



CODE OCEAN

**Pierre Montagano**

Director of Business Development

[pierre@codeocean.com](mailto:pierre@codeocean.com)

# Before Code Ocean



1. Find the code
2. Acquire the right hardware
3. Set up the environment
4. Import the right files
5. Installing all dependencies...packages, versions, OS etc...
6. Errors.. Debugging.. Errors.. Debugging
7. Run
8. Results

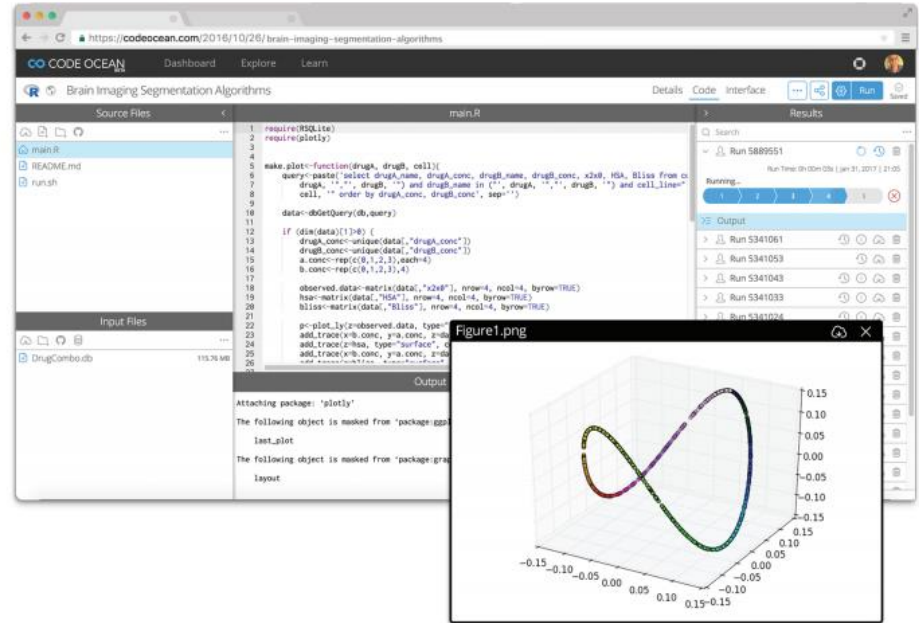
# After Code Ocean

Steps 1 to 6 are already configured, just press

Run

And see the results!

1. Find the code ✓
2. Acquire the right hardware ✓
3. Set up the environment ✓
4. Import the right files ✓
5. Installing all dependencies ✓
6. Errors.. Debugging.. Errors.. Debugging ✓



Code

causality

prediction

- GP\_causalityTimeFramesB.py
- GP\_query\_trueVar.py
- GP\_train.py
- GPpredict.m
- GPtrain.m
- predictSequence\_time.m
- predictSequence\_win.m

