Opening the Science

Using a Mozilla Study Group to help grow an open science culture

Murray Cadzow, Tom Kelly, James Boocock, Mik Black
Department of Biochemistry, University of Otago

9 February 2016

Overview

- Background
  - Where are we coming from?
  - What are our needs?
- What are we doing?
  - Software Carpentry
  - Mozilla Study Group
  - Research Bazaar
- What next?



Pre-background: meet your speaker

- My background is statistics
  - collaborative genomics research
  - Linux and R for data analysis
  - research group with eclectic skill set
- Infrastructure initiatives: building resources
  - NZGL
  - NeSI

Where were we in 2013?

- Small research group of relatively junior graduate students and research assistants
  - mix of computer science, statistics and biology/genetics backgrounds.
  - some had come through the NeSI Summer of eResearch programmes in 2011 and 2012.
- Similar/related research projects, and common needs in terms of skills development

What were/are our needs?

- Unix shell
  - general usage for data manipulation
  - scripting for basic automation
- R
  - general statistical analysis (esp. linear models)
  - genetics/genomics data analysis techniques
  - data visualisation
  - reproducible research
- HPC cluster access
  - simulations and permutations/resampling
  - embarrassingly parallel…

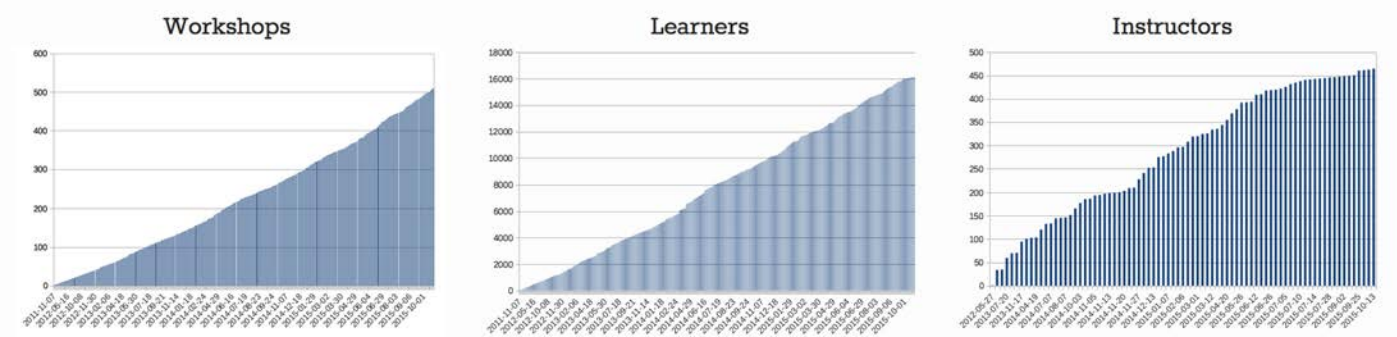
Sounds (vaguely) familiar?

- Software Carpentry (SWC):
  - Unix Shell
  - R/Python
  - Git
  - MySQL
- First NZ SWC workshop at eResearch NZ, June 2013 (Canterbury)
  - Unix Shell, version control, Python
  - we missed the boat on that one, luckily NeSI didn't…

Software Carpentry

"Since 1998, Software Carpentry has been teaching researchers in science, engineering, medicine, and related disciplines the computing skills they need to get more done in less time and with less pain."

- Trained instructors
- Comprehensive lessons



http://software-carpentry.org

Also in June 2013

- Mozilla Science Lab announced:
  - https://www.mozillascience.org
  - "The Mozilla Science Lab is an initiative of the Mozilla Foundation exploring how the power of open source can change the way science is done on the web."
- Director Kaitlin Thaney
  - Keen advocate for Open Science
  - Keynote speaker at eResearch NZ in June 2014 (Hamilton): NeSI is on the ball again…

mozilla Science Lab

More developments in 2014

- In May 2014, the first "Data Carpentry bootcamp" is taught:
  - "Data Carpentry develops and teaches workshops on the fundamental data skills needed to conduct research."
  - sibling organisation to Software Carpentry
  - http://www.datacarpentry.org/
- In October 2014, the Software Carpentry Foundation is announced. Partners include:
  - Mozilla Science Lab
  - Software Sustainability Institute
  - Lawrence Berkeley National Laboratory
  - NumFocus

What were we doing at this point?

