The challenge of reproducible research in the computer age

K. Jarrod Millman
Brain Imaging Center
University of California, Berkeley

millman@berkeley.edu

SIAM Conference on Computational Science and Engineering (CSE11)
Verifiable, reproducible research and computational science Mini Symposium
March 4, 2011

"Science is the belief in the ignorance of experts"

Science alone of all the subjects contains within itself the lesson of the danger of belief in the infallibility of the greatest teachers in the preceding generation... Learn from science that you must doubt the experts.

— Richard Feynman, What is Science? (1969)

"... none absolutely certain"

When a scientist doesn't know the answer to a problem, he is ignorant. When he has a hunch as to what the result is, he is uncertain. And when he is pretty darned sure of what the result is going to be, he is in some doubt... Scientific knowledge is a body of statements of varying degrees of certainty — some most unsure, some nearly sure, none ***absolutely*** *certain.*

— Richard Feynman, What Do You Care What Other People Think? (1988)

Scientific method

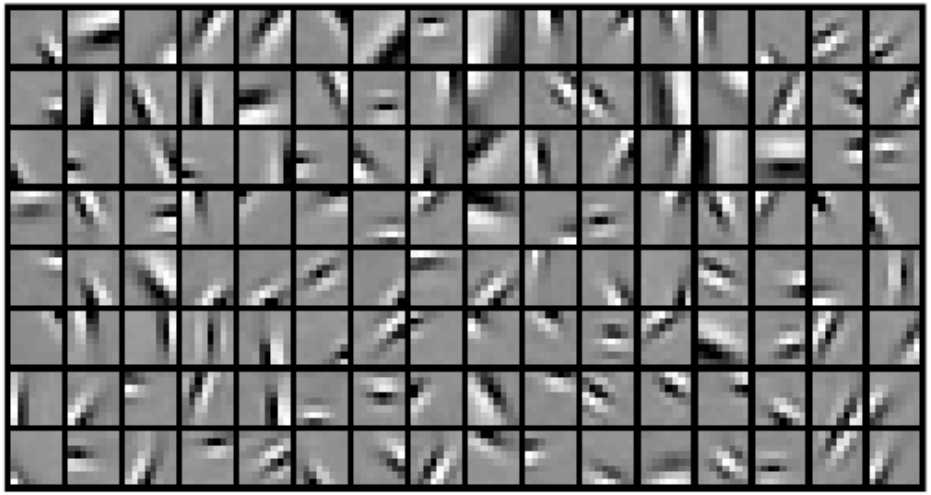
1590 — controlled experiment (Bacon)
1665 — repeatability (Boyle)
1687 — hypothesis/prediction (Newton)
1920 — falsifiability (Popper)
1926 — randomized design (Fisher)
1937 — controlled placebo*
1946 — computer simulation
1950 — double blind experiment*

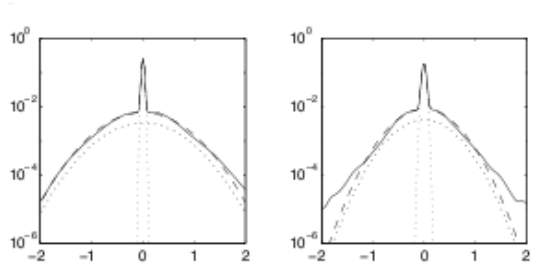
"Take nobody's word for it"

