The Carpentries
Are we using the right tools for data driven research?

Mistaken Identifiers: Gene name errors can be introduced inadvertently when using Excel in bioinformatics

Barry R Zeeberg¹, Joseph Riss¹, David W Kane, Kimberly J Bussey, Edward Uchio, W Marston Linehan, J Carl Barrett and John N Weinstein II

¹ Contributed equally

Received: 05 March 2004 | Accepted: 23 June 2004 | Published: 23 June 2004

Automatic conversion of gene symbols to dates and floating-point numbers is a problematic feature of Excel software.
— Zieman et al.
What are the right tools?

With growing rate of data accumulation, there is an acute need for all researchers to learn about:

— Repeating common tasks
— cleaning/reading/processing data
— sharing code for common tools and methods
— collaborating with code
Every discipline is accumulating data at unprecedented rates.

Skills lag far behind needs in many most disciplines.

These skills are incredibly marketable in both science & business careers.
Technical consulting doesn't scale
Online training and MOOCs can leave people frustrated, without a community to help them.

The Carpentries
Our approach is to build community and capacity through peer instruction.
What does it take to have a workshop with IMPACT

The Carpentries
Quality LESSONS
Motivated LEARNERS
Trained INSTRUCTORS
Lead to workshops with IMPACT
It's working

$n = 1,350$ on all 7 continents yes, even Antarctica
Impactful Workshops

- Trained instructors who are themselves researchers
- Community developed lessons
- Passion to apply the skills and spread the community
- A mission to change the culture of how research is done
Instructors
Instructor training

— Instructor training prepares researchers to impactfully teach technical skills:
  — live coding
  — take into account learner's background
  — reducing cognitive load
  — improve self as an instructor over time
  — give and receive constructive feedback

The Carpentries
Teaching as performance art

— Excitement
— Engagement
— Passion for the topic
— Improv
— Lessons are a loosely sketched script
Lessons

The Carpentries
Preparing the unconscious mind

— Two day constraint
— N.R.F. Maier's two cords experiment
— Problem solving skills can be inaccessible to the conscious mind
— You can't teach all the things
Jugyokenkyu - lesson study

- Coordinated collaboration, testing and continuous improvement of lessons.
- Collaboration on lessons, conversation about teaching of lessons
- Instructor community that discusses the ongoing improvement of a lessons
- Kaizen of teaching

The Carpentries
Workshops
<table>
<thead>
<tr>
<th>Lesson</th>
<th>Site</th>
<th>Repository</th>
<th>Reference</th>
<th>Maintainer(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Unix Shell</td>
<td></td>
<td></td>
<td></td>
<td>Gabriel Devenyi, Ashwin Srinath</td>
</tr>
<tr>
<td>Version Control with Git</td>
<td></td>
<td></td>
<td></td>
<td>Ivan Gonzalez, Daisie Huang</td>
</tr>
<tr>
<td>Version Control with Mercurial</td>
<td></td>
<td></td>
<td></td>
<td>Doug Latornell</td>
</tr>
<tr>
<td>Using Databases and SQL</td>
<td></td>
<td></td>
<td></td>
<td>Abigail Cabunoc Mayes, Sheldon McKay</td>
</tr>
<tr>
<td>Programming with Python</td>
<td></td>
<td></td>
<td></td>
<td>Trevor Bekolay, Valentina Staneva</td>
</tr>
<tr>
<td>Programming with R</td>
<td></td>
<td></td>
<td></td>
<td>Daniel Chen, Harriet Dashnow</td>
</tr>
<tr>
<td>R for Reproducible Scientific Analysis</td>
<td></td>
<td></td>
<td></td>
<td>Thomas Wright, Naupaka Zimmerman</td>
</tr>
<tr>
<td>Programming with MATLAB</td>
<td></td>
<td></td>
<td></td>
<td>Isabell Kiral-Kornek, Ashwin Srinath</td>
</tr>
<tr>
<td>Automation and Make</td>
<td></td>
<td></td>
<td></td>
<td>Gerard Capes</td>
</tr>
<tr>
<td>Instructor Training</td>
<td></td>
<td></td>
<td></td>
<td>Greg Wilson</td>
</tr>
</tbody>
</table>