- I went on sabbatical…
- Then from mid-2014, lab meetings/workshops on:
  - dplyr/tidyr
  - reproducible research
  - ggplot2
  - ggvis
  - shiny
  - Bayesian modelling with JAGS
  - genomic data visualisation
- That was exhausting…

And then came SWC and ResBaz

- SWC instructor training, Melbourne, Jan 2015
  - Led by Bill Mills (MSL community manager)
  - Two groups members (Tom Kelly and James Boocock - thanks NeSI) and myself attended.
  - Strong NeSI presence at training, followed by SWC affiliate status
- Inaugural Research Bazaar (ResBaz), Melbourne, Feb 2015
  - postgraduate students and early/mid-career researchers
  - SWC training + many other workshops
- NeSI ran NZ SWC workshops soon after (Auckland and Christchurch)

Saved by SYSKA

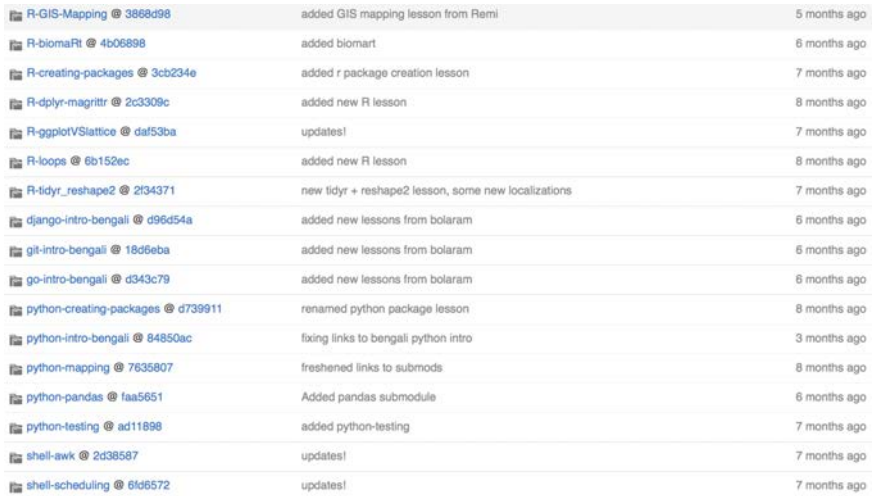
- We now had three trained SWC instructors in our extended research group
  - the students are taking over!
  - the senior students were now able to train others… and so were the junior students
- Time for SYSKA: Sh*t You Should Know About
  - "hosted" by Murray Cadzow
  - rotating weekly slots
  - short (20-30 minute) presentation (by student) to group on something useful or topical
  - Python Tricks, Vim vs Emacs, NeSI HPC, Shell tricks, RMySQL, LaTeX, dplyr (again)…

Soon, everyone wants to play

- SWC workshop at Otago in June 2015
  - Training for many of the wet-lab genetics students within our extended research group (and many others).
  - Merriman Lab decides to move to a Git-based system for lab documentation
  - Suddenly SYSKA is popular
- Time to grow:
  - Initially expand group to encompass students doing "wet-lab science" within our extended research group
  - Aim to create a stable learning environment before further expansion.

Mozilla Study Groups

- Announced in April 2015 by Bill Mills:
  - skill sharing
  - idea discovery
  - community support
- Lots of introductory lessons:



R-GIS-Mapping @ 3868d98	added GIS mapping lesson from Remi	5 months ago
R-biomaRt @ 4b06898	added biomart	6 months ago
R-creating-packages @ 3cb234e	added r package creation lesson	7 months ago
R-dplyr-magrittr @ 2c3309c	added new R lesson	8 months ago
R-ggplotVSlattice @ daf53ba	updates!	7 months ago
R-loops @ 6b152ec	added new R lesson	8 months ago
R-tidyr_reshape2 @ 2f34371	new tidyr + reshape2 lesson, some new localizations	7 months ago
django-intro-bengali @ d96d54a	added new lessons from bolaram	6 months ago
git-intro-bengali @ 18d6eba	added new lessons from bolaram	6 months ago
go-intro-bengali @ d343c79	added new lessons from bolaram	6 months ago
python-creating-packages @ d739911	renamed python package lesson	8 months ago
python-intro-bengali @ 84850ac	fixing links to bengali python intro	3 months ago
python-mapping @ 7635807	freshened links to submods	8 months ago
python-pandas @ faa5651	Added pandas submodule	6 months ago
python-testing @ ad11898	added python-testing	7 months ago
shell-awk @ 2d38587	updates!	7 months ago
shell-scheduling @ 6fd6572	updates!	7 months ago

https://github.com/mozillascience/studyGroupLessons

Handy instructions

https://mozillascience.github.io/studyGroupHandbook/

Even we could follow them:

