



# Introduction to Library Carpentry and Building a Carpentries Consortium

NNLM Research Data Management Webinar Series  
February 7, 2019

Chris Erdmann @libcce  
Kate Nyhan @KDNyhan  
Julie Goldman @jgolds2

# Outline

**Introduction to Library Carpentry — Chris Erdmann**

**Demonstration of OpenRefine — Kate Nyhan**

**Background and Information About NESCLiC — Julie Goldman**

**Questions & Discussion**

# Introduction to Library Carpentry

Chris Erdmann

Community and Development Director  
Library Carpentry

The IMLS Data Science in Libraries Shifting to Data Savvy Report highlights the increasing demand for software and data skills across the spectrum of library roles and the importance of Library Carpentry and The Carpentries to meet this need.

Burton, M., Lyon, L. Erdmann, C., Tijerina, B. (2018) Shifting to Data Savvy: The Future of Data Science In Libraries. Project Report. University of Pittsburgh, Pittsburgh, PA. <http://d-scholarship.pitt.edu/33891>



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

# Why Library Carpentry?

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.



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

How your library can be a hub for data science activities in your community.



By Elaine L. Westbrook



# Workshops



- 2-days, active/hands-on learning
- Trained/certified instructors
- Feedback to learners throughout the workshop (Surveys, post-its, +/-)
- Friendly learning environment (Code of Conduct)
- Flexibility is important to Library Carpentry

# 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



# Core Curriculum/Lessons

**Data Intro for Librarians:** An introduction to data structures, regular expressions, and computing terms (*Jargon Busting & Pattern Matching*)

