Software Carpentry: teaching computing skills to researchers

Iñigo Aldazabal Mensa

Centro de Física de Materiales (CSIC-UPV/EHU)
Software & Data Carpentry Instructor

HPCKP’16 Barcelona, 27-28th April 2016

1inigo.aldazabal@ehu.eus
1. **Software Carpentry**
   - About
   - Lessons
   - Workshops
   - Instructors & helpers

2. **Running a workshop**
   - Requesting a workshop
   - Operations Guide
   - San Sebastian 2015 SC Workshop
1 Software Carpentry
- About
- Lessons
- Workshops
- Instructors & helpers

2 Running a workshop
- Requesting a workshop
- Operations Guide
- San Sebastian 2015 SC Workshop
Somewhere back in the 2000’s...
Overview

Many scientists and engineers spend much of their lives programming, but only a handful have ever been taught how to do this well. As a result, they spend their time wrestling with software, instead of doing research, but have no idea how reliable or efficient their programs are.

This course is an intensive introduction to basic software development practices for scientists and engineers that can reduce the time they spend programming by 20-25%. All of the material is open source: it may be used freely by anyone for educational or commercial purposes, and research groups in academia and industry are actively encouraged to adapt it to their needs.

Upcoming Offerings

January-April 2007, University of Toronto. For more information, contact Greg Wilson.

Acknowledgments

This work has been made possible by the Python Software Foundation, and the University of Toronto. See Acknowledgments for more details, and License for terms of re-use.

Download

Via Subversion: http://svn.scipy.org/svn/swc (You probably only want trunk or unstable, but you might be interested in recent releases.)

Lectures

- License
- Introduction
- Shell Basics
- More Shell
- Version Control
- Automated Builds
- Basic Scripting
- Strings, Lists, and Files
- Functions and Libraries
- Style
- Quality Assurance
- Sets, Dictionaries, and Complexity
- Debugging
- Object-Oriented Programming
- More on Objects
- Unit Testing
- Regular Expressions
- Binary Data
- XML
- Relational Databases
- For more learning...
Introduction

- This course will teach you how to design, build, maintain, and share programs efficiently
- Focus on the equivalent of good laboratory technique
  - The 20% of ideas that account for 80% of real world use
  - Software carpentry, rather than software engineering
    - About putting an extension on the house, rather than building the Channel Tunnel
- Everything that will make you more productive will also improve the quality of what you build
  - Help computational science deserve the second half of its name
Version 1122 (Fri Sep 15 09:37:42 2006)

Lectures

- License
- Introduction
- Shell Basics
- More Shell
- Version Control
- Automated Builds
- Basic Scripting
- Strings, Lists, and Files
- Functions and Libraries
- Style

- Quality Assurance
- Sets, Dictionaries, and Complexity
- Debugging
- Object-Oriented Programming
- More on Objects
- Unit Testing
- Regular Expressions
- Binary Data
- XML
- Relational Databases
- Spreadsheets
- Integration
- Web Client Programming
- Web Server Programming
- Security
- The Development Process
Where's the Real Bottleneck in Scientific Computing?

Scientists would do well to pick up some tools widely used in the software industry

Gregory V. Wilson

American Scientist
Volume: 94 Number: 1 Page: 5
DOI: 10.1511/2006.1.5
link

When I first started doing computational science in 1986, a new generation of fast, cheap chips had just ushered in the current era of low-cost supercomputers, in which multiple processors work in parallel on a single problem. Suddenly, it seemed as though everyone who took number crunching seriously was rewriting his or her software to take advantage of these new machines. Sure, it hurt—the compilers that translated programs to run on parallel computers were flaky, debugging tools were nonexistent, and thinking about how to solve problems in parallel was often like trying to solve a thousand crossword puzzles at once—but the potential payoff seemed enormous. Many investigators were positive that within a few years, computer modeling would let scientists
Now in 2016...

software-carpentry.org
About Us

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. Our volunteer instructors have run hundreds of events for thousands of scientists in the past two and a half years. All of our lesson materials are freely reusable under the Creative Commons - Attribution license.

The Software Carpentry Foundation and its sibling organization Data Carpentry are members of NumFOCUS, a 501(c)3 non-profit incorporated in the United States.

Supporters
Software Carpentry is made possible by the generous support of our partners and by the hard work of our volunteers. We offer several levels of institutional engagement, and individuals can become instructors.

Workshops
You can host a workshop or attend one that someone else is hosting. Our code of conduct and operations guides describe how our workshops are run, and our answers other questions.

