



LIBER Webinar:
Library Carpentry
*– Teaching Data Science
Skills & Upcoming Instructor
Training*



WEBINAR HOST



Dr Birgit Schmidt
Göttingen State and University Library

Co-Chair of LIBER's *Research Data Management Working Group*; member of the *Digital Skills for Library Staff & Researchers Working Group*

bschmidt@sub.uni-goettingen.de

SPEAKERS



Chris Erdmann

Library Carpentry Community &
Development Director at The
Carpentries/UC3

Christopher.Erdmann@ucop.edu



Juliane Schneider

Team Lead/Lead Data Curator,
Harvard Catalyst

Juliane_Schneider@hms.harvard.edu



David Kane

Systems Librarian,
Waterford Institute of Technology

DKANE@wit.ie



NOTES

- **The webinar is being recorded.** A recording will be shared with all attendees, and posted on YouTube.
- **Slides can be downloaded on Zenodo.** Check the chatbox for the link.
- **Questions?** Put them in the chat box. We'll put them to the speakers, following the presentations.

LIBER Webinar: Library Carpentry – Teaching Data Science Skills & Upcoming Instructor Training

Chris Erdmann | Juliane Schneider | David Kane

April 8, 2019



University of California

CDL

California Digital Library



Skills and perspectives to work with software and data are increasingly important as we generate more data.

With the emergence of our ability to generate increasing amounts of data, research and work in almost every domain has a data and computational component, including the whole new field of data science.

REALIZING THE POTENTIAL OF DATA SCIENCE

**Final Report from the National Science Foundation Computer and
Information Science and Engineering Advisory Committee Data Science
Working Group**

Francine Berman and Rob Rutenbar, co-Chairs
Henrik Christensen, Susan Davidson, Deborah Estrin, Michael Franklin, Brent
Hailpern, Margaret Martonosi, Padma Raghavan, Victoria Stodden, Alex Szalay

December 2016



http://msdse.org/files/Creating_Institutional_Change.pdf



Providing researchers with the skills and competencies they need to practise Open Science

Open Science Skills Working Group Report

Written by the Working Group on Education and Skills under Open Science
July – 2017



Rise of data science initiatives in academia

From the Data Science Community Newsletter by Noren & Stenger:

Brigham Young University, Caltech, Carnegie Mellon, College of Charleston, Columbia, Cornell, Dartmouth UMass, George Mason University, Georgetown University, Georgia Tech, Harvard, Illinois Wesleyan University, Johns Hopkins, Mid America Nazarene University, MIT, Northeastern University, Northern Kentucky University, Northwestern, Northwestern College in Iowa, Ohio State University, Penn State University, Princeton, Purdue, Stanford, Tufts University, UC Berkeley, UC Davis, UC Irvine, UC Merced, UC Riverside, UC San Diego, UCLA, UIUC, University of Iowa, University of Michigan, University of Oregon, University of Pennsylvania, University of Rochester, University of San Francisco, University of Warwick, University of Washington, UT Austin, UW Madison, Vanderbilt University, Virginia Tech, Washington University in St. Louis, Middle Tennessee State University, NYU, Amherst College, Brown, CU Boulder, Duke, Illinois Institute of Technology, Lehigh University, Loyola University - Maryland, Rice University, SUNY at Stony Brook, UC Santa Barbara, UC Santa Cruz, UCSF, UMass Amherst, UNC - Wilmington, University of Vermont, University of Arizona, University of British Columbia, University of Chicago, University of Virginia, USC, Worcester Polytechnic, Yale

69% of business leaders in the United States will prefer job applicants with data skills by 2021.

23% of college and university leaders say their graduates will have those skills.

pwc.com/us/dsa-skills

Investing in America's data science and analytics talent

The case for action

April 2017



Importance of research software & training

- 92% of academics use research software
- 69% say that their research would not be practical without it
- 56% develop their own software (worryingly, 21% of those have no training in software development)

S.J. Hettrick et al, UK Research Software Survey 2014 [Data set].
Zenodo. <http://doi.org/10.5281/zenodo.14809>

Our path to better science in less time using open science tools

Reproducibility has long been a tenet of science but has been challenging to achieve—we learned this the hard way when our old approaches proved inadequate to efficiently reproduce our own work. Here we describe how several free software tools have fundamentally upgraded our approach to collaborative research, making our entire workflow more transparent and streamlined. By describing specific tools and how we incrementally began using them for the **Ocean Health Index** project, we hope to encourage others in the scientific community to do the same—so we can all produce better science in less time.