**Unix Shell:** An introduction to command line interfaces and task automation using the Unix shell (*Text-based*)

**Git Intro for Librarians:** An introduction to version control using Git and GitHub for collaboration (*GitHub Focus*)

**OpenRefine:** An introduction to cleaning up and enhancing a dataset using OpenRefine (*Journal Metadata Cleaning*)

**...SQL, Tidy Data, Python, Webscraping, Wikidata, FAIR, R, Digital Preservation**

# NLM Library Carpentry Pilot Report

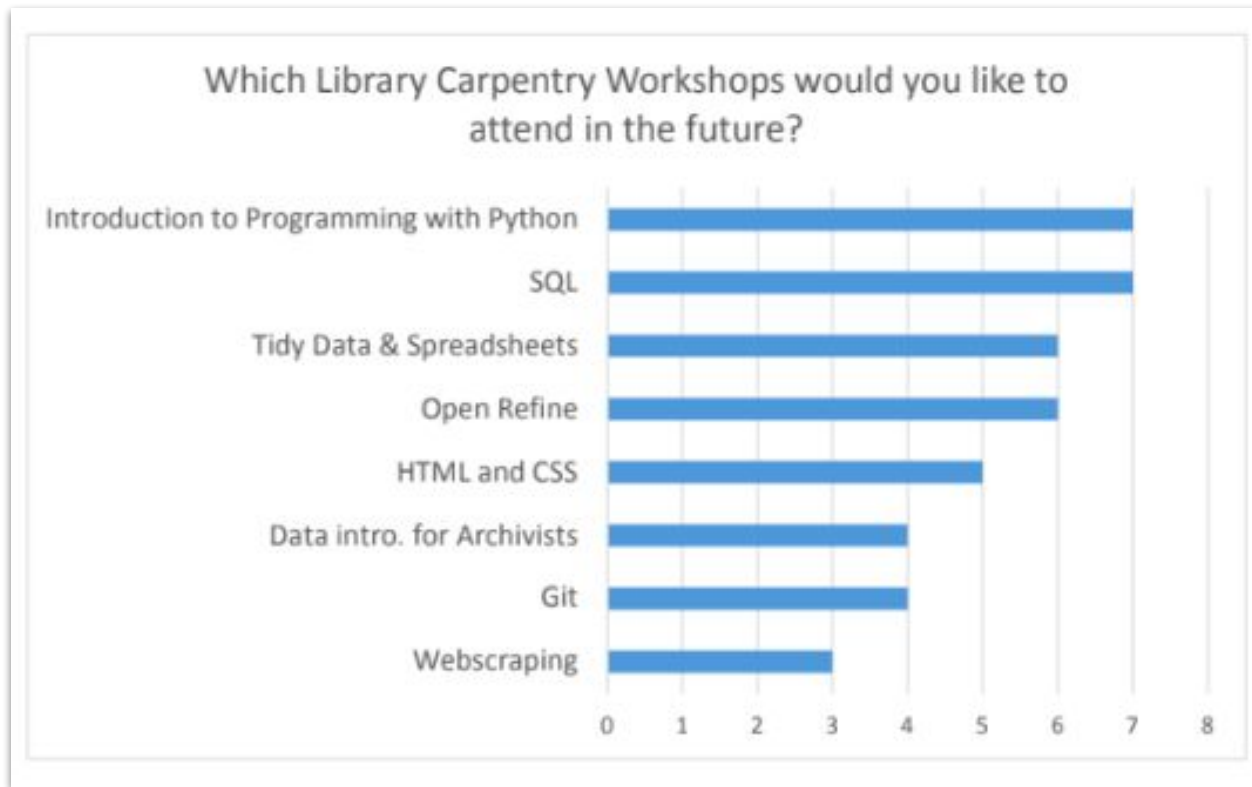
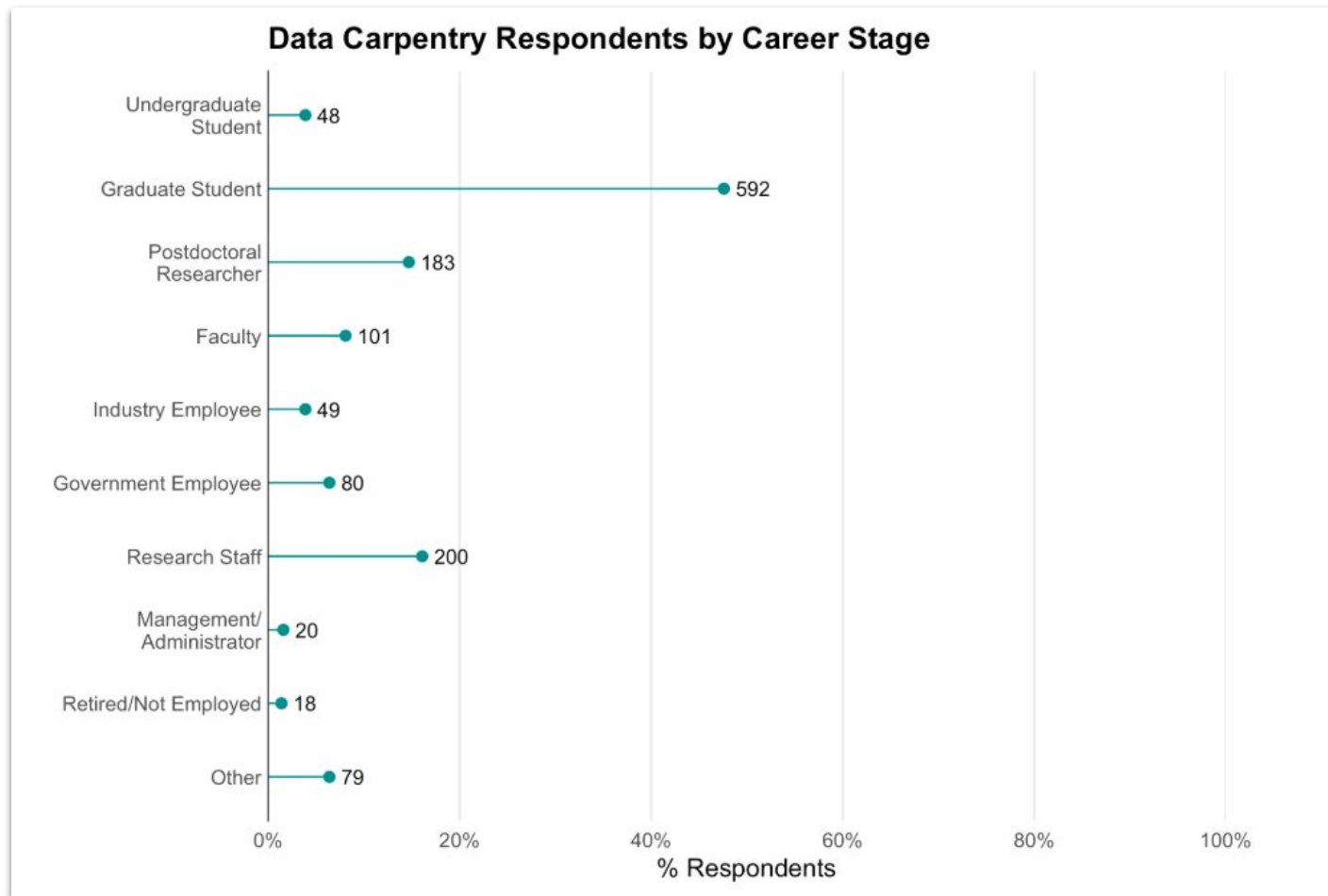