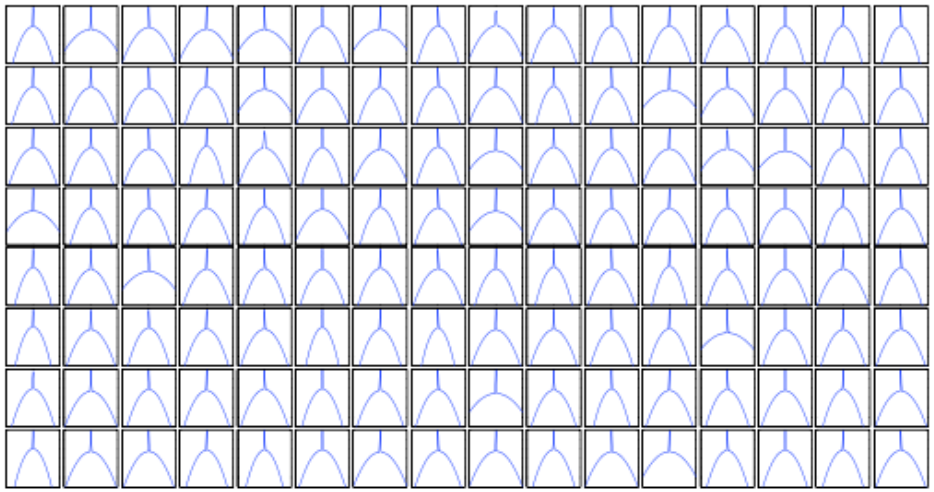
Scientific journals

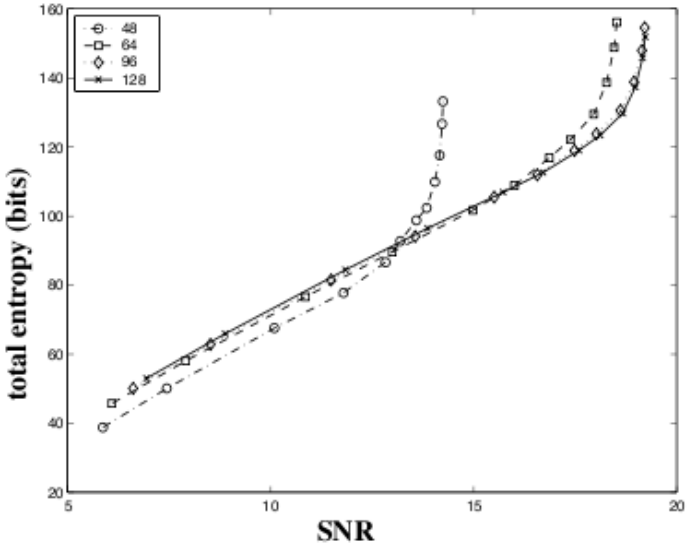
1660s — first scientific journals
1950s — wide-spread adoption of peer review
1960s — open access movement begins
1990s — open access becomes more prevalent