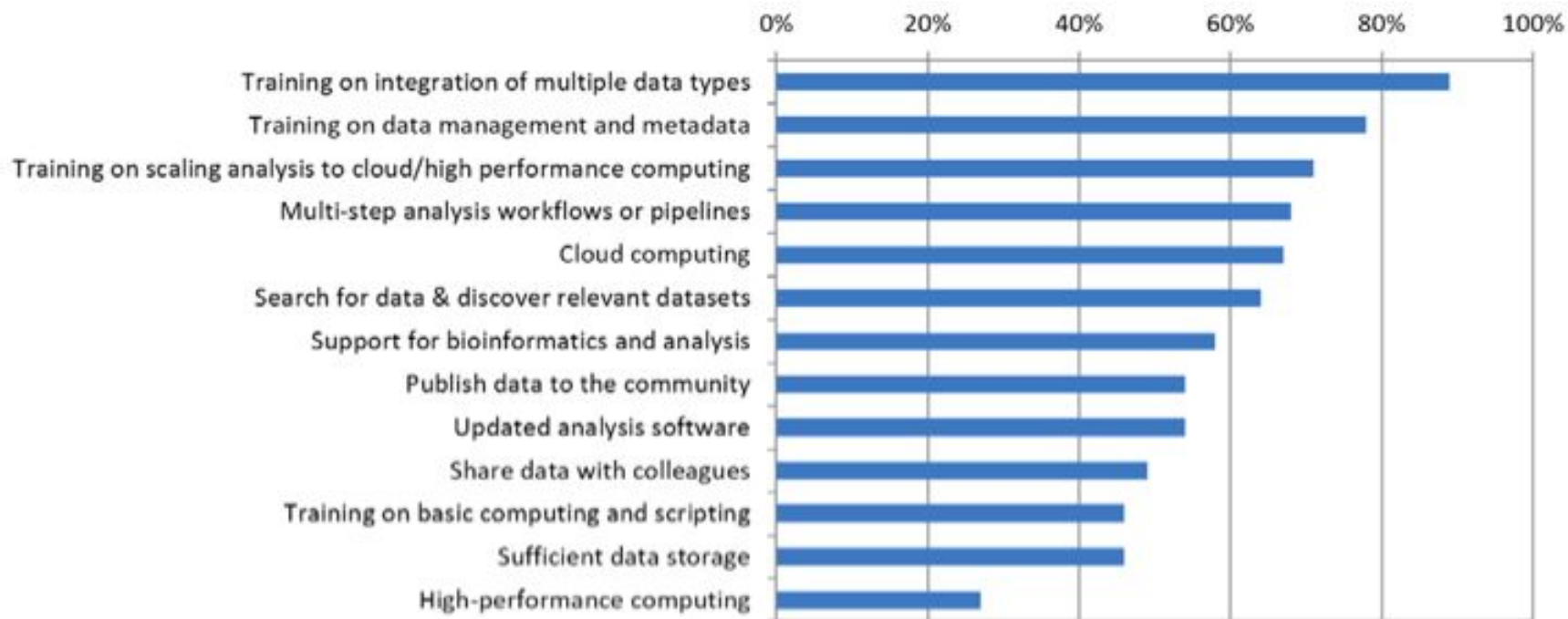
Lowndes, Julia S. Stewart, et al. "[Our path to better science in less time using open data science tools.](#)" *Nature ecology & evolution* 1.6 (2017): 160.

Researchers are very interested in learning these skills

Survey by Bioinformatics Resource Australia on what it would be most useful for them to offer



Current unmet needs



38 mentions of The Carpentries as an example and recommendation in the *Shifting to Data Savvy* report.