Figure: Perception of Workshop Impact



Analysis of the Carpentries Long-Term Feedback Survey: Perception of Workshop Impact

[https://carpentries.github.io/assessment/learner-assessment/archives/2018/code/2018\\_January\\_long\\_term\\_report.html](https://carpentries.github.io/assessment/learner-assessment/archives/2018/code/2018_January_long_term_report.html)



Analysis of Software and Data Carpentry's Pre- and Post-Workshop Surveys: Learners' career stage  
<https://carpentries.github.io/assessment/learner-assessment/reports/2018-07-learners-pre-post-surveys.html>

# Pathways

## **Instructor Certification**

Apply / Member ⇒ 2 day-online training (pedagogy) ⇒ Checkouts/Demo ⇒ Host/Help/Teach

[Apply to be an Instructor](#)

[Become a Member](#)

## **Request a Workshop**

2-day workshop ⇒ Library, Data, and/or Software? ⇒ Recruit instructors ⇒ Host Workshop

\* Cost is \$2500

[Request a Workshop](#)

[Carpentries Handbook](#)

**Get involved:** [https://librarycarpentry.org/get\\_involved](https://librarycarpentry.org/get_involved)

# Carpentries



Non-profit organization that:

- Trains people in software development and data science skills for more effective work and career development
- Builds community and local capacity for teaching and learning these skills and perspectives

Note: “Carpentry” means “the basics” like learning how to nail two boards together or put up a wall straight.





The Carpentries 2018 Annual Report

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

# Demonstration of OpenRefine

Kate Nyhan

Research and Education Librarian

Yale University

# Carpentry Teaching Skills

- Instructors learn from peers
- See teaching practices in action
- Opportunities to practice teaching
- Provide and receive useful & thoughtful feedback
- Incorporate feedback into teaching practices



# Library Carpentry Demo: OpenRefine

## **Today:**

Importing data -- OpenRefine layout -- data cleaning -- data export -- reconciliation -- feedback


See the whole Library Carpentry lesson:

<https://librarycarpentry.org/lc-open-refine>

Here's the Library Carpentry summary of what you can do with OpenRefine:

<https://librarycarpentry.org/lc-open-refine/reference.html>

# Facet/Filter and Undo/Redo

 **OpenRefine** doaj article sample csv [Permalink](#)

Facet / Filter Undo / Redo 2 / 2

Refresh Reset All Remove All Show as: rows records Show: 5 10 25 50 rows

**Language** change

4 choices Sort by: name count Cluster

EN 871

English 107

ES 7

FR 1

(blank) 15

Facet by choice counts

**Publisher** change

7 choices Sort by: name count Cluster

Akshantala Enterprises 13

Aurel Vlaicu University Editing

House 17

Consejo Superior de Investigaciones

Científicas 11

International Union of


Crystallography 858

MDPI AG 23

1001 rows

All Title Authors

☆	🗨	1.	The Fisher Thermodynamics of Quasi-Probabilities	Flavia Pennini Angelo Plastino
☆	🗨	2.	Aflatoxin Contamination of the Milk Supply: A Pakistan Perspective	Naveed Aslam Peter C. Wynn
☆	🗨	3.	Metagenomic Analysis of Upwelling-Affected Brazilian Coastal Seawater Reveals Sequence Domains of Type I PKS and Modular NRPS	Rafael R. C. Cuadral Juliano C. Cury Alberto M. R. D'Ávila
☆	🗨	4.	Synthesis and Reactivity of a Cerium(III) Scorpionate Complex Containing a Redox Non-Innocent 2,2â€²-Bipyridine Ligand	Fabrizio Ortul Hao Zhu Marie-Emmanuelle Boulon David P. Mills
☆	🗨	5.	Performance and Uncertainty Evaluation of Snow Models on Snowmelt Flow Simulations over a Nordic Catchment (Mistassibi, Canada)	Magali Troin Richard Arsenault François Brissette
☆	🗨	6.	Dihydrochalcone Compounds Isolated from Crabapple Leaves Showed Anticancer Effects on Human Cancer Cell Lines	Xiaoxiao Qin Yun Feng Xing Zhiqin Zhou Yuncong Yao

