

# BUILDING THE FOUNDATIONS FOR OPEN-SOURCE GEOPHYSICS

Leonardo Uieda

University of Liverpool - 2019-10-23



Feel free to photograph and share this presentation.

BACKGROUND



## Rio de Janeiro

State of Rio de Janeiro  
Brazil

Rio de Janeiro is a huge seaside city in Brazil, famed for its Copacabana and Ipanema beaches, 38m Christ the Redeemer statue atop Mount Corcovado and for Sugarloaf Mountain, a granite peak with cable cars to its summit. The city is also known for its sprawling favelas (shanty towns). Its raucous Carnival festival, featuring parade floats, flamboyant costumes and samba dancers, is considered the world's largest.

[rio.rj.gov.br](http://rio.rj.gov.br)

US Dept of State Geographer  
© 2018 Google  
Image Landsat / Copernicus  
Data SIO, NOAA, U.S. Navy, NGA, GEBCO

Google Earth





Data LDEO-Columbia, NSF, NOAA

Data SIO, NOAA, U.S. Navy, NGA, GEBCO

Google Earth





THE GENERIC MAPPING TOOLS

Data LDEO-Columbia, NSF, NOAA

Data SIO, NOAA, U.S. Navy, NGA, GEBCO

Google Earth

University of Hawai'i at Mānoa

Hawaii



“I do some coding  
for my research”



“I am a self-taught coder”

Programming *languages*



# Goals

Why good code matters

How to do it

What you get in return

WHY

# The Reinhart-Rogoff error – or how not to Excel at economics

April 22, 2013 9.40pm BST

Data and computer code should be made publicly available at an early stage – or else ... esarastudillo

Email

Twitter

118

Facebook

Facebook

464

LinkedIn

LinkedIn

Print

Last week we learned a famous [2010 academic paper](#), relied on by political big-hitters to bolster arguments for austerity cuts, contained significant errors; and that those errors came down to misuse of an Excel spreadsheet.

Sadly, these are not the first mistakes of this size and nature when handling data. So what on Earth went wrong, and can we fix it?

Harvard's [Carmen Reinhart](#) and [Kenneth Rogoff](#) are two of the most respected and influential academic economists active today.

Or at least, they were. On April 16, doctoral student Thomas Herndon and professors Michael Ash and Robert Pollin, at the [Political Economy Research Institute](#) at the University of Massachusetts Amherst, released the results of their [analysis](#) of two 2010 papers by Reinhard and Rogoff, papers that also provided much of the grist for the 2011 bestseller [Next Time Is Different](#).

## Authors



**Jonathan Borwein (Jon)**

Laureate Professor of Mathematics, University of Newcastle



**David H. Bailey**

PhD; Senior Scientist, Lawrence Berkeley Laboratory (retired) and Research Fellow, University of California, Davis

## Disclosure statement

Jonathan Borwein (Jon) receives funding from ARC

Bailey does not receive any grant from Australian sources, nor does he have any other financial interest.

## Partners

## The Reinhart-Rogoff error – or how not to Excel at economics

April 22, 2013 9.40pm BST

Data and

✉ Email

🐦 Twitter

📘 Facebook

🌐 LinkedIn

🖨️ Print

***“The most serious was that, in their Excel spreadsheet, Reinhart and Rogoff had not selected the entire row when averaging growth figures...”***

handling data. So what on Earth went wrong, and can we fix it?

Harvard’s [Carmen Reinhart](#) and [Kenneth Rogoff](#) are two of the most respected and influential academic economists active today.

Or at least, they were. On April 16, doctoral student Thomas Herndon and professors Michael Ash and Robert Pollin, at the [Political Economy Research Institute](#) at the University of Massachusetts Amherst, released the results of their [analysis](#) of two 2010 papers by Reinhard and Rogoff, papers that also provided much of the grist for the 2011 bestseller [Next Time Is Different](#).



Jonathan Borwein (Jon) receives funding from Berkeley Laboratory (retired) and Research Fellow, University of California, Davis

### Disclosure statement

Jonathan Borwein (Jon) receives funding from ARC

Bailey does not receive any grant from Australian sources, nor does he have any other financial interest.

### Partners



HOW

# Developer superpowers



*Infinite undo:* Version Control

*X-ray vision:* Unit Tests

*How to write good:* Design Patterns

*Robot servants:* Continuous Integration

# *Infinite undo:* Version Control



# *Infinite undo:* Version Control



git

GitHub



# "FINAL".doc



FINAL.doc!