Conversations
You can send us email, join our mailing lists, read our blog, follow us on Twitter, or browse our repositories on GitHub.
Our Lessons

Our lessons are developed collaboratively on GitHub. You can check the status of each lesson on our dashboard, or view the nightly build. You may also enjoy Data Carpentry’s lessons, which focus on data organization, cleanup, analysis, and visualization.

All of our lessons are freely available under the Creative Commons - Attribution License. You may re-use and re-mix the material in any way you wish, without asking permission, provided you cite us as the original source (e.g., provide a link back to this website).
<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, Azalee Bostroem</td>
</tr>
<tr>
<td>Programming with R</td>
<td></td>
<td></td>
<td></td>
<td>John Blischak, Denis Haine</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>Mike Jackson</td>
</tr>
<tr>
<td>Instructor Training</td>
<td></td>
<td></td>
<td></td>
<td>Greg Wilson</td>
</tr>
</tbody>
</table>
Software Carpentry Worskshops

A Software Carpentry workshop is hands-on two-day event that covers the core skills needed to be productive in a small research team. Short tutorials alternate with practical exercises, and all instruction is done via live coding.
Software Carpentry Workshops core topics

- automating tasks using the Unix shell,
- structured programming in Python, R, or MATLAB, and
- version control using Git or Mercurial.

Software Carpentry Workshops prerequisites

- covers the core topics,
- has at least one certified Software Carpentry instructor teaching,
- runs the standardised pre- and post-workshop assessment forms and provides us with the results, and
- sends back summary information about attendees.
Workshop flow

- First half an hour check that everybody has everything correctly configured ([workshop webpage instructions]).
- The lessons are typed live by the instructor and the learners follow along. (Almost) no slides!
- Learners use green and red sticky notes to show progress or report problems.
- Helpers provide support when someone reports a problem ([red sticker]), so the class does not get interrupted.
- Learners are encouraged to collaborate with nearby learners.
- **Etherpad** collaborative online note taking editor is used for sharing links, resources or tips, asking questions, etc.
Software Carpentry Workshops members

- **The host**, who is the principal local contact for the workshop.
- **The instructors**, who present the tutorials and lead the practical exercises.
- **The lead instructor**, who is in charge of deciding what will be taught by whom and ensuring the teaching aspects of the workshop go smoothly.
- **Helpers**, who provide assistance during practical sessions, update and answer questions on the workshop Etherpad, and help with various practical matters.
- **The Software Carpentry administrator** advises and helps the host with workshop organization, matches hosts with instructors, sets up registration and assessments, etc.
- **Learners** come to the workshop to learn.
Software Carpentry Worskshops members

- The **host**, who is the principal local contact for the workshop.
- The **instructors**, who present the tutorials and lead the practical exercises.
- The **lead instructor**, who is in charge of deciding what will be taught by whom and ensuring the teaching aspects of the workshop go smoothly.
- **Helpers**, who provide assistance during practical sessions, update and answer questions on the workshop Etherpad, and help with various practical matters.
- The **Software Carpentry administrator** advises and helps the host with workshop organization, matches hosts with instructors, sets up registration and assessments, etc.
- **Learners** come to the workshop to learn.
Instructors

The instructors often are scientists themselves, and many are workshop alumni, who have gone through a free online training course specifically tailored at people who would like to teach Software Carpentry. All of the instructors are volunteers, so the host only pays for their travel and accommodation, not their time.
Helpers

Helpers provide assistance to workshop learners during practical sessions. They can be alumni of previous workshops or locals who have the right skills to help out.

Their main job is to help the attendees during practical sessions. Helping at a workshop is also an opportunity to learn the material well enough to teach it yourself in future.

Helpers are the unsung heroes of Software Carpentry. They need to have not only technical competence but also the interpersonal skills to spot people who are confused and help them get unstuck.

Helpers are vital to keeping workshops flowing smoothly.
Software Carpentry

1 Software Carpentry
   - About
   - Lessons
   - Workshops
   - Instructors & helpers

2 Running a workshop
   - Requesting a workshop
   - Operations Guide
   - San Sebastian 2015 SC Workshop
San Sebastian Software Carpentry Workshop - 15-16th June

Third floor room 3.4
If you consider hosting a workshop, think about:

- **Where it will be:** Is a good venue available? Will it cost anything to book?
- **When it will be:** Think about what makes sense for you and your audience. Be flexible, because instructor availability at our end will be a factor too.
- **How you will fund it:** The cost of running a workshop includes instructors’ travel and lodging, the cost of booking the room, catering, and any donation you would like to make to help keep Software Carpentry going.
- **Who the audience will be:** Will the workshop be open to people in your lab, in your department, at your university or company, attending a conference, or anyone nearby? How many learners can you accommodate?
A workshop is requested by fulfilling an online form. Once a site requests a workshop, Software Carpentry Foundation locates instructors that can come teach. All of the instructors are volunteers, so the host only pays for their travel and accommodation, not their time.