 **OpenRefine** doaj article sample csv [Permalink](#)

Facet / Filter

Undo / Redo 2 / 2

Extract... Apply...

Filter:

0. Create project

1. Split multi-valued cells in column Authors

2. Join multi-valued cells in column Authors

1001 rows

Show as: rows records Show: 5

All Title

☆	🗨	1.	The Fisher Thermodynamics of Quasi-Probabilities
☆	🗨	2.	Aflatoxin Contamination of the Milk Supply: A Pakistan Perspective
☆	🗨	3.	Metagenomic Analysis of Upwelling-Affected Brazilian Coastal Seawater Reveals Sequence Domains of Type I PKS and Modular NRPS
☆	🗨	4.	Synthesis and Reactivity of a Cerium(III) Scorpionate Complex Containing a Redox Non-Innocent 2,2â€²-Bipyridine Ligand
☆	🗨	5.	Performance and Uncertainty Evaluation of Snow Models on Snowmelt Flow Simulations over a Nordic Catchment (Mistassibi, Canada)
☆	🗨	6.	Dihydrochalcone Compounds Isolated from Crabapple Leaves Showed Anticancer Effects on Human Cancer Cell Lines

# Background and Information About NESCLiC

Julie Goldman

Research Data Services Librarian  
Harvard Medical School



# Community of Practice

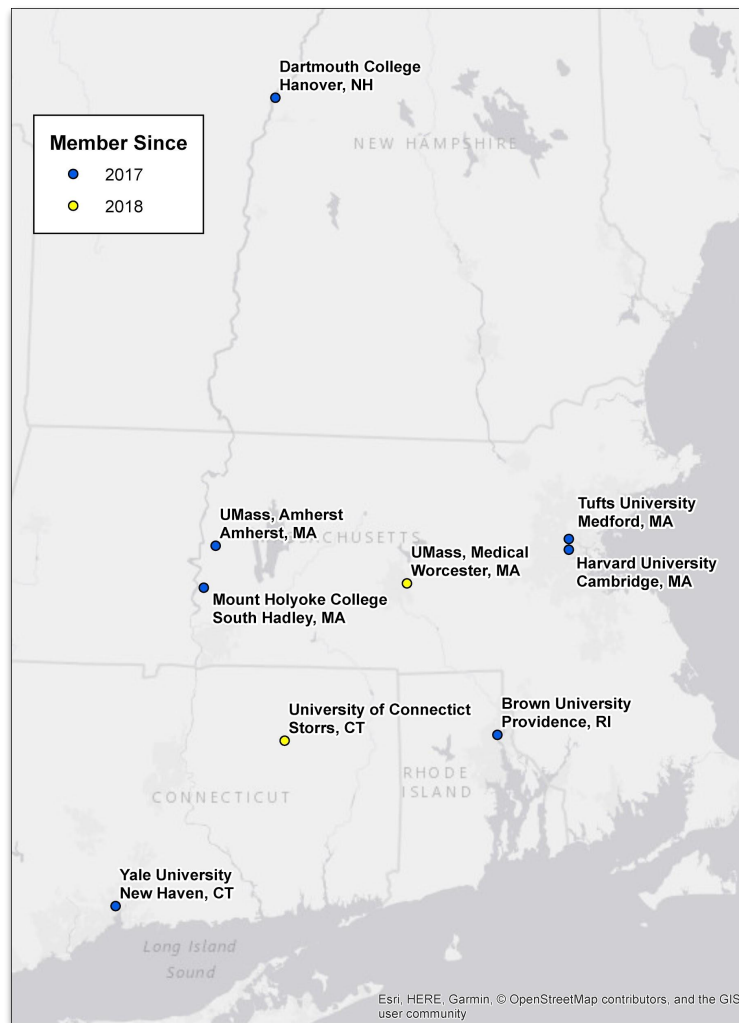
Could we bring together institutional contributions and the staff time of already-busy professionals to increase the number of available Carpentries instructors in the region?