**Shifting to Data Savvy:
The Future of Data Science
In Libraries**

Matt Burton
Liz Lyon
Chris Erdmann
Bonnie Tijerina

The Strategic Value of Library Carpentry and The Carpentries to Research Libraries



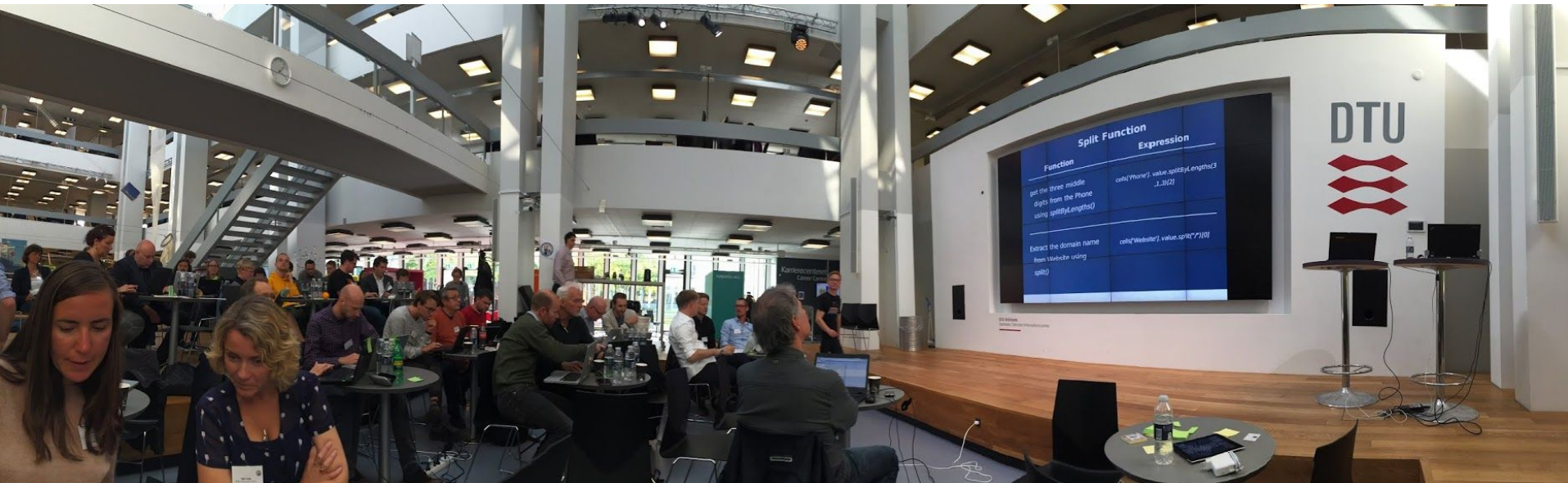
By Elaine L. Westbrooks



University of California
CDL
California Digital Library



Minding the gaps - Scaling software and data skills training



Carpentries: Building Skills and Community

- Non-profit teaching data science skills for more effective work and career development
- Creating training ‘in the gaps’ that is accessible, approachable, aligned and applicable (the practical skills you need in your work)
- Volunteer instructors, peer-led hands-on intensive workshops
- Open and collaborative lesson materials
- Creating and supporting community, local capacity for teaching and learning these skills and perspectives

Workshops



- 2-days, active learning
- Feedback to learners throughout the workshop
- Trained, certified instructors
- Friendly learning environment (Code of Conduct)

Focus

Data Carpentry

Domain-specific, research data-related

Software Carpentry

Domain agnostic, research workflow/software-related

Library Carpentry

Library and information/workflow-related, Carpentries onboarding, community outreach and advocacy-driven

Workshop Goals

- Teach skills
- Get people started and introduce them to what's possible
- Build confidence in using these skills
- Encourage people to continue learning
- Positive learning experience

Our workshops.
Our learners.