Mozilla Study Groups September Update

Posted on September 4, 2015 by Bill Mills

Welcome New Study Groups

Welcome to our newest Study Groups in Otago, New Zealand and North West University, South Africa! More Study Groups are spooling up around the world; we'll note them here and add them to the Study Group Map as their first events come on line.

https://www.mozillascience.org/mozilla-study-groups-september-update

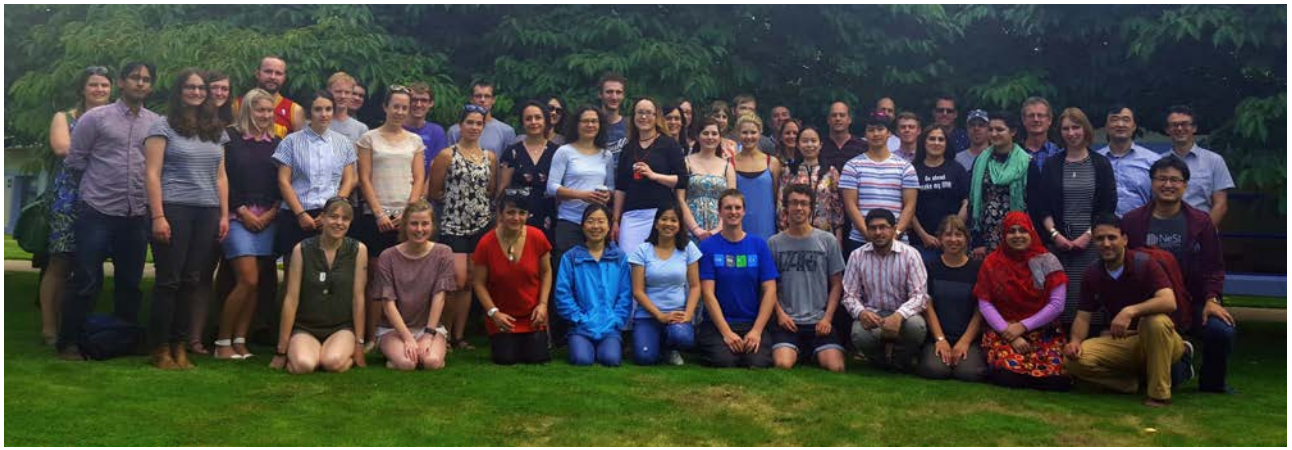
Fun times!

Mozilla Study Group format

- Fortnightly meetings
- 4 session rotating format:
  - 2 weeks of nominated topics: hands-on coding
  - 1 week of hacky hour
  - 1 week of 5x5 lightning SYSKA
- Lightning SYSKA!
  - 5 presenters, 5 minutes each
  - Present a cool topic
  - Use topics/interest to decide content for future lessons

Otago Research Bazaar 2016 (Feb 2-4)

- ~80 attendees: 2/3 postgraduate students, 1/3 early-mid career researchers.
  - 70/30 split: SWC vs intermediate lessons
- Strongly supported by NeSI, and University of Otago



Where are we now?

- Senior group members are competent SWC instructors or helpers
- Good study group attendance
  - Extended research group is becoming competent with core digital research tools (Shell, R, Git)
  - We have a solid collection of training materials (both general, and domain-specific)
  - Presentations are hands-on: major advantage (and a good step forward)
- Reproducible Research and Open Science concepts/techniques are starting to be used more frequently.

What are the challenges?

- Academic environment: student/staff turnover (although this is why we NEED SWC and MSG)
    - sustainability
    - research focus can change
- Research project focus: teaching others isn't what students (or research assistants) are being paid to do
    - good for CVs
    - not necessarily good for projects
    - need to ensure appropriate balance: graduating students is still our core business
- Resourcing: institutional/departmental buy-in

What next?

- Expand the study group(s)
  - open existing MSG to ResBaz attendees (and others)
  - encourage new groups to form as needed
- More Software/Data Carpentry training
  - 400 level Biochemistry/Genetics/Statistics students (late Feb 2016)
  - work SWC material into 3rd and 4th year curriculum (Genetics and Statistics)
  - intermediate/advanced SWC workshop mid-2016?
- Potential flow-on into increased NeSI HPC access

Summary

- Over the past year we have transformed the way in which members of our research group gain new digital skills
  - student/peer led
  - increased emphasis on hands-on training
  - links to international training initiatives and community resources
- Software Carpentry and Mozilla Study Group have been major steps forward for us
  - extended training outside immediate research group
  - created a learning community, and an associated support network
  - powerful framework for delivering information/training on new tools and techniques