FINAL\_rev.2.doc



FINAL\_rev.6.COMMENTS.doc



FINAL\_rev.8.comments5.  
CORRECTIONS.doc



FINAL\_rev.18.comments7.  
corrections9.MORE.30.doc



FINAL\_rev.22.comments49.  
corrections.10. #@\$%WHYDID  
ICOMETOGRADSCHOOL?????.doc

JORGE CHAM © 2012

# *X-ray vision: Unit Tests*

```
def range_overlap(ranges):
    '''Return common overlap among a set of [left, right] ranges.'''
    max_left = 0.0
    min_right = 1.0
    for (left, right) in ranges:
        max_left = max(max_left, left)
        min_right = min(min_right, right)
    return (max_left, min_right)

def test_range_overlap():
    assert range_overlap([ (0.0, 1.0), (5.0, 6.0) ]) == None
    assert range_overlap([ (0.0, 1.0), (1.0, 2.0) ]) == None
    assert range_overlap([ (0.0, 1.0) ]) == (0.0, 1.0)
    assert range_overlap([ (2.0, 3.0), (2.0, 4.0) ]) == (2.0, 3.0)
    assert range_overlap([ (0.0, 1.0), (0.0, 2.0), (-1.0, 1.0) ]) == (0.0, 1.0)
    assert range_overlap([]) == None
```

Branch: master ▾

verde / verde / tests /

Create new file

Upload files

Find file

History



leouieda Update CI scripts and fix linting errors (#203) ...

Latest commit 047029a 6 days ago

..

📁 baseline	Add wind data from Texas (#104)	last year
📄 __init__.py	Start the repository with basic template files	2 years ago
📄 test_base.py	Add Azure Pipelines and remove AppVeyor (#176)	7 months ago
📄 test_blockreduce.py	Add option to reduce extra coordinates in BlockReduce (#198)	3 months ago
📄 test_chain.py	Rename Components to Vector (#130)	last year
📄 test_coordinates.py	Add longitude_continuity func to support slicing longitudes (#181)	3 months ago
📄 test_datasets.py	Flake8 check, simplify Makefile and Travis builds (#139)	last year
📄 test_distances.py	New median_distance func for near neighbors (#163)	10 months ago
📄 test_io.py	Surfer reader feature (#169)	6 months ago
📄 test_mask.py	Simplify distance_mask and add a projection arg (#125)	last year
📄 test_minimal.py	Add Azure Pipelines and remove AppVeyor (#176)	7 months ago
📄 test_scipy.py	Remove default shape for BaseGridder.grid (#140)	last year
📄 test_spline.py	Fix breaking Spline tests: extra warnings and type tests (#194)	3 months ago
📄 test_trend.py	Allow zero degree polynomials in Trend (a mean) (#162)	10 months ago
📄 test_utils.py	Update CI scripts and fix linting errors (#203)	6 days ago
📄 test_vector.py	Make more functions part of the base API (#156)	11 months ago

# *Robot servants:* Continuous Integration

Run tests every time something changes

TravisCI, Microsoft Azure, etc

Update website, publish releases, etc



[Dashboard](#)[Changelog](#)[Documentation](#)[Help](#)

fatiando / pooch



build passing

[Current](#)[Branches](#)[Build History](#)[Pull Requests](#)> [Build #308](#)[More options](#)✓ **v0.6.0** Some tweaks to Py2.7 warning and date of release (#110)

The changelog has yesterday's date as the release date. Also add the 2.7

[Commit e1df9d4](#)[Compare v0.6.0](#)[Tag v0.6.0](#)

Leonardo Uieda

🔑 #308 passed

[Restart build](#)

Ran for 2 min 36 sec

Total time 7 min 1 sec

a day ago

[Build jobs](#)[View config](#)

✓ # 308.1	AMD64		Linux - Python 3.7	1 min 12 sec	
✓ # 308.2	AMD64		Linux - Python 3.7 [optional]	1 min 16 sec	
✓ # 308.3	AMD64		Linux - Python 3.6	1 min 39 sec	
✓ # 308.4	AMD64		Linux - Python 3.5	1 min 50 sec	
✓ # 308.5	AMD64		Linux - Python 2.7	1 min 4 sec	

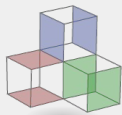
RETURN ON  
INVESTMENT

# (Geo)Scientific Python stack

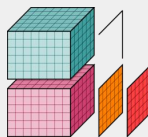
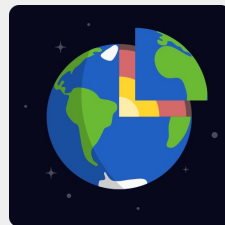


ObsPy

A Python Framework for Seismology

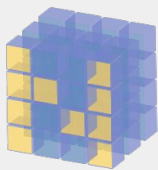


simpeg



xarray

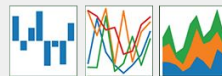
matplotlib



NumPy

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



python<sup>TM</sup>



# Fatiando a Terra

open-source tools for geophysics





Spatial data processing and interpolation (**gridding**) using Green's functions (or radial basis functions) with a machine learning inspired interface.