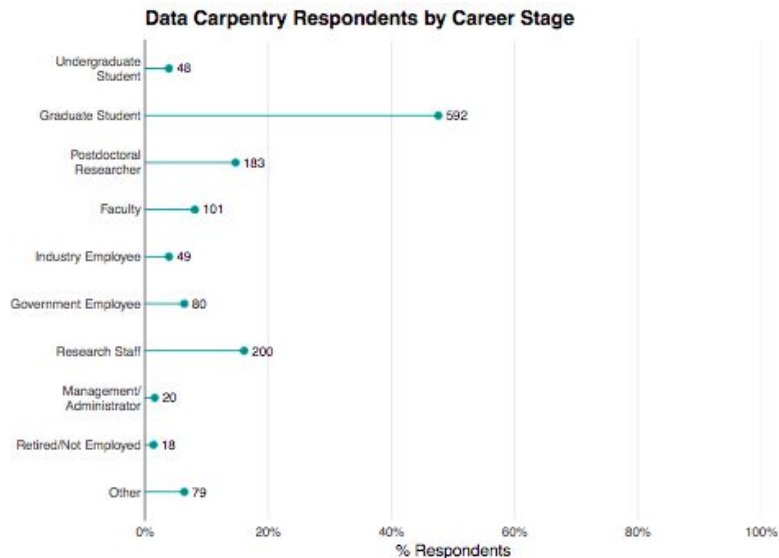


The Carpentries 2018 Annual Report

<https://carpentries.org/files/assessment/TheCarpentries2018AnnualReport.pdf>

66% of the Data Carpentry workshop attendees are early career.

Analysis of Software and Data Carpentry's Pre- and Post-Workshop Surveys
<https://doi.org/10.5281/zenodo.1325463>



Instructor Training

Educational pedagogy is the focus of Instructor training program:

1. 2-days on online training
2. Edit a lesson
3. 1-hr discussion
4. Demo

More information: <http://carpentries.github.io/instructor-training/>

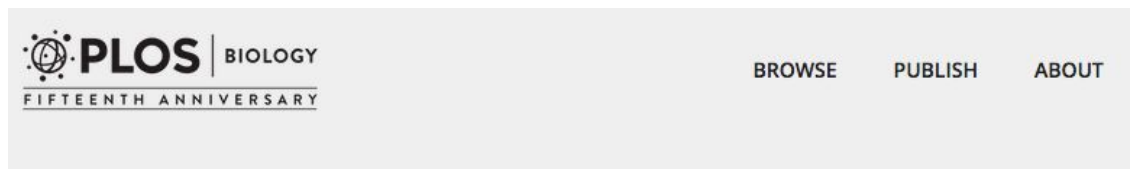
Community

A group of people excited about software and data skills and about sharing them with others

- Mentoring program and instructor onboarding
- Discussion groups and community calls
- Email lists
- Teaching at other institutions

Outcomes

Short and long term surveys show that people are learning the skills, putting them into practice in their work and have more confidence in their ability to do computational work.



 OPEN ACCESS

COMMUNITY PAGE

Developing a modern data workflow for regularly updated data

Glenda M. Yenni , Erica M. Christensen, Ellen K. Bledsoe, Sarah R. Supp, Renata M. Diaz, Ethan P. White, S. K. Morgan Ernest

Version 2



Published: January 29, 2019 • <https://doi.org/10.1371/journal.pbio.3000125>

Figure: Perception of Workshop Impact



Library Carpentry Core Objectives

Library Carpentry workshops teach people working in library- and information-related roles how to:

- Cut through the jargon terms and phrases of software development and data science and apply concepts from these fields in library tasks;
- Identify and use best practices in data structures;
- Learn how to programmatically transform and map data from one form to another;
- Work effectively with researchers, IT, and systems colleagues;
- Automate repetitive, error prone tasks.



Our Audience

Library Carpentry speaks to those working with information and data:

- Who want to explore this material to determine if it is helpful or relevant to their work
- Who face new challenges surrounding data such as learning new tools and practices
- Who want to improve their workflows and reduce time consuming tasks
- Who need to gain familiarity and comfort with language surrounding data
- Who would like to collaboratively develop lesson material around information and data
- Who are generally interested in data scholarship
- Who are looking for a community where they can discuss data related challenges



University of California
CDL
California Digital Library





23 Research Data Things allowed thousands of librarians to familiarize themselves with research data topics. Library Carpentry aims to do the same for software and data skills.

