ASEG Workshop

August 15, 2013
Melbourne, Australia
Software as a Research Tool

Connections

Power
Outline

• State of the Madagascar project

• Madagascar philosophy
“Abandoning the habit of secrecy in favor of process transparency and peer review was the crucial step by which alchemy became chemistry. In the same way, it is beginning to appear that open-source development may signal the long-awaited maturation of software development as a discipline.”  

Eric Raymond
In a Nutshell, Madagascar…

…has had 9,517 commits made by 70 contributors representing 629,407 lines of code

…is mostly written in C

with an average number of source code comments

…has a well established, mature codebase

maintained by a very large development team

with increasing Y-O-Y commits

…took an estimated 170 years of effort

starting with its first commit in May, 2003

ending with its most recent commit about 2 hours ago
Contributors

Stable Versions

- 0.9 – 06/2006
- 1.0 - 07/2010
- 1.5 – 07/2013
Madagascar Schools

- 2006 – Vancouver, Canada
- 2007 – Austin, Texas, USA
- 2008 – Golden, Colorado, USA
- 2009 – Delft, Netherlands; Bahia, Brazil
- 2010 – Houston, Texas, USA
- 2011 – Beijing, China
- 2012 – Austin, Texas, USA
- 2013 – Melbourne, Australia
Migration Gallery
Future Goals

• Release 1.6
  ◦ Migration gallery

• Release 2.0
  ◦ Parallel computing examples
  ◦ Field data processing examples
Outline

- State of the Madagascar project
- Madagascar philosophy
Research Pyramid

Implement

Test

Publish

Implement
Research Pyramid

- 1,000 Programs
- 500 Workflows
- 5,000 Figures
- 150 Papers

Languages and Tools:
- LaTeX
- Python
- SCons
- Unix
- C
What is Science?
“Science is the belief in the ignorance of experts.”

Richard Feynman

What is Science? (1966)
What is Science?

Science is the systematic enterprise of gathering knowledge about the universe and organizing and condensing that knowledge into testable laws and theories. The success and credibility of science are anchored in the willingness of scientists to expose their ideas and results to independent testing and replication by other scientists. This requires the complete and open exchange of data, procedures and materials.
Claerbout’s Principle

“An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.”

(Buckheit and Donoho, 1995)
Reproducible Research

“It is a big chore for one researcher to reproduce the analysis and computational results of another [...] I discovered that this problem has a simple technological solution: illustrations (figures) in a technical document are made by programs and command scripts that along with required data should be linked to the document itself [...] This is hardly any extra work for the author, but it makes the document much more valuable to readers who possess the document in electronic form because they are able to track down the computations that lead to the illustrations.”

(Claerbout, 1991)
Reproducible Research

Addressing the Need for Data and Code Sharing in Computational Science

By the Yale Law School Roundtable on Data and Code Sharing