The Carpentries
Community

— Cultivating a community and culture requires ongoing and intentional work
  
  — Code of Conduct and enforcement

  — Active mentoring of instructors

  — Collaborative lesson development
Getting Started

- Run a local workshop
- Build a support coalition
  - Libraries, Deans, Chancellors, Department chairs
- Train local instructors
- Build it into your organization
- Connect to the global community
Learners
Our workshops are welcoming and relevant to all disciplines

<table>
<thead>
<tr>
<th>Research Domain</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Sciences (Genetics, genomics, bioinformatics)</td>
<td>24.9</td>
</tr>
<tr>
<td>Life Science - Organismal/systems (ecology, botany, zoology, microbiology, neuroscience)</td>
<td>24.0</td>
</tr>
<tr>
<td>Planetary sciences (geology, climatology, oceanography, etc.)</td>
<td>6.6</td>
</tr>
<tr>
<td>Mathematics/statistics</td>
<td>6.0</td>
</tr>
<tr>
<td>Physics</td>
<td>5.8</td>
</tr>
<tr>
<td>Civil, mechanical, chemical, or nuclear engineering</td>
<td>4.5</td>
</tr>
<tr>
<td>Medicine and/or Pharmacy</td>
<td>4.3</td>
</tr>
<tr>
<td>Chemistry</td>
<td>4.0</td>
</tr>
<tr>
<td>Social sciences</td>
<td>4.0</td>
</tr>
<tr>
<td>Library and information science</td>
<td>3.2</td>
</tr>
<tr>
<td>Economics/business</td>
<td>2.6</td>
</tr>
<tr>
<td>Humanities</td>
<td>2.6</td>
</tr>
<tr>
<td>Psychology</td>
<td>2.4</td>
</tr>
<tr>
<td>Education</td>
<td>2.1</td>
</tr>
<tr>
<td>High performance computing</td>
<td>2.1</td>
</tr>
<tr>
<td>Space sciences</td>
<td>0.9</td>
</tr>
</tbody>
</table>
We achieve a very good gender balance
Long-term outcomes for learners

The Carpentries
Adapting to learners

How did you perceive the pace of the workshop?

Answered: 2,064  Skipped: 298

- Too slow
- Slightly slow
- Just right
- Slightly fast
- Too fast
Balancing demonstration and application
Building confidence

Respondents are More Confident Post-Workshop

- I’m less confident now: 6 respondents
- I’m equally confident now: 93 respondents
- I’m more confident now: 333 respondents

Change in Confidence

% Respondents
Sentiment analysis of "Do you have specific comments about the instructors or helpers?"
Self-assessment of workshop impact on practices

Respondents' Programming Usage Increased

<table>
<thead>
<tr>
<th>Programming Usage</th>
<th>Before Workshop</th>
<th>After Workshop</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have not been using</td>
<td>75</td>
<td>22</td>
</tr>
<tr>
<td>tools like these.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than once per year</td>
<td>52</td>
<td>18</td>
</tr>
<tr>
<td>Several times per year</td>
<td>61</td>
<td>48</td>
</tr>
<tr>
<td>Monthly</td>
<td>70</td>
<td>48</td>
</tr>
<tr>
<td>Weekly</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>Daily</td>
<td>119</td>
<td>119</td>
</tr>
</tbody>
</table>

The Carpentries
Collaborative lesson development
caseyyoungflesh commented 10 days ago

Changed some of the language to be more clear - changes in sentence structure, defining jargon. Also changed the use of the 'tape' analogy to 'recording' (everyone knows what a recording is!) in reference to issue #461.

Add more commits by pushing to the youngflesh-branch branch on caseyyoungflesh/git-novice.

This branch has no conflicts with the base branch
Merging can be performed automatically.

You can also open this in GitHub Desktop or view command line instructions.
Version control systems start with a base version of the document and

-then **save just the changes you made** at each step of the way. You can
-think of it as a tape: if you rewind the tape and start at the base
-document, then you can play back each change and end up with your
-**latest** version.

