Production is not the application of tools to materials.  

It is the application of logic to work.  


K. Jarrod Millman  
Helen Wills Neuroscience Institute  
University of California, Berkeley  

Applied Mathematics Perspectives 2011  
Reproducible Research: Tools and Strategies for Scientific Computing
Make the dirt fly!

The challenge of reproducible research in the computer age
Culture matters

The challenge of reproducible research in the computer age
Build quality into the process
The radical novelty of computing

The concept of radical novelties is of contemporary significance because, while we are ill-prepared to cope with them, science and technology have now shown themselves expert at inflicting them upon us.

Better, faster, cheaper

- Are we doing a good (enough) job? How would we know?
- How long does it take to go from the idea as presented in (say) lab meeting to the paper being submitted?
- What proportion of measured data makes it to publication?
- Are we duplicating work that other people have done already?
- Are we doing work for other people because they don’t know how to do it?
- Are there tasks that can be automated?
“truth will sooner come out of error than from confusion.”

...so when a man tries all kinds of experiments without method or order, this is mere groping in the dark; but when he proceeds with some direction and order in his experiments, it is as if he were led by the hand...

— Francis Bacon, Novum Organum (1620)
Neuroimaging

The challenge of reproducible research in the computer age
Dee magic begins here...

- Specialization
- Lack of patience
- Lack of understanding
- Confusion, frustration, and helplessness
The purpose of NIPY is to make it easier to do better brain imaging research. We believe that neuroscience ideas and analysis ideas develop together. Good ideas come from understanding; understanding comes from clarity, and clarity must come from well-designed teaching materials and well-designed software. The software must be designed as a natural extension of the underlying ideas.

We aim to build software that is:

- clearly written
- clearly explained
- a good fit for the underlying ideas
- a natural home for collaboration
The process

- How many mistakes do you make?
- What do they cost?
- Could you have made mistakes you don’t know about?
Data & code sharing

- Could you send someone else in the lab an email with all the information they need to rerun your analysis?
- How long would it take to write that email?
Git for everything

Git is a free & open source, distributed version control system designed to handle everything from small to very large projects with speed and efficiency.

Every Git clone is a full-fledged repository with complete history and full revision tracking capabilities, not dependent on network access or a central server. Branching and merging are fast and easy to do.

Git is used for version control of files, much like tools such as Mercurial, Bazaar, Subversion, CVS, Perforce, and Team Foundation Server.

Projects using Git
- Git
- Linux Kernel
- Perl
- Eclipse
- Gnome
- KDE
- Qt
- Ruby on Rails
- Android
- PostgreSQL
- Debian
- X.org

Download Git
The latest stable Git release is v1.7.6

Release notes (2011-06-26)

Windows Mac OSX Source

Other Download Options
Git Source Repository

The challenge of reproducible research in the computer age
Perspective & Plan
Tools & Practices
Invitations

Python

The challenge of reproducible research in the computer age
“Literate programming”

- Sweave: \LaTeX{} & R
- Sphinx: reStructuredText & Python
Automate, automate, automate
Programming as a first class citizen

- Read programming articles, books, etc.
- Learn new languages
Agile methodology

- Test driven development
- Pair programming
- Metaprogramming
Since 1997, **Software Carpentry** has taught scientists and engineers the concepts, skills, and tools they need to use and build software more productively. All of the content is freely available under a Creative Commons license, and we are constantly adding and updating lectures, videos, and exercises.

Can Software Carpentry help you? These [comments from former students](http://software-carpentry.org/about) and our [three-minute pitch](http://software-carpentry.org/about), can help you decide.

Can you help Software Carpentry? We are an open source/open license project, and there are many ways in which volunteers can [contribute](http://software-carpentry.org/contributing). We also have several [funding models](http://software-carpentry.org/funding), so if you would like to help people around the world solve the problems they face today, and prepare them to tackle the larger challenges of parallelism, cloud computing, reproducible research, and global-scale collaboration, please get in touch.
The challenge of reproducible research in the computer age

SciPy (pronounced "Sigh Pie") is open-source software for mathematics, science, and engineering. It is also the name of a very popular conference on scientific programming with Python. The SciPy library depends on NumPy, which provides convenient and fast N-dimensional array manipulation. The SciPy library is built to work with NumPy arrays, and provides many user-friendly and efficient numerical routines such as routines for numerical integration and optimization. Together, they run on all popular operating systems, are quick to install, and are free of charge. NumPy and SciPy are easy to use, but powerful enough to be depended upon by some of the world's leading scientists and engineers. If you need to manipulate numbers on a computer and display or publish the results, give SciPy a try!
The challenge of reproducible research in the computer age
Open Research Computation

Editor-in-Chief
Cameron Neylon (UK)

*Open Research Computation* is an open access journal that publishes articles describing the development, capacities, and uses of software for researchers in any field. The journal also encourages submissions that review or describe developments relating to software based research tools. All software source code published in *Open Research Computation* is made available under an Open Source Initiative compliant license.

Submit your manuscript and benefit from:
- High visibility for articles through unrestricted online access
- No limits on article length, additional files, colour figures or movies
- Immediate open access publication on acceptance
- Expert peer review

[www.openresearchcomputation.com](http://www.openresearchcomputation.com)

The challenge of reproducible research in the computer age