README.md

Data

example\_workspace.mat 261.04 KB

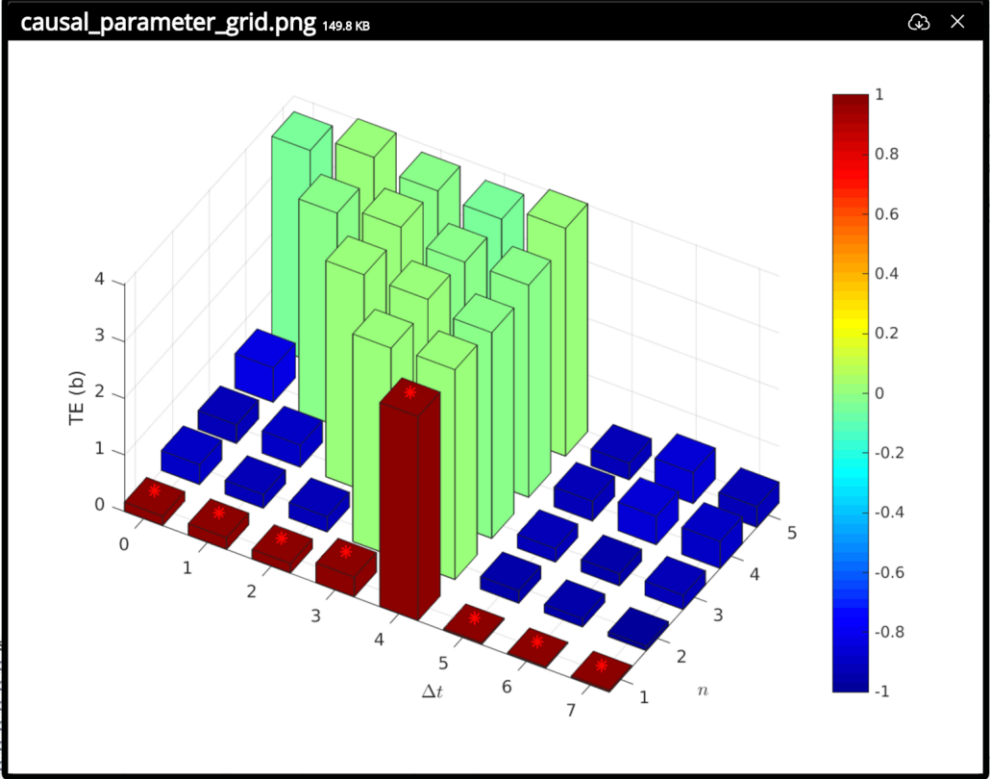
GP\_query\_trueVar.py

```
1 '''
2 Script to query a pre-trained gaussian process. Outputting the expected radius
3 for a set of normalised points on the objects surface.')
4
5 command line args:
6 -i input filepath
7 -o output filepath
8 -d num_dimensions
9 -n num_pts
10 -v verbose
11
12 return codes:
13 0 success
14 1 invalid csv file dimensionality
15 2 File IO failure
16
17 Created on Jun 17, 2013
18
19 @author: Simon
20 '''
21
22 import argparse
23 from numpy import *
24 import GPpy
25 from pickle import Unpickler
26 import sys
27
28 def vprint(flag, text):
29     if flag==1:
30         for item in text:
31             print "verbose: ",item
32             print "\n"
33
34 if __name__ == '__main__':
35     #read command line args
36     parser = argparse.ArgumentParser(description='Script to query a pre-trained gaussian process. Outputting the expected radius for a set of normalised points on the objects surface.')
37     parser.add_argument('-f',nargs=1,type=str)
38     parser.add_argument('-i',nargs=1,type=str)
39     parser.add_argument('-o',nargs=1,type=str)
40     parser.add_argument('-d',nargs=1,type=int)
41     parser.add_argument('-n',nargs=1,type=int)
42     parser.add_argument('-v',nargs=1,type=int)
```

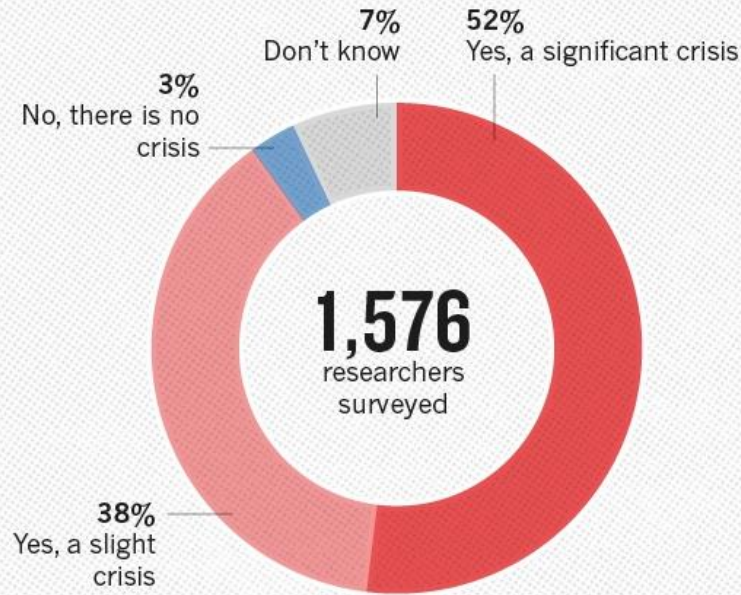
Results

Search

Published Result



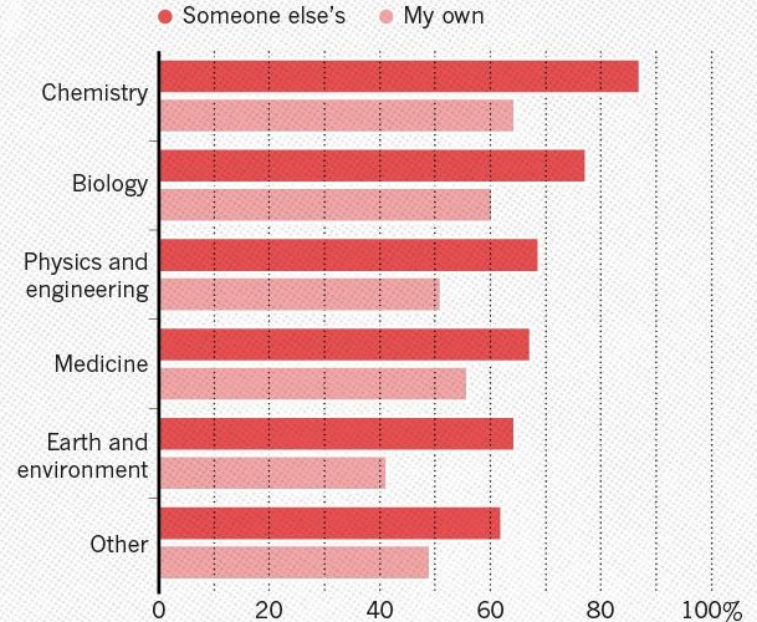
## IS THERE A REPRODUCIBILITY CRISIS?