Statistics of Natural Images



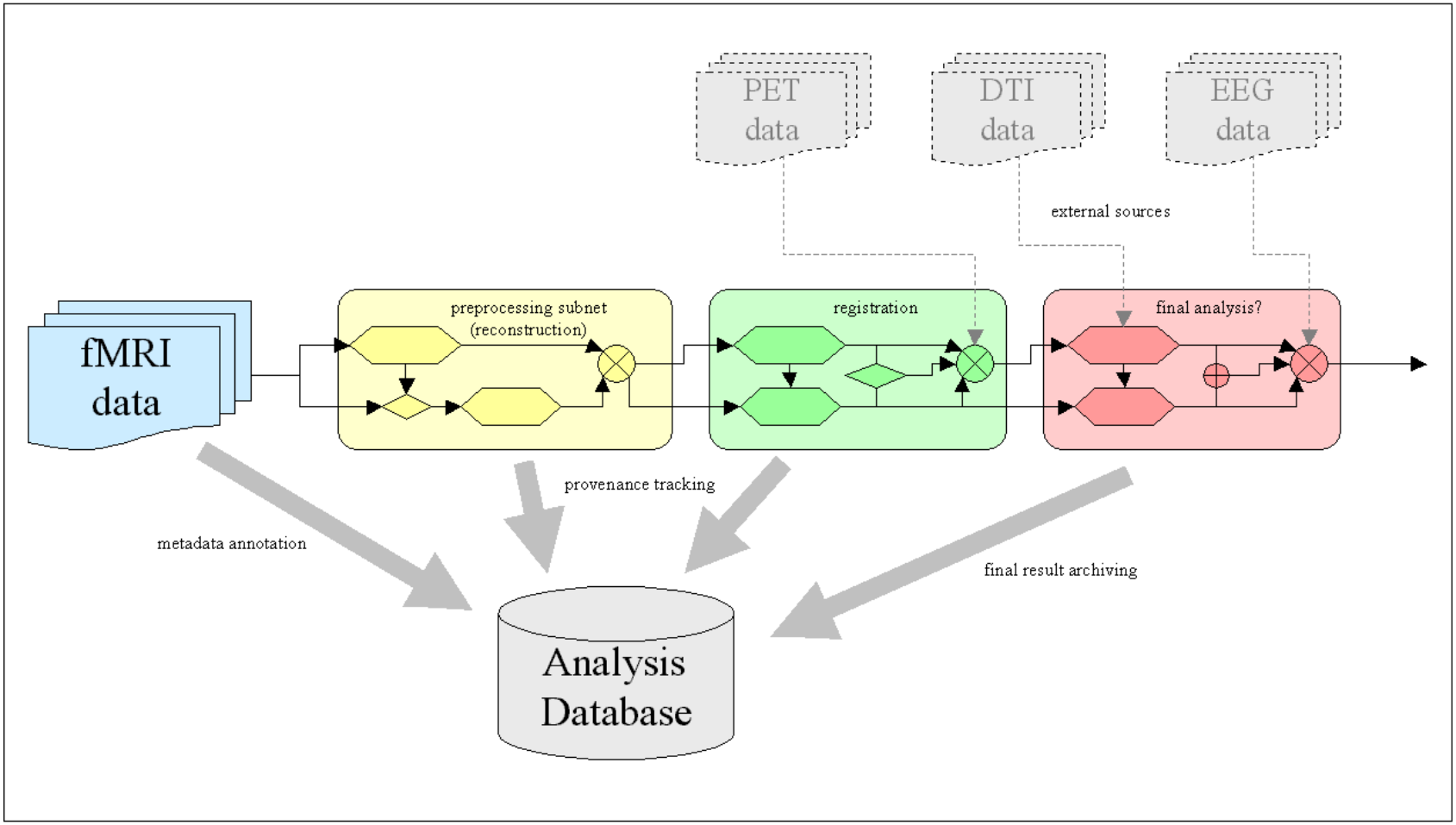






Berkeley's Brain Imaging Center

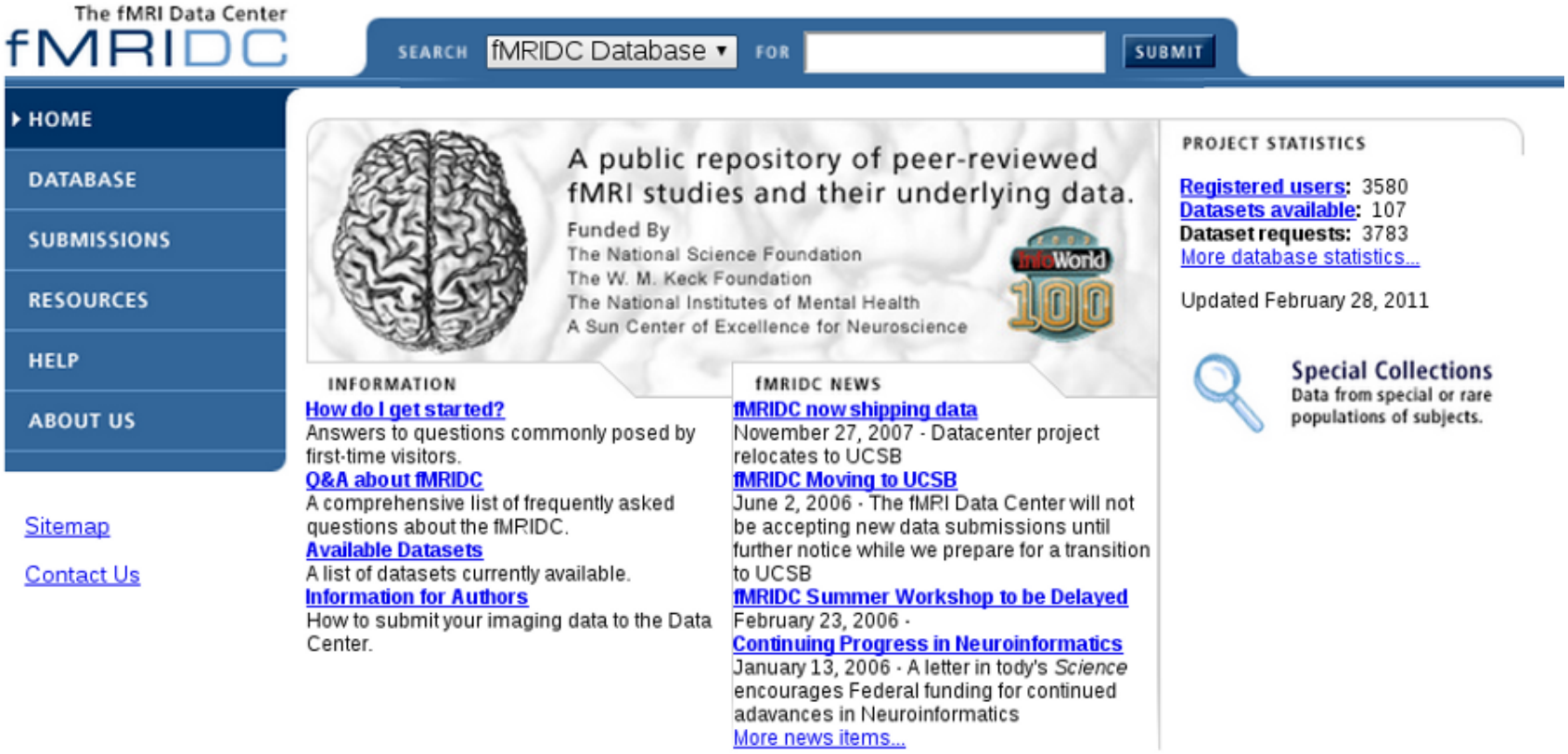
FMRI Data Analysis



Deep magic begins here...

- deep magic
- black magic
- shotgun debugging
- rain dance/waving the dead chicken

FMRI Data Center



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.

— Edsger Dijkstra, The Cruelty of Really Teaching Computer Science (1988)