 [fatiando/verde](https://github.com/fatiando/verde)

 [www.fatiando.org/verde](http://www.fatiando.org/verde)

 doi: [10.21105/joss.00957](https://doi.org/10.21105/joss.00957)


 Stable and ready for use



Processing and modeling **gravity** and **magnetic** data, like terrain correction, upward continuation, equivalent layers, 3D inversion, and more.

 [fatiando/harmonica](https://github.com/fatiando/harmonica)

 [www.fatiando.org/harmonica/dev](http://www.fatiando.org/harmonica/dev)


 Early development and design



Manages the download of sample data files over HTTP from a server and storing them in a local directory. Used by our other libraries.

 [fatiando/pooch](https://github.com/fatiando/pooch)

 [www.fatiando.org/pooch](http://www.fatiando.org/pooch)


 Ready for use but still changing



Download geophysical models and datasets (PREM, CRUST1.0, ETOPO1) and load them into Python. Relies on Pooch to manage the downloads.

 [fatiando/rockhound](https://github.com/fatiando/rockhound)

 [www.fatiando.org/rockhound](http://www.fatiando.org/rockhound)

 Ready for use but still changing

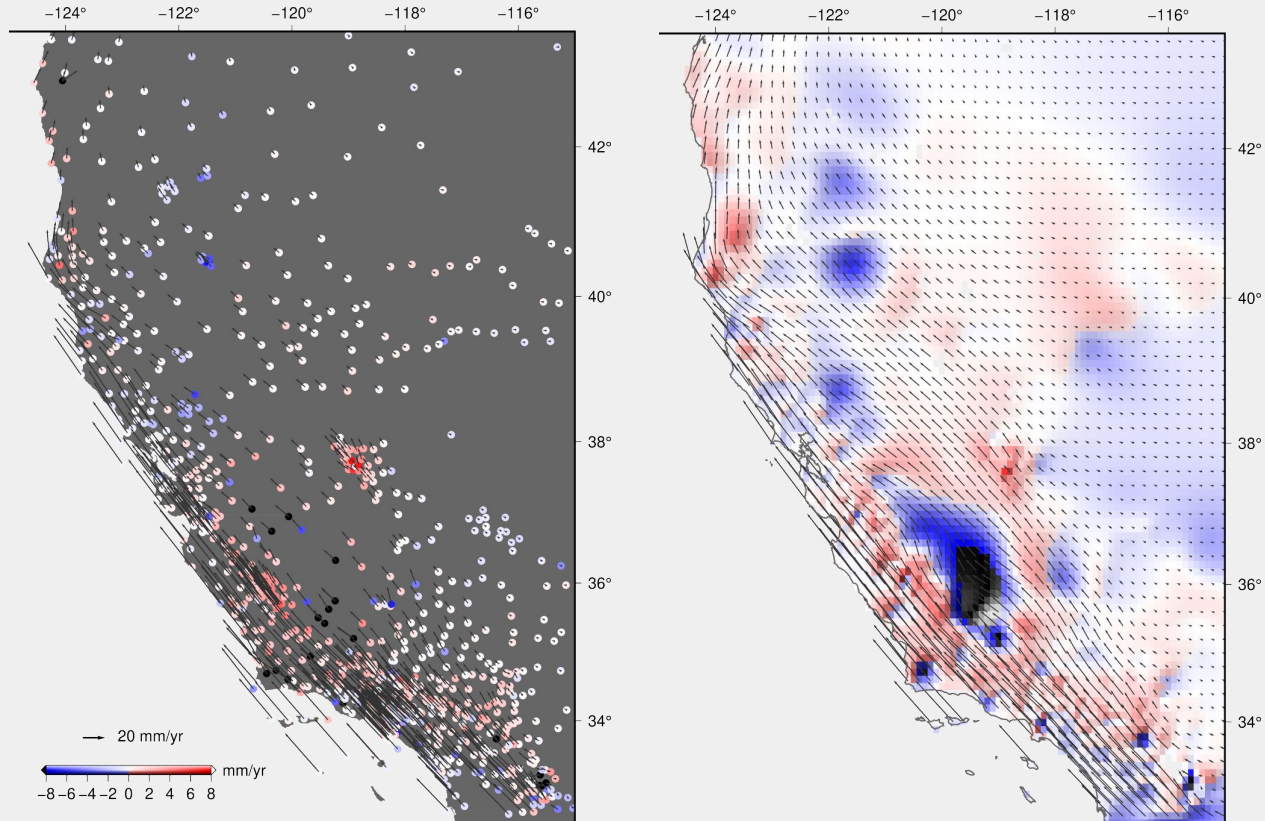
# Prototype new methods

Solid base to test new ideas

Diverse collaborators

Combine tools in new ways

# Example: GPS interpolation



# Reproducible research

Same tools for papers

Single command to generate results

Capture methods, parameters, etc

Reusable code

[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)[pinga-lab](#) / [paper-moho-inversion-tesseractoids](#)[Unwatch](#)

3

[★ Unstar](#)

12

[Fork](#)

6

[Code](#)[Issues](#) 0[Pull requests](#) 0[Projects](#) 0[Wiki](#)[Security](#)[Insights](#)[Settings](#)

Source code, data, and model results for "Fast non-linear gravity inversion in spherical coordinates with application to the South American Moho". Published in the Geophysical Journal International. <https://doi.org/10.1093/gji/ggw390>

[Edit](#)[geophysics](#)[geoscience](#)[earth-science](#)[python](#)[fatiando-a-terra](#)[science](#)[open-science](#)[jupyter-notebook](#)[Manage topics](#)[460 commits](#)[1 branch](#)[3 releases](#)[1 contributor](#)[View license](#)Branch: [master](#)[New pull request](#)[Create new file](#)[Upload files](#)[Find file](#)[Clone or download](#)[leouieda](#) Format manuscript without GJI style for archival

Latest commit b7ba65b 19 days ago

<a href="#">code</a>	Compliment and reorganize code README	3 years ago