©nature

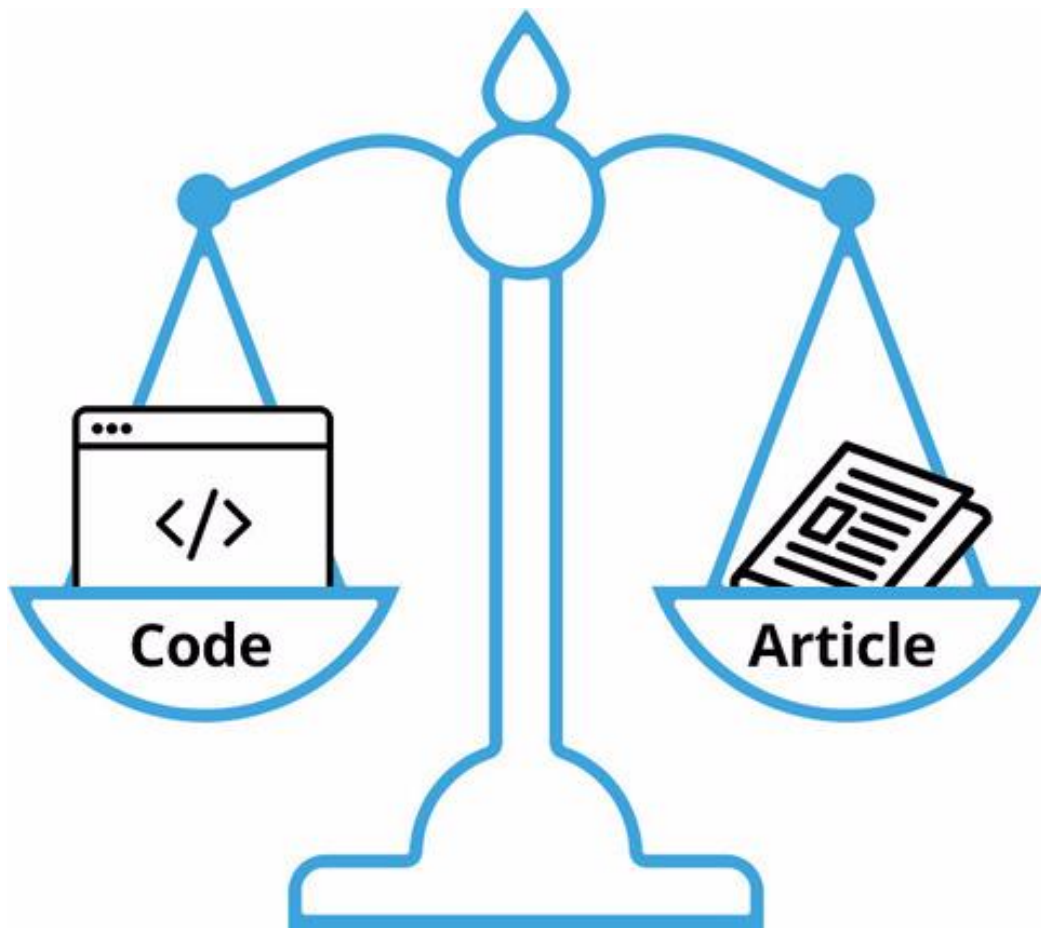
## HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.