Who at our institutions would want this training?

What would be the benefits to our institutions?

# The New England Software Carpentry Library Consortium

Institution	Number of Participants Total (2017/2018)
Brown University	4 (2/2)
Dartmouth College	4 (2/2)
Harvard University	3 (2/1)
Mount Holyoke College	2 (2/0)
Tufts University	5 (3/2)
University of Connecticut	1 (0/1)
University of Massachusetts, Amherst	3 (1/2)
University of Massachusetts, Worcester	2 (0/2)
Yale University	6 (3/3)



# Membership and Benefits

- Gold Tier Membership - \$15,000 per year
  - 15 people get instructor training
  - 6 workshops with administrative support by the Carpentries
- Pilot
  - Try out Carpentries membership with less risk
- Community building
  - Across organizational and institutional boundaries
  - Many backgrounds and job titles
  - Teach together - help fill skill gaps
  - Group to ask for help!

# Carpentry Instructor Training

- Introduce you to evidence-based best-practices of teaching
- Teach you how to create a positive environment for learners at your workshops
- Provide opportunities for you to practice and build your teaching skills
- Help you become integrated into the Carpentries community
- Prepare you to use these teaching skills in teaching Carpentries workshops



# Activities

- Develop instructional materials and contribute back to the community
- Be part of a network of data fluent researchers and instructors
- Refine skills by participating as instructors at other member institutions
- Learn new skills and improve on existing ones

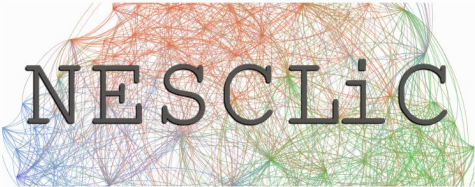


October 2018 Library Carpentry Workshop at Brown University Library

# nesclib.github.io/home

NESCLIC

AboutMembersPublicationsWorkshops



The **New England Software Carpentry Library Consortium** or **NESCLIC** is an association of academic libraries, joined as a community of practice focused on building data science skills in research computing, and extracting, wrangling, storing, analyzing, and visualizing data. NESCLIC allows member institutions to share in a Gold-level membership with the Carpentries.

### Our Model

NESCLIC member institutions share a Gold-level Carpentries membership, which affords NESCLIC 6 free coordinated workshops, a 50% discount on the cost for additional coordinated workshops, no charge for self-organized workshops, and fifteen instructors trained, per year.

NESCLIC provides the opportunity for members to:

- Develop instructional materials
- Build a network of data fluent researchers and instructors
- Refine skills by participating as instructors at other member institutions

**Financial commitment:** Cost of shared Carpentry membership (\$1000/person) plus training expenses (travel or hosting)

**Time commitment:** (2) – two-day training workshops plus committee work

NESCLIC has established bylaws and members contribute to the consortium by participating in one of several roles:

- Carpentries liaison
- Membership/Recruiting
- Assessment
- Outreach
- Workshop coordination

### Become a Member

Interested in joining NESCLIC?

Email: [nesclib@googlegroups.com](mailto:nesclib@googlegroups.com)

## Workshops

Check back for future workshops offered by NESCLIC members.

*This list includes official Carpentries' workshops and workshops adapted from the Carpentries' materials where NESCLIC members served as hosts, instructors, or helpers.*

Venue	Dates	Lessons
<a href="#">Tufts University</a>	April 12-13, 2019	Unix, Git, Programming with Python
<a href="#">Tufts University</a>	February 15-16, 2019	Data Organization in Spreadsheets, OpenRefine for Data Cleaning, R for Social Scientists
<a href="#">Brown University</a>	January 16-17, 2019	Data Organization in Spreadsheets, OpenRefine for Data Cleaning, Introduction to R, Data Analysis and Visualization in R, Data Management with SQL
<a href="#">Harvard University</a>	January 15-16, 2019	Data Intro for Librarians, Shell Lessons for Librarians, Tidy Data for Librarians, OpenRefine for Librarians
<a href="#">UMass Amherst</a>	January 8-9, 2019	Unix, Python, Git
<a href="#">Dartmouth College</a>	November 29-30, 2018	Unix, Git, Python
<a href="#">Tufts University</a>	November 9-10, 2018	Unix, Git, Programming with Python
<a href="#">Brown University Library (Hosts &amp; Helpers), Yale and Tufts (Instructors)</a>	October 22-23, 2018	Data Intro for Librarians, Shell Lesson for Librarians, Git Intro for Librarians, OpenRefine for Librarians
<a href="#">Yale University Library</a>	August 15-16, 2018	Tidy Data for Librarians, Shell Lesson for Librarians, OpenRefine for Librarians, Python Intro for Librarians