23 things [research data]
Australia

Do as many or as much as you want; do them in any order; do them by yourself, or form a group to learn together. There are many activities to choose from — whether you are new to data; need to extend your knowledge, or want a challenge!

Extend your knowledge and skills about research data by exploring activities, links and much more at ands.org.au/23-things

Ready, set, data!
Build your knowledge of key concepts and issues in research data management.

Repositories for data
Learn about repositories for depositing, managing and discovering research data.

Data citation & impact
Extract value from research data: data citation, impact and metrics.

Rights, ethics & sensitive data
Learn about responsible sharing and reuse, and the importance of forming research data.

Metadata & more
Research descriptor, controlled vocabularies, linked data and crosswalks.

Let's talk data!
Start a data conversation: data interviews and data management plans. Data stories from publishers and research funders.

Hands on with data & tools
Take introduction specific data and use software tools for research data. Our page aimed to be something new!

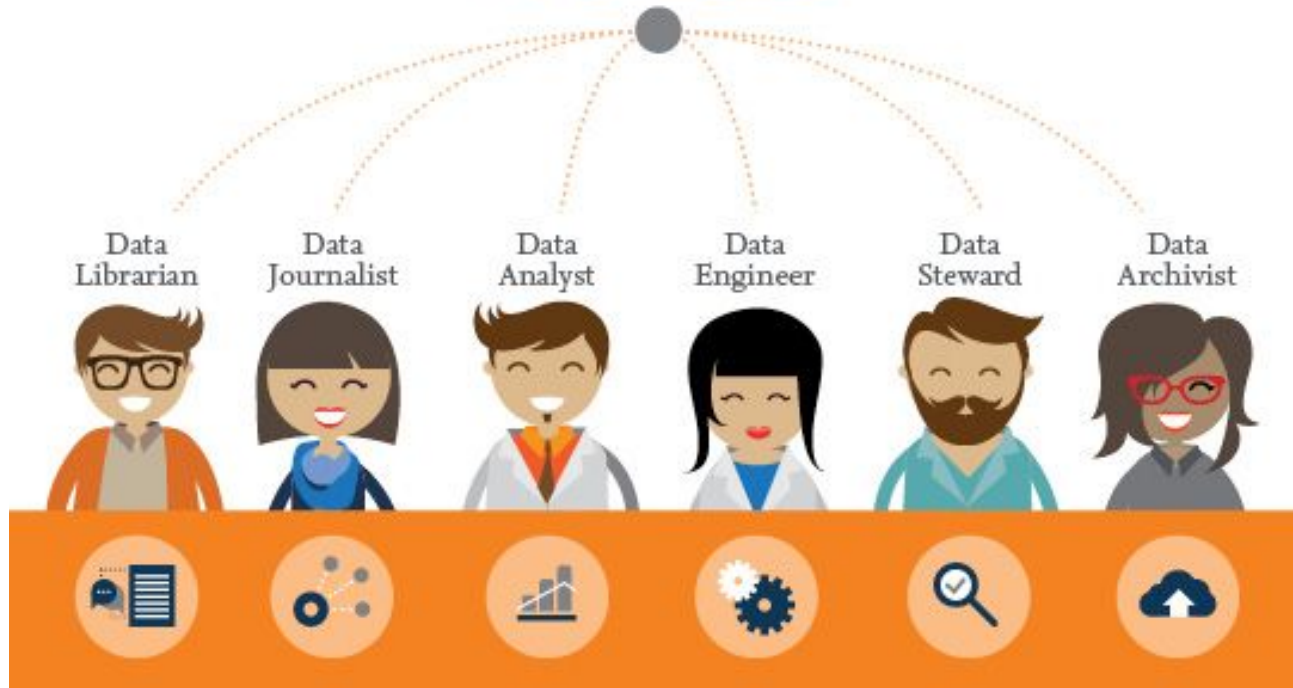
Data communities
Which who in the research data management zoo? Connecting with colleagues and continuing to learn.