+then **record changes you make** each step of the way. You can
+think of it as a recording of your progress: you can rewind to start at the base
+document, play back each change **you made**, eventually arriving at your
+**more recent** version.
How to get your organization involved?

http://software-carpentry.org/membership/
# Partnership Tiers

<table>
<thead>
<tr>
<th></th>
<th>Bronze</th>
<th>Silver</th>
<th>Gold</th>
<th>Platinum</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Coordinated Workshops</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>negotiable</td>
</tr>
<tr>
<td>Discount for additional</td>
<td>20%</td>
<td>33%</td>
<td>50%</td>
<td>negotiable</td>
</tr>
<tr>
<td>coordinated workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-organized workshops at</td>
<td>no-charge</td>
<td>no-charge</td>
<td>no-charge</td>
<td>no-charge</td>
</tr>
<tr>
<td>partner organization **</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of instructors</td>
<td>0</td>
<td>6 online</td>
<td>15 with possibility for in-person training event</td>
<td>negotiable</td>
</tr>
<tr>
<td>trained ***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seat on the SCF Advisory Board</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Train an in-house instructor</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Available</td>
</tr>
<tr>
<td>trainer at partner org</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesson development services</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Available</td>
</tr>
<tr>
<td>Membership Dues (annual)</td>
<td>$5,000</td>
<td>$7,500</td>
<td>$15,000</td>
<td>Contact us</td>
</tr>
</tbody>
</table>
What you Support with Membership

— Instructor training
— Instructor mentorship
— Workshop management database
— Curriculum Development and Maintenance
— Metrics / Assessment infrastructure
— Global communities of practice
CarpentryCon

— When: May 30-June 1, 2018
— Where: Dublin, Ireland
— Moreinfo: carpentrycon.org
Introduction
- Introduction
- Building or improving your own local community
- What is a community of practice?
- Codes of Conduct

Initiation
- Getting started

Activities
- Workshops
- Help Sessions
- H hacky Hour
- Carpentries Study Groups
- Events
- Research Bazaar (ResBaz)
- CarpentryCon
- SatRdays
- THATCamp
- International Events
- Vendor Training

Tools
- Communications

Support Structures
- Support Structures
- Sourcing funding for workshops

Appendix
- The Carpentries
- Appendix

Community Cookbook
https://cookbook.carpentries.org
Thank You to our supporters!

55 Organizations in 10 countries are currently members.

1300 workshops have been taught since 2012 for 35,000 learners by our 1200 instructors from 39 countries.
Other Related Initiatives

— coderefinery.org
— Research Software Engineer Association
— Library Carpentry
— Mozilla Science Lab

Many of our community members and member organizations are active in the above initiatives as well.
Thank You!

Jonah Duckles
Director of Membership
Software Carpentry
jduckles@carpentries.org
@jduckles

Become a member organization:
http://software-carpentry.org/membership/

Get involved: http://software-carpentry.org/join/

Twitter: @swcarpentry @jduckles
RESERVE SLIDES
Mastery Rubric

- Naive
- Intermediate
- Expert

KSAs

KSA = Knowledge Skills & Abilities
Mastery Rubric

- Naive
- Intermediate
- Expert

KSAs
KSAs
KSAs

Development Trajectory
Mastery Rubric

- **Naive**
  - Description of KSA as comprehended at that level of mastery.

- **Intermediate**
  - Description of KSA as comprehended at that level of mastery.

- **Expert**
  - Description of KSA as comprehended at that level of mastery.

**Intervention:**
- Workshop
- Mentoring
- Practice
- Other
Mastery Rubric

- Naive
- Intermediate
- Expert

- Unix Shell
- Programming
- Version Control
- Visualization

SWC Today

The Carpentries
Rochelle Tractenberg's paper about a Mastery Rubric for Statistical Literacy:

http://www.mdpi.com/2227-7102/7/1/3/pdf
How to get yourself involved?

http://software-carpentry.org/join/

— newsletter
— mailing list
— github
— open calls for instructor training