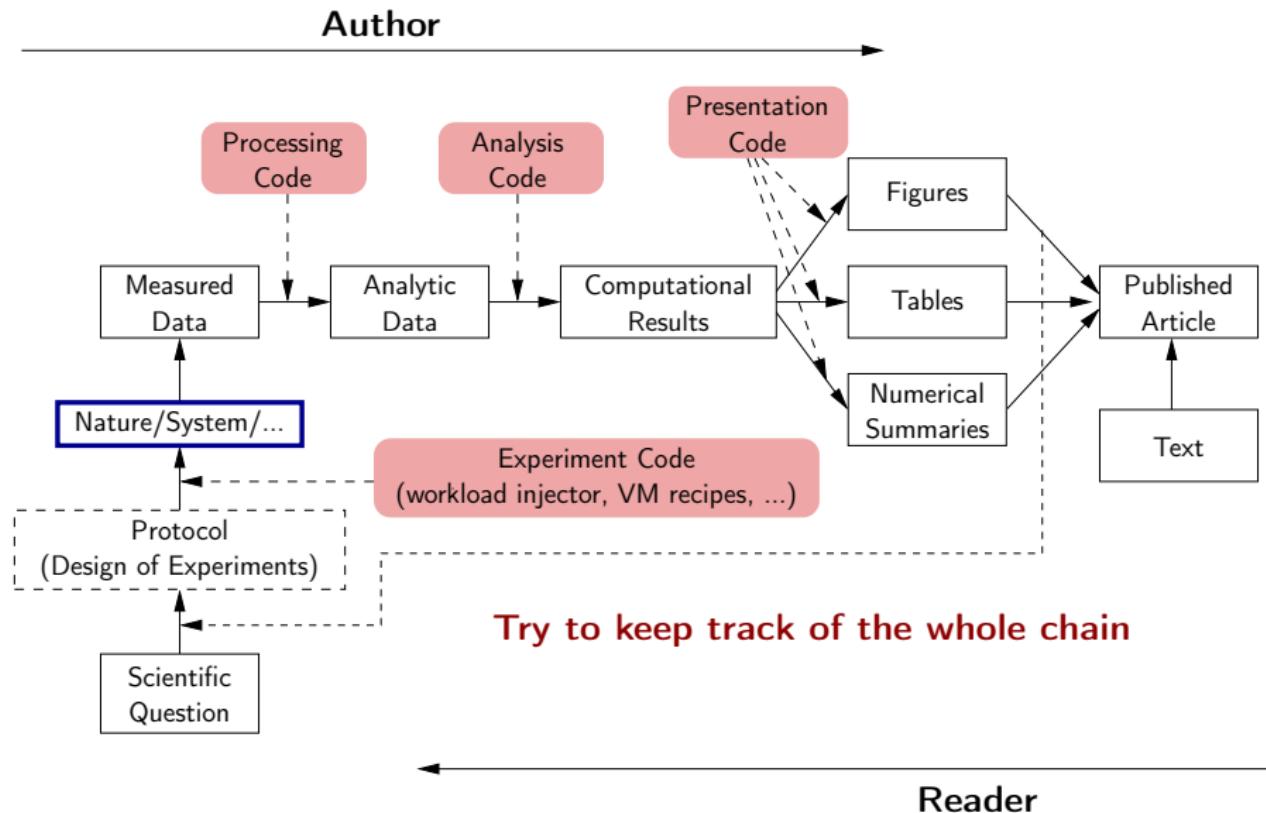


Reader

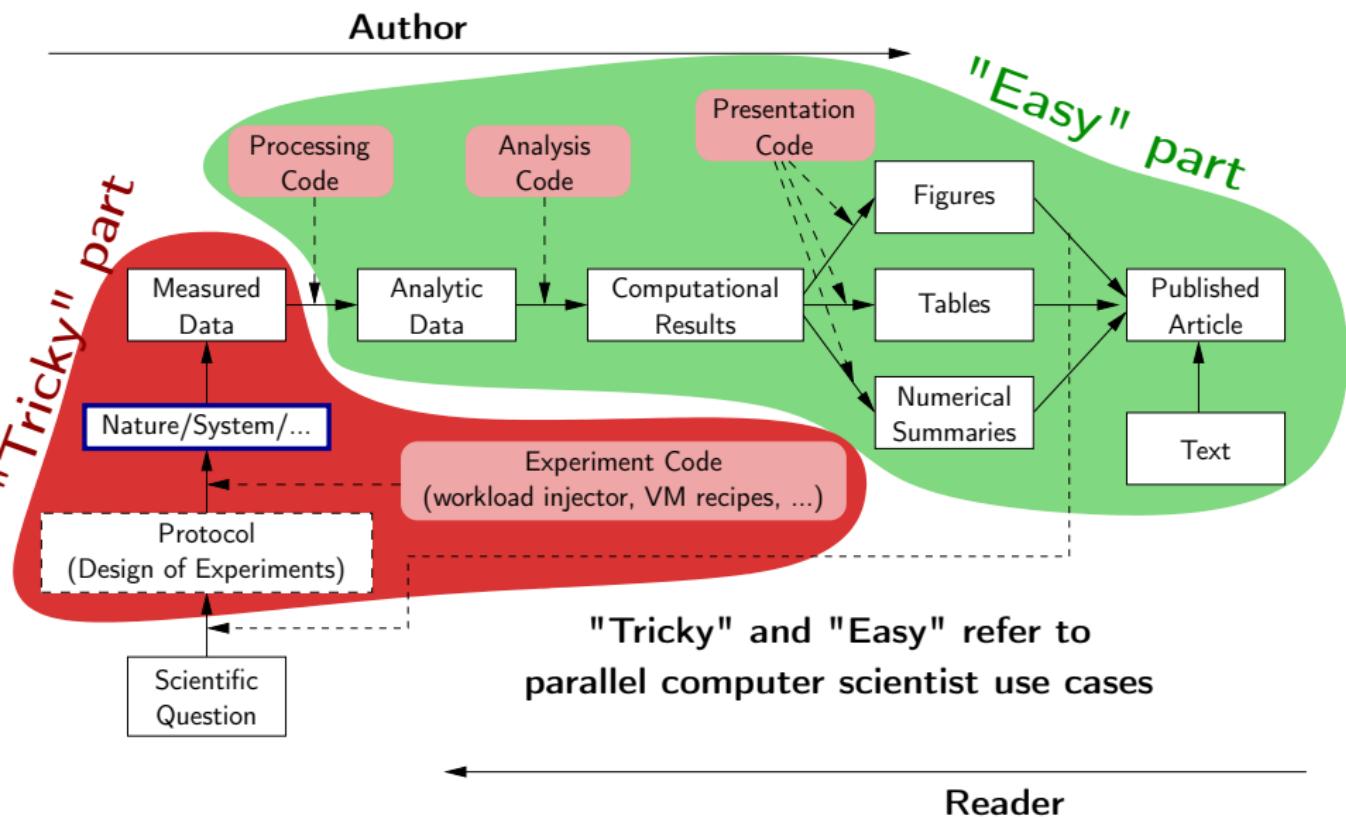
Reproducible Research: Trying to Bridge the Gap



Reproducible Research: Trying to Bridge the Gap



Reproducible Research: Trying to Bridge the Gap



Controlling what's happening

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

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

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

And this landscape is rapidly evolving

Outline

- ① A Few Motivating Examples
- ② What about Computer Science ?
- ③ Reproducible Research/Open Science: Illustrating Nice Ideas Through Different Tools

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

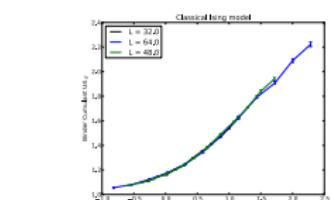
⁴Department of Physics and Astronomy, University of Wyoming, Laramie, Wyoming 82071, USA