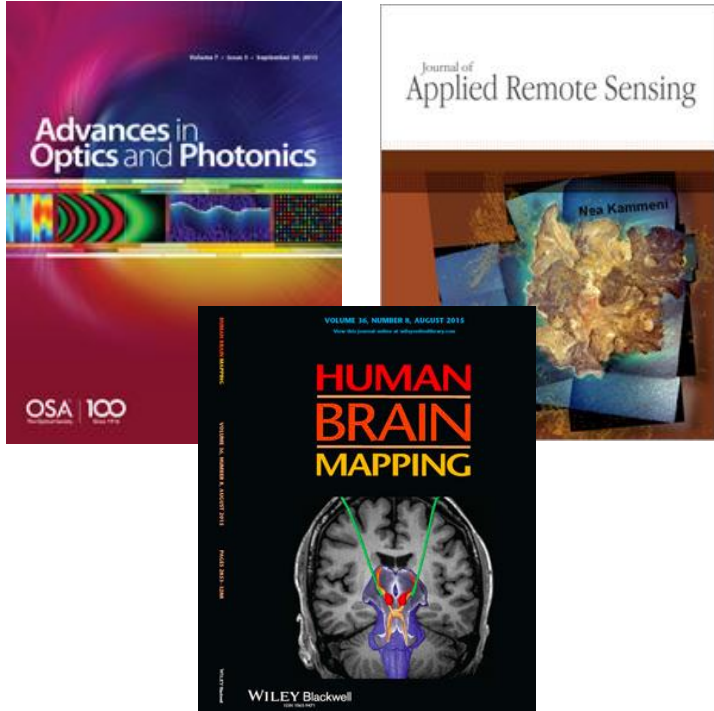


Nature **533**, 452–454 (26 May 2016) doi:10.1038/533452a

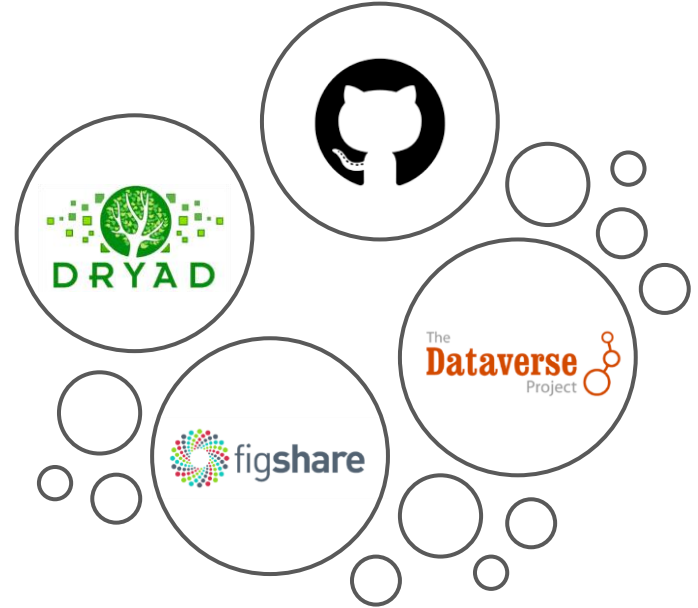




# Improving user experience



Articles



Data and Code



RESEARCH ARTICLE

# A reanalysis of mouse ENCODE comparative gene expression data [version 1; referees: 3 approved, 1 approved with reservations]

 Yoav Gilad, Orna Mizrahi-Man

 [Author details](#)

 [Grant information](#)



This article is included in the [Preclinical Reproducibility and Robustness gateway](#)

## Abstract

Recently, the Mouse ENCODE Consortium reported that comparative gene expression data from human and mouse tend to cluster more by species rather than by tissue. This observation was surprising, as it contradicted much of the comparative gene regulatory data collected previously, as well as the common notion that major developmental pathways are highly conserved across a wide range of species, in particular across mammals. Here we show that the Mouse ENCODE gene expression data were collected using a flawed study design, which confounded sequencing batch (namely, the assignment of samples to sequencing flowcells and lanes) with species. When we account for the batch effect, the corrected comparative gene expression data from human and mouse tend to cluster by tissue, not by species.

 Corresponding author: Yoav Gilad

How to cite: Gilad Y and Mizrahi-Man O. A reanalysis of mouse ENCODE comparative gene expression data [version 1; referees: 3 approved, 1 approved with reservations]




METRICS

25047

 VIEWS

2294

 DOWNLOADS

 Get PDF

 Get XML

 Cite





 Export

 Track

 Email

 Share

## Open Peer Review

Referee Status:    

### Invited Referees

Version(s)	1	2	3	4
Version 1 published 19 May 2015	 read report	 read report	 read report	 read report

- 1 **Rafael Irizarry**, Harvard School of Public Health, USA
- 2 **Michael Eisen**, University of California, Berkeley, USA
- 3 **Mick Watson**, University of Edinburgh, UK
- 4 **Lior Pachter**, University of California, Berkeley, USA

All reports (4), Responses and comments (1)

## Comments on this article

All comments (33)

[Add a Comment](#)

## Subscribe to Table of Content Alert

Your email address

**SUBSCRIBE**



# Why is it important?



## Reproducibility

Code Ocean allows other researchers to run the code, to test for reproducible results, and also execute the code with new input values.



## Impact

There is a positive correlation between the availability of code and software and the number of citations



## Reuse

Allows users to run multiple versions of the algorithm augmenting the code or/and using alternative inputs.



## Move beyond the pdf

Brings authors work alive in an executable environment and provides a new level of engagement to end users.



Thank you for your time



**Pierre Montagano**  
Director of Business Development  
[pierre@codeocean.com](mailto:pierre@codeocean.com)