Cargo cult science

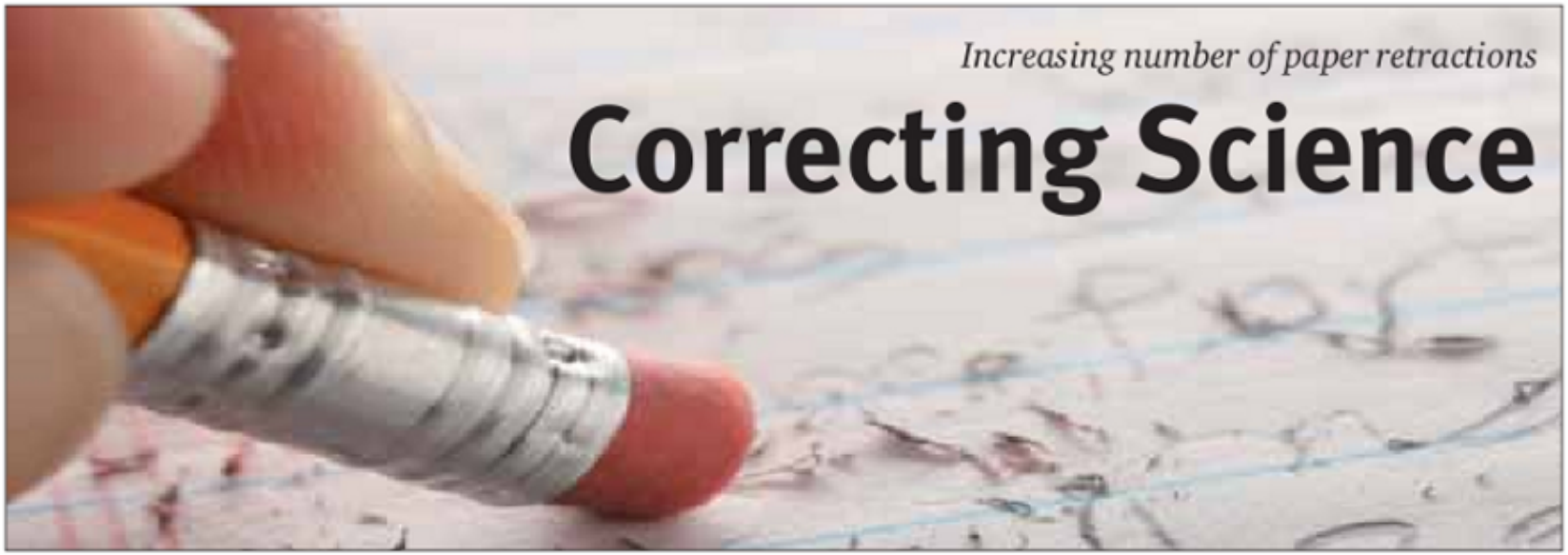
In summary, the idea is to try to give all of the information to help others to judge the value of your contribution; not just the information that leads to judgment in one particular direction or another.

— Richard Feynman, Cargo Cult Science (1974)

Publish or perish

- Software and data not shared
- Minimal publishable unit
- Not reproducible
- Not verifiable

High profile article retractions



Lately, it would seem that published research results have to be taken with a pinch of salt. It could very well happen that only a couple of weeks or months later those very same results are taken back or retracted. Has the search for truth temporarily lost its way?

As a curious scientist, you usually search on websites like *Nature News* or *Science Insider* to read about extraordinary discoveries or achievements in science. Recently, though, you are, more often than not, rather unpleasantly surprised with headlines, such as "Gene therapy researcher retracts four papers", "Nobel-winning brain researcher retracts two papers" or "Highly cited Harvard stem cell scientist retracts *Nature* paper".

But let's go back to the more serious cases! The first retractions ever issued by scientific journals date back to the 1970s. Back then it was a clear misconduct case, which led to several papers being retracted. A PhD student, first in the lab of Charles Rowe at Birmingham University, then in the lab of Bernd Hamprecht at the Max Planck Institute for Biochemistry in Munich, confessed to have simply "invented" his results. Normally, your

"...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)

What is needed for reproducible research?

- Data
- Code
- Provenance
- Details of non-automated steps

Software crisis and silver bullets

The major cause of the software crisis is that the machines have become several orders of magnitude more powerful! To put it quite bluntly: as long as there were no machines, programming was no problem at all; when we had a few weak computers, programming became a mild problem, and now we have gigantic computers, programming has become an equally gigantic problem.

— Edsger Dijkstra, The Humble Programmer (1972)

What tools are necessary for reproducible research?

- version control/collaboration tools
- better scientific programming languages/libraries
- data/code archives/libraries

How do we train the next generation of scientists?

- computational literacy
- software engineering
- scientific integrity

What role should journals play?

- supplemental material
- peer review

How will all this impact scientists?

- funding
- academic merit review

Questions

... when all have wandered from the path, quitting it entirely, and deserting experience, or involving themselves in mazes, and wandering about...

— Francis Bacon, Novum Organum (1620)