

# A DIFFERENT KIND OF CARPENTRY

**Software Carpentry, Data Carpentry**, and giving researchers the skills they need to tackle large data and computing challenges

 Feel free to tweet or post slide photos from this talk

Christina Koch, UW Madison

August 4, 2017

ACI-REF Virtual Residency

# What are Software and Data Carpentry?

  
software carpentry



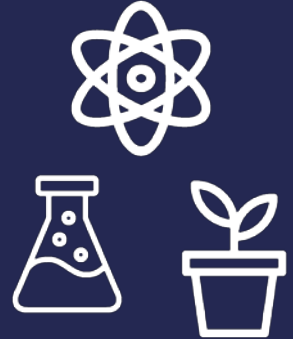
Two formal **organizations\***  
+ international **community** of  
volunteers



that develop **lessons**, train **instructors** +  
organize **workshops**



in **data**  
management and **software**  
development **best practices**



For and by **researchers**  
across  
disciplines

\* soon to be restructured into a combined organization