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

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

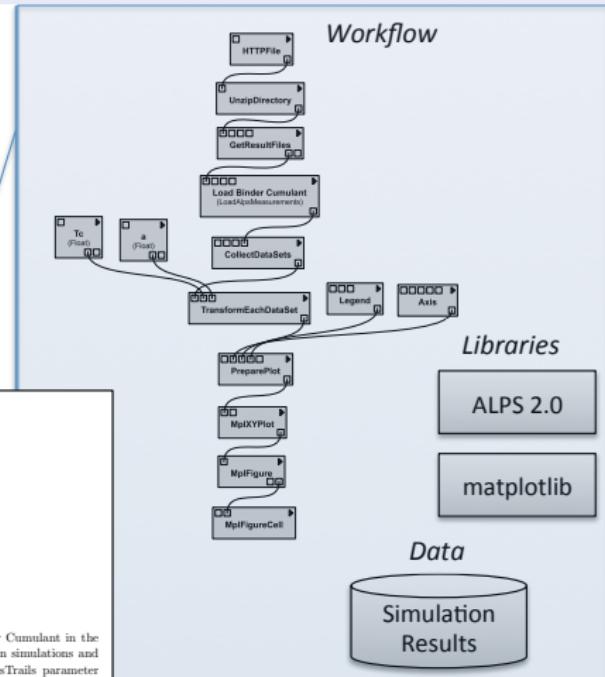
⁷Columbia University, New York, NY 10027, USA

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) ISTAT 2011

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}
```

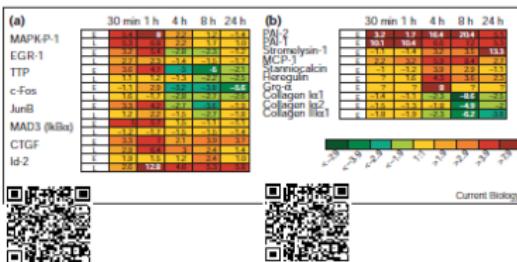
Permanently bind printed graphics to underlying result content

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].

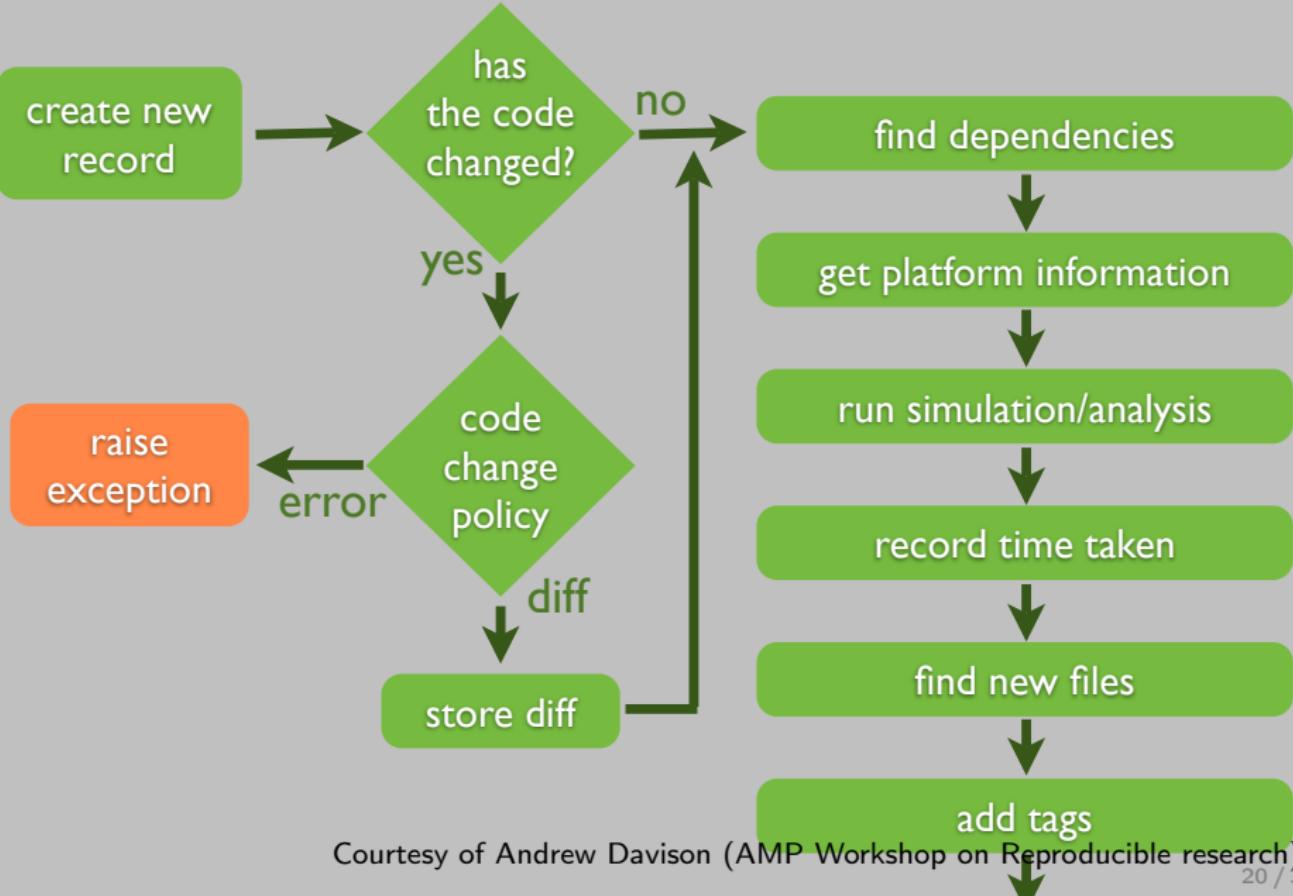
Deficiencies in the response of senescent cells to serum stimulation have been reported, and include an inability to induce the expression of *c-fos* mRNA [31] and markers of late G1 and S phase [32]. In response to serum, expression of inflammatory chemokines, matrix-degrading proteases and their modulators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

Courtesy of Matan Gavish and David Donoho (AMP Workshop on Reproducible research) 19 / 36

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 screenshot shows the Jupyter Notebook interface. On the left, there is a sidebar with the Jupyter logo and a "Welcome to the..." message. In the center, a main window titled "jupyter Lorenz Differential Equations (autosaved)" displays the content of a notebook. The content includes:

- A title section: "Exploring the Lorenz System".
- A text block: "In this Notebook we explore the [Lorenz system](#) of differential equations:
- Three differential equations:
$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = \rho x - y - xz$$
$$\dot{z} = -\beta z + xy$$
- A descriptive paragraph: "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."
- An "In [7]:" cell containing Python code:

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

 This cell also contains four sliders for "angle", "max_time", "σ", "β", and "ρ".
- A 3D plot of the Lorenz attractor, showing a complex, butterfly-shaped trajectory in three dimensions.
- Code cells at the bottom, including:

```
In [ ]: %matplotlib inline  
import pandas as pd  
import numpy as np  
import matplotlib
```

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 😊, Figshare, Zenodo, ActivePapers 😊, Elsevier executable paper 😞, ...

A Difficult Trade-off

Many different tools/approaches developed in various communities

But mainly two approaches:

- **Automatically keeping track of everything**
 - the code that was run (source code, libraries, compilation procedure)
 - processor architecture, OS, machine, date, ...
- **Ensuring others can understand/adapt what was done**
 - Why did I run this? Does it still work when I change this piece of code for this one?

A Difficult Trade-off

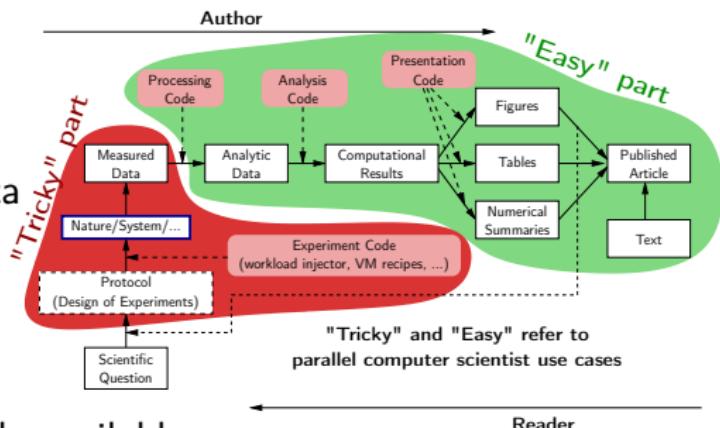
Many different tools/approaches developed in various communities

But mainly two approaches:

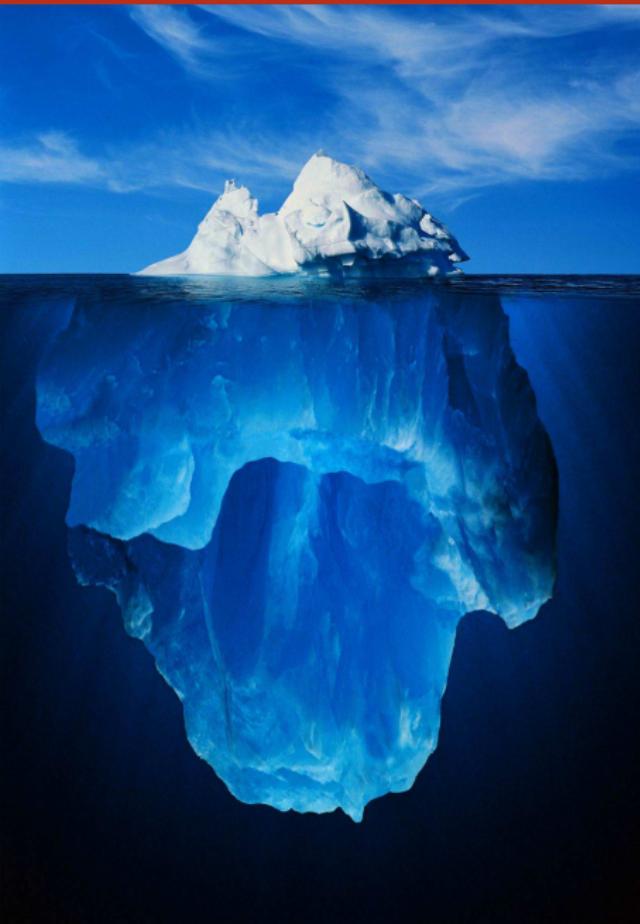
- Automatically keeping track of everything
 - the code that was run (source code, libraries, compilation procedure)
 - processor architecture, OS, machine, date, ...
- Ensuring others can understand/adapt what was done
 - Why did I run this? Does it still work when I change this piece of code for this one?

And the following key points:

- ① Replicable article
- ② Logging your activity
- ③ Logging and backup your data
- ④ Organizing your data
- ⑤ Mastering your environment
- ⑥ Controlling your experiments
- ⑦ Making your data/code/article available



Remember the general picture



The article is only the top of the iceberg, we need a way to **dive** and **unveil** what's behind every graphics and number...

1. Replicable article (Literate programming)

Donald Knuth: explanation of the program logic in a natural language interspersed with snippets of macros and traditional source code.

I'm way too 3133t to program this way 😊 but that's exactly what we need for writing a reproducible article/analysis!

Org-mode (my favorite! requires emacs though)

- Org-mode is plain text, very smooth, works both for html, pdf, ...
- Allows to combine all my favorite languages

Ipython/Jupyter notebook

Python user ↪ go for Jupyter. Web app, easy to use/setup... Writing replicable article may be tricky though

Knitr (a.k.a. Sweave)

For R and emacs users. Easy replicable articles with a modern IDE (e.g., Rstudio)

Note that this generation depends on a computational environment whose preservation is not addressed here (see for example activepapers).

2. Logging your activity (Laboratory Notebook)

- Jupyter project (formerly known as the IPython notebook)
- Org-mode again!
 - Capture mechanism (notes, todo, ...)
 - Babel favors code reuse, ssh connections in sessions, meta-programming
 - Tagging mechanism to structure the journal
 - Link mechanism, Todo, Calendar views, Tables, ...

Tips:

- Spending **more than an hour without** at least **writing** what you're working on **is not right...** Take a 5 min break and ask yourself what you're doing, what is keeping you busy and where all this is leading you
- While working on something, you will often notice/think about something you should fix/improve but you just don't want to do it now. Take 20 seconds to write a **TODO** entry
- There are moments where you have to **wait for something** (e.g. compiling) → perfect time for improving your notes
- **By the end of the day:** daily (and weekly) **review!**

Pros and Cons of these three tools

- Ipython notebook:
 - 😊 Easy to set up, user-friendly, machine readable format (JSON), easy sharing on the cloud
 - 😟 Writing an article, JSON, not fully polyglot
- knitR/Rstudio:
 - 😊 Easy to set up, user-friendly, writing articles, easy publishing on [rpubs](#)
 - 😟 not fully polyglot
- Emacs/Org-mode:
 - 😟 Emacs, steep learning curve
 - 😊 Powerful and versatile, yields control to power users, works both for writing articles and a notebook, good integration on github

The ultimate tool would combine an engine in an editor that allows collaborative interactive edition

3. Logging and backup your data

What are the options?

- Nothing 😞 (remember the funny examples from the beginning... 😊)
- Incremental backup mechanisms (e.g., time machine)
- The cloud! (e.g., Dropbox and Google Drive 😞 ...)
- Flexible version control systems (e.g. git 😊) where you're in control of what's happening
 - Use a crontab if you really do not want to think about it