1 Getting started with research data
2 Issues in research data management
3 Data in the research lifecycle
4 Data discovery
5 Data sharing
6 Long-lived data: curation & preservation
7 Data citation for access & attribution
8 Citation metrics for data
9 Licensing data for reuse
10 Sharing sensitive data
11 What's my metadata scheme?
12 Vocabularies for data description
13 Walk the crosswalk
14 Identifiers and linked data
15 Data management plans
16 What are publishers & funders saying about data?
17 Data literacy & outreach
18 Data interviews: ask the talk
19 Exploring APIs & apps
20 Find it with data!
21 Tools of the trade
22 What's in a name?
23 Making connections

ands.org.au/23-things

NCRIS National Research Infrastructure for Australia
Australian National Data Service 2016
ands

DATA SCIENCE ROLES



Together, as a community



The New England Software Carpentry Library Consortium (NESCLiC)





ZB MED are running Library Carpentry workshops and hacky hours.





Carpentries-based Workshop
“FAIR Data and Software”
July 9 - 13, 2018 in Hannover



Instructors

Katrin Leinweber, Angelina Kraft, Konrad Förstner, Martin Hammitzsch, Luke Johnston, Mateusz Kuzak

Helpers

Chris Erdmann

General Information

This workshop aimed to train junior scientists in implementing the FAIR principles for research data & software management & development. We want to help you identify similarities and differences between these two scientific objects and apply respectively appropriate good practices in preparing, publishing and archiving your work.

It was a new, experimental workshop format that contextualises the highly practical lesson material from the [Software](#) and [Data Carpentries](#) with the [FAIR principles](#)

University of California

CDL
California Digital Library





Australian Research Data Commons

Top 10 FAIR Data & Software Things

about github repository download/cite #top10fair

Oceanography

Research Software

Research Libraries

Research Data Management Support

International Relations

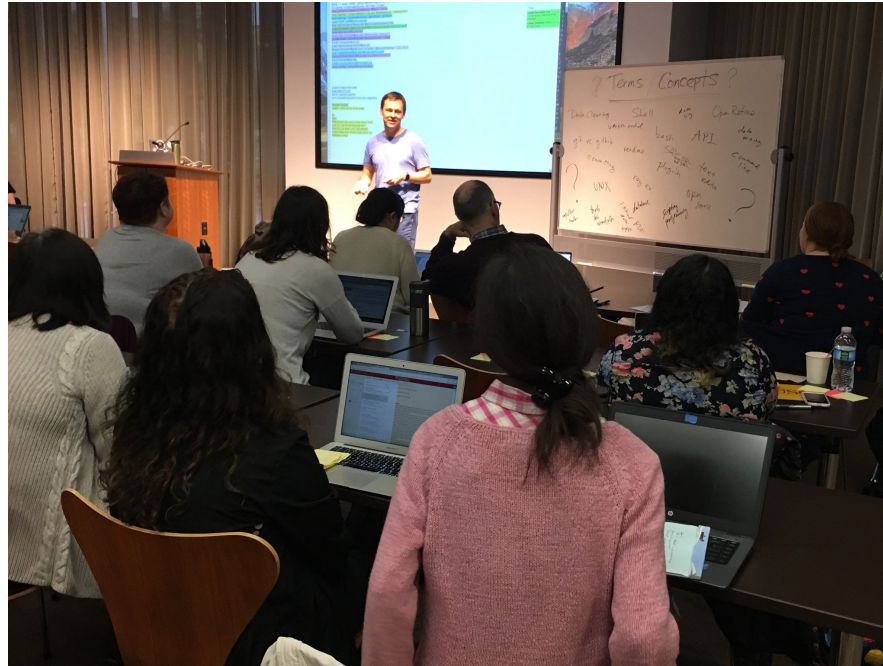
<https://librarycarpentry.org/Top-10-FAIR/>

How can I get started? Contribute to a lesson.