<a href="#">data</a>	Update data README with all files	3 years ago
<a href="#">manuscript</a>	Format manuscript without GJI style for archival	19 days ago
<a href="#">model</a>	Result fig with seismic diffs for README	3 years ago
<a href="#">.env</a>	Environment file for autoenv	4 years ago
<a href="#">.gitignore</a>	Simplify Makefile to make it run on windows	3 years ago
<a href="#">LICENSE.md</a>	Include binder figshare + personal site links	3 years ago
<a href="#">README.md</a>	Add publication doi	3 years ago
<a href="#">environment.yml</a>	Register exact versions of packages used	3 years ago
<a href="#">screenshot-jupyter-notebook.png</a>	Add a screenshot of the jupyter notebook	4 years ago

CHALLENGES



# Social vs Technical

Code is the “easy” part

Building software is hard

Nurturing a community is even harder

# Incentives

Pushing AGU for support

Increased recognition

Desired in industry



The Journal of Open Source Software is a  
**developer friendly**, open access journal for  
research software packages.

Committed to publishing quality research software with zero article  
processing charges or subscription fees.

[Submit a paper to JOSS](#)[👤 Volunteer to review](#)[📖 Explore Papers](#)[📖 Documentation](#)[📖 Learn More](#)

## Recently Published Papers 727

**PUBLISHED** Published about 2 hours ago

**Hasasia: A Python package for Pulsar Timing Array Sensitivity Curves**

Python



@Hazboun6

DOI [10.21105/joss.01775](https://doi.org/10.21105/joss.01775)

# Training

Full curriculum

Qualified instructors

Workshops to the rescue



Teaching basic lab skills  
for research computing



Our Workshops ›  
Find or host a workshop.



Our Lessons ›  
Have a look at what we teach.



Get Involved ›  
Help us help researchers.

## Recent Blog Posts

In addition of the posts below, find out what's happening in our community through [The Carpentries blog](#), a great resource that collates posts from Data Carpentry, Library Carpentry, and Software Carpentry, and publishes updates of general interest to the community.

### Git lesson using worksheets

Parikshat Nanda / 2018-05-26

## Upcoming Workshops

### [The Bioinformatics Institute, A\\*STAR](#)

Oct 21 - Oct 23, 2019

Instructors: Ashar Malik

### [Institut Cavanilles de Biodiversitat i Biologia Evolutiva](#)

Oct 22 - Oct 23, 2019

Instructors: José Ignacio Lucas Lledó

FINALLY



# Take home

Good software matters.

So many tools exist! Use them.

Invest now, benefit in the future.