The host also pays an administrative fee to the Software Carpentry Foundation which is used to cover the organizational and development work needed to keep things running.
The workshop Operators Guide provide detailed instructions and checklists for everybody involved in a workshop.

It also has a extensive description of the story of a workshop and general tips and guidelines.

Just follow it and everything will run smoothly!
Operations Guide: Host’s role

- Coordinate dates, travel and accommodation for the instructors
- Decide, together with instructors, workshop content
- Find the helpers
- Setup registration and manage it (acceptances, cancellations, etc)
- Book an adequate room (plugs, networks, space,...)
- Book catering if needed
- Check that all necessary gear is available
- Advertise the workshop
- Manage social events
- ...

(See the full checklist)
Operations Guide: Helpers

Helpers provide assistance to workshop learners during practical sessions. Helpers are typically not reimbursed for travel expenses or incidentals.

- Read over the material the instructors will cover in the sessions you’re helping with.
- Work through the setup instructions for the workshop: the first thing you will encounter on the first day will be setup problems.
- During exercises, walk around the room and look for learners who need help.
- If you can’t answer an attendee’s question, tell them so and ask an instructor for help. Do not dive into a 20-minute debugging session.
Materials Physics Center - University of the Basque Country

Jun 15-16, 2015
9:00 am - 5:30 pm

Instructors: Luis Pedro Coelho, Jean-Christophe Leyder

Helpers: Iñigo Aldazabal, Alex Savio, Oier Echaniz, Diego Lasa, Garikoitz Aguirregabiria, Ivor Loncaric

General Information
San Sebastian 2015 SC Workshop

- 1 host (Iñigo Aldazabal)
- 2 Instructors (Luis Pedro Coelho, Jean-Christophe Leyder)
- 6 Helpers (Iñigo Aldazabal, Alex Savio, Oier Echaniz, Diego Lasa, Garikoitz Aguirregabiria, Ivor Loncaric)
- 44 Learners (20 CFM & related, 24 University & other research centers)
- 2 Sponsors (Bull/Atos, Python Software Foundation)
- Topics: bash, git and Python
San Sebastian Software Carpentry Workshop  15-16th June 2015

WELCOME!

WIFI: EHU-wGuest

Go to eg. www.google.com in your web browser and use the credentials

scwss15 / softwarecarpentry

Etherpad:  http://swc-ss.etherpad.mozilla.org/3

Use the power plug on your left (except the first column)
- content (very useful)
- level of language (very simple)
- straight forward for beginners
- assistance (good!)

- timetable
- help from assistants
- good instructors

1. making figures
2. numpy
3. exercises

1. numpy user-friendly
2. exercises during theory expl.
3. plotting

- picture
- plotting data
- running scripts and realizing that they actually work,

* The lectures were easy to follow and understand.
* The assistants were very helpful and kind.
* The level of knowledge of the lecturers on the topic was good enough to handle the questions.

* The plot stuff was nice
* It was very well explained
* The helping people

Making plots
The teaching speed
The teachers
- Some things were too basic
- Food was quite bad
- Weather was horrible 😞

Slow pace with Shell lesson
Didn't complete more advanced Shell subjects
Didn't cover sharing with git

Rain, Sidseria, Programming 😁

1. The helpers should have known the presentations in advance
2. Homework could have helped to bring questions for day #2.
3. The rain.

Slow/We didn't
Suggestion: It was an introduction, we didn't teach how to use tools on our own.

I NEED MORE DAYS (MORE DATA)
FREE LAPTOPS (LOL)

No negative

Lunch
Maybe 1 day more?
San Sebastian 2016 SC Workshop!

Materials Physics Center - University of the Basque Country

Jun 27-29, 2016
9:00 am - 5:30 pm

**Instructors:** Iñigo Aldazabal
Mensa, Andrés Diaz-Gil

**Helpers:** Ivor Loncaric, Diego Lasa, Oier Echaniz, David de Sancho
Thank you for your attention!