The screenshot shows the GitHub repository page for `LibraryCarpentry / lc-open-refine`. At the top, there are navigation options: `<> Code`, `Issues 7`, `Pull requests 0`, `Projects 0`, `Wiki`, `Insights`, and `Settings`. Below this, the repository name is displayed with a link to `https://librarycarpentry.github.io/lc...` and an `Edit` button. A horizontal bar shows repository statistics: `1,078 commits`, `14 branches`, `0 releases`, and `63 contributors`. Below the bar, there are buttons for `Branch: gh-pages`, `New pull request`, `Create new file`, `Upload files`, `Find file`, and `Clone or download`. The main content area shows a list of recent commits:

Commit	Description	Time
<code>ccronje Merge pull request #27 from LibraryCarpentry/libcce-patch-direct-down...</code>		Latest commit e123614 9 days ago
<code>.github</code>	Suggest template language	a year ago
<code>_episodes</code>	Merge pull request #26 from jt14den/dennis-fix-cluster-exercise-ep05	14 days ago
<code>_episodes_rmd</code>	move data/ into _episodes/ and _episodes_rmd/	2 years ago
<code>_extras</code>	Merge pull request #4 from LibraryCarpentry/ccronje-patch-4	4 months ago
<code>_includes</code>	resolve descriptin conflict	2 months ago
<code>_layouts</code>	use favicons for workshop page	4 months ago

How can I get started? Host, Help, Teach.



How can I get started? Become a member.

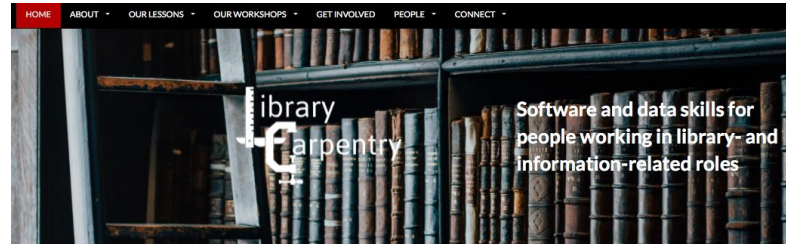
Our Current Member Organisations

A Member Organisation is an organisation that has made a financial commitment to the growth and sustainability of The Carpentries and is building local capacity for training. See more on [becoming a Member Organisation](#).

Platinum Member Organisations

- Compute Canada
- University of Arizona/CyVerse
- Macquarie University
- Manchester University
- New Zealand eScience Infrastructure (NeSI)
- Nordic e-Infrastructure Collaboration (NeIC)
- North West University South Africa
- Queensland Cyber Infrastructure Foundation
- Software Sustainability Institute
- Stanford University
- University of California Davis
- University of Oklahoma

Library Carpentry Website



What we do

Library Carpentry develops lessons and teaches workshops for and with people working in library- and information-related roles. Our goal is to create an on-ramp to empower this community to use software and data in their own work as well as be advocates for and train others in efficient, effective and reproducible data and software practices. Our workshops are based on our lessons. Workshop hosts, Instructors, and



Who we are

We are a diverse, global community of volunteers. Our community includes Instructors, helpers, Trainers, Maintainers, Mentors, community champions, member organisations, supporters, workshop organisers, and staff. Library Carpentry is guided by a Advisory Group and a Curriculum Advisory Committee. Our audience are primarily people working in library- and information-related roles.



Get involved

See all the ways you can engage and get involved with Library Carpentry. Follow us on Twitter.