# Workshop Feedback

We collect feedback at the end of each day:

## YELLOW (What worked well?)

- “I liked the availability of further examples”
- “Excellent introduction”
- “Very helpful to learn about collaborating on Github & using OpenRefine. I can see application to my work in the future”

## RED (What could be better?)

- “Wish we had more time”
- “More on how different parts of the workshop could work together”
- “More practice time”

# Goals & Updates

- Keep alumni involved
- Recruit new members
  - Launched 2nd cohort
- Run more LC Workshops
- Develop new curriculum
  - Digital Humanities at Yale
  - Biodiversity at Harvard



Carpentry Instructors Training at UMass Amherst Library

# Get Involved!

Learn how the instructors are staying involved in this consortium, building their own carpentry skills, giving back to their communities, and collaborating around New England (and beyond)!

**Email:** [nesclic@googlegroups.com](mailto:nesclic@googlegroups.com)

**Website:** <https://nesclic.github.io/home>






**Github:** <https://github.com/NESCLiC>

**Publications:** <https://nesclic.github.io/home/publications>

# Upcoming Workshops

## Upcoming Workshops

This page includes upcoming workshops and previous workshops, sorted by upcoming workshops. Most are first advertised to the organization running the workshop, but some are open to other participants. Don't see a workshop in your area, but wish there was one? Become a host and [request a workshop!](#)

	<b>University of Toronto</b> Instructors: Rachel Wang Helpers: Jordan Pedersen, Sarah Gorman, Angel Su, Dominique Glassman	11 March 2019
	<b>NEPTTP WITS Library Carpentry Workshop</b> Instructors: Juan Steyn Helpers: Retha Langa, Pelumi Oladokun	30 January - 1 February 2019
	<b>Harvard University</b> Instructors: Kristin Lee Helpers: Bill Barthelmy, Ceilyn Boyd, Jess Cohen-Tanugi, Cole Crawford, Julie Goldman, Jeremy Guillette, Hugh Truslow	15-16 January 2019
	<b>ZB MED</b> Instructors: Silvia Di Giorgio, Konrad Förstner Helpers:	15-16 January 2019
	<b>Middle Tennessee State University</b> Instructors: Chris Erdmann, Juliane Schneider, Meredith Broadway, Anna Oates Helpers: Mary Ellen Sloane, Suzanne Mangrum, Robert T. Wilson	7-9 January 2019

[https://librarycarpentry.org/upcoming\\_workshops](https://librarycarpentry.org/upcoming_workshops)

# Connect with Library Carpentry

- Main email communication channel is our [Library Carpentry Topicbox Group](#)
- Chat in our [Gitter Chatroom](#) -- join with your GitHub username or Twitter handle
- New to the community and have questions? Reach out to the Library Carpentry Community and Development Director, [Chris Erdmann](#)
- Want a workshop at your institution? [Request one](#)! Note your interest in Library Carpentry!
- You can raise issues and/or pull requests on any of the lesson repositories and this site through the [Library Carpentry GitHub organisation](#)
- Follow the [@LibCarpentry Twitter](#) to learn about what community members are doing!
- Find out about the many ways you can [connect with The Carpentries](#) community via social media, Slack, email lists, blog, community calendar, and newsletter.
- For general information, inquiries or comments, get in touch via [team@carpentries.org](mailto:team@carpentries.org)

# Questions & Discussion



**Chris Erdmann**  
**@libcce**

Community and Development Director  
Library Carpentry



**Kate Nyhan**  
**@KDNyhan**

Research and Education Librarian  
Yale University



**Julie Goldman**  
**@jgolds2**

Research Data Services Librarian  
Harvard Medical School