 - We have come up with a specific **git branching workflow** for managing experimental results

4. Organizing your data

- Use the machine readable CSV format
- Provide raw data and meta data, not just statistical outputs
- Organization
 - Explain your conventions (e.g., src/, data/, script/, journal.org)
 - Git submodules
- Never do data manipulation and statistical tests by hand
- Use R, Python or another free software to read and process raw data.
 - Use a workflow that documents both data and process
 - The org-mode tangling mechanism may help

5. Mastering your environment

What are the options?

- Nothing 😊
- Restrict your tools/dependencies to the bare minimum (e.g., python)
 - List them all manually in a README
 - Use custom shell scripts or sosreport that log all the dependencies you are aware. Ask your friends to check whether this is sufficient...
 - Combine everything in activepapers, i.e., an HDFS5 file combining datasets and programs working on these datasets in a single package, along with meta data, history, ...
- Create and distribute your own virtual image (VM, docker, Singularity)
- Have tools that automatically keep track of dependencies/files and packages up the Code, Data, and Environment
 - CDE (Guo et al., 2011) ReproZip (Freire et al., 2013), CARE (Janin et al., 2014),
 - See Preserve the Mess or Encourage Cleanliness? (Thain et al., 2015)
- Use a specific tool to generate customized appliances (kvm, LXC, Virtualbox, iso, ...): recipes with steps and aliases, execution in contexts, checkpoints, ... (*Kameleon*)

6. Controlling your experiments

- Naive way: sh + ssh + ...

Parallel/distributed experiments differ from computational science and come with their own difficulties

- **Expo** (2007-, G5K)
 - **XPflow** (2012-, G5K)
 - **Execo** (2013-, G5K)
 - Plush (2006-, PlanetLab)
 - OMF (2009-, Wireless testbeds and Planetlab)
 - Splay (2008, distributed algorithm comparison), ...
- } although nothing specific to G5K

They differ in the underlying paradigms and the platforms for which they have been designed

- A survey of general-purpose experiment management tools for distributed systems, T. Buchert, C. Ruiz, L. Nussbaum, O. Richard, FGCS, 2014

7. Making your data/code/article available

- Your webpage 😞
- Figshare, Zenodo 😊, ...
- Companion websites ([elsevier executable paper](#) 😞, [runmycode](#), [exec&share](#) 😊, ...)
- Inria Forge/Gitlab, Github (damn, they're good! 😊), ...

This may seem easy but is more tricky than it looks like:

- Arbitrary limits can make your life painful
- Perennity
 - CodeSpaces murdered on Amazon, Google Code termination, Gitorious shutdown, ...
 - Disruption of the web of reference: URLs decay (half-life of 4 years), DOIs have little guarantee, ...

Where are we standing now?

- Changes in **funding agency** requirements
 - Starting? I hardly see how they could really enforce things
- Changes in journal/conferences **publication requirements**
 - Several attempts (artifact review and branding)
- **Cultural changes** in our **relation to publication**

Where are we standing now?

- Changes in **funding agency** requirements
 - Starting? I hardly see how they could really enforce things
- Changes in journal/conferences **publication requirements**
 - Several attempts (artifact review and branding)
- **Cultural changes** in our **relation to publication**
- **We** should care. What are the incentives?
 - Reproducible papers are **more cited?** 😊
 - Definitely **more efficient** (not only in the long run and for the community)
 - It's simply **more satisfying...** 😊
- **Train** researchers and **students** to use better tools, better research methodology, statistics/design of experiments, performance evaluation, ...

Next steps

Examples of reproducible articles with Org-mode:

- Fast and Accurate Simulation of Multithreaded Sparse Linear Algebra Solvers at ICPADS 2015. Reproducible version.
- Analyzing dynamic task-based applications on hybrid platforms: an agile scripting approach at VPA/SC 2016. Reproducible version.

Series of webinars on Reproducible Research:

Slides and Videos: https://github.com/alegrand/RR_webinars

Thank you

Contact:

- Arnaud Legrand: arnaud.legrand@inria.fr
- Lucas Mello Schnorr: schnorr@inf.ufrgs.br
- Vinícius Garcia Pinto: vinicius.pinto@inf.ufrgs.br

Slides

- <https://github.com/viniciovgp/pesquisareprodutivo1>