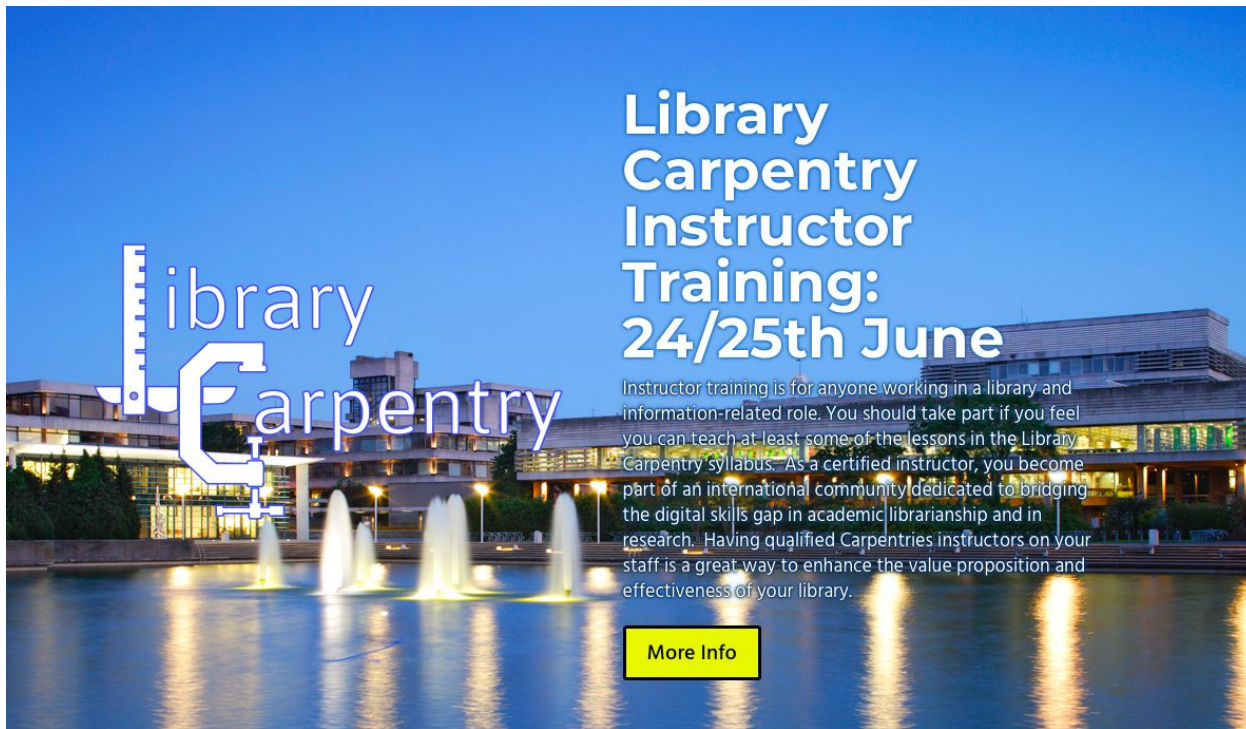
[More >](#)

<https://librarycarpentry.org>



University of California
CDL
California Digital Library





Library Carpentry Instructor Training: 24/25th June

Instructor training is for anyone working in a library and information-related role. You should take part if you feel you can teach at least some of the lessons in the Library Carpentry syllabus. As a certified instructor, you become part of an international community dedicated to bridging the digital skills gap in academic librarianship and in research. Having qualified Carpentries instructors on your staff is a great way to enhance the value proposition and effectiveness of your library.

[More Info](#)

<https://lirgroup.heanet.ie/index.php/2019/03/04/liber-lc-instructortraining2019/>



University of California
CDL
California Digital Library



Instructor Training Schedule

<https://carpentries.github.io/instructor-training/>

Sample Schedule

		Setup	Download files required for the lesson
		Pre-training survey	Please fill out our pre-training survey before the start of the course.
Day 1	09:00	Welcome	Who are we and how do we approach teaching? What should you expect from this workshop?
	09:15	Building Skill With Practice	How do people learn? Who is a typical Carpentries learner? How can we help novices become competent practitioners?
	10:20	Expertise and Instruction	What type of instructor is best for novices? How are we (as instructors) different from our learners and how does this impact our teaching?
	11:10	Morning Break	Break
	11:25	Memory and Cognitive Load	What is cognitive load and how does it affect learning? How can we design instruction to work with, rather than against, memory constraints?
	12:10	Building Skill With Feedback	How can I get feedback from learners? How can I use this feedback to improve my teaching?
	12:30	Lunch	Break
	13:30	Motivation and Demotivation	Why is motivation important? How can we create a motivating environment for learners?
	14:35	Mindset	How does mindset influence learning? How should we praise our learners? How should we talk about errors? What are successful habits of lifelong learners?



Mozilla-Library Carpentry Global Sprint, 30-31 May, 2019

Upcoming Sprint to improve and develop Library Carpentry material

<https://librarycarpentry.org/blog/2019/03/lc-mozilla-global-sprint/>

Thank you

Contact Chris Erdmann if you have further questions at Christopher.Erdmann@ucop.edu or @libcce on Twitter.



THANKS!

Questions?

Please put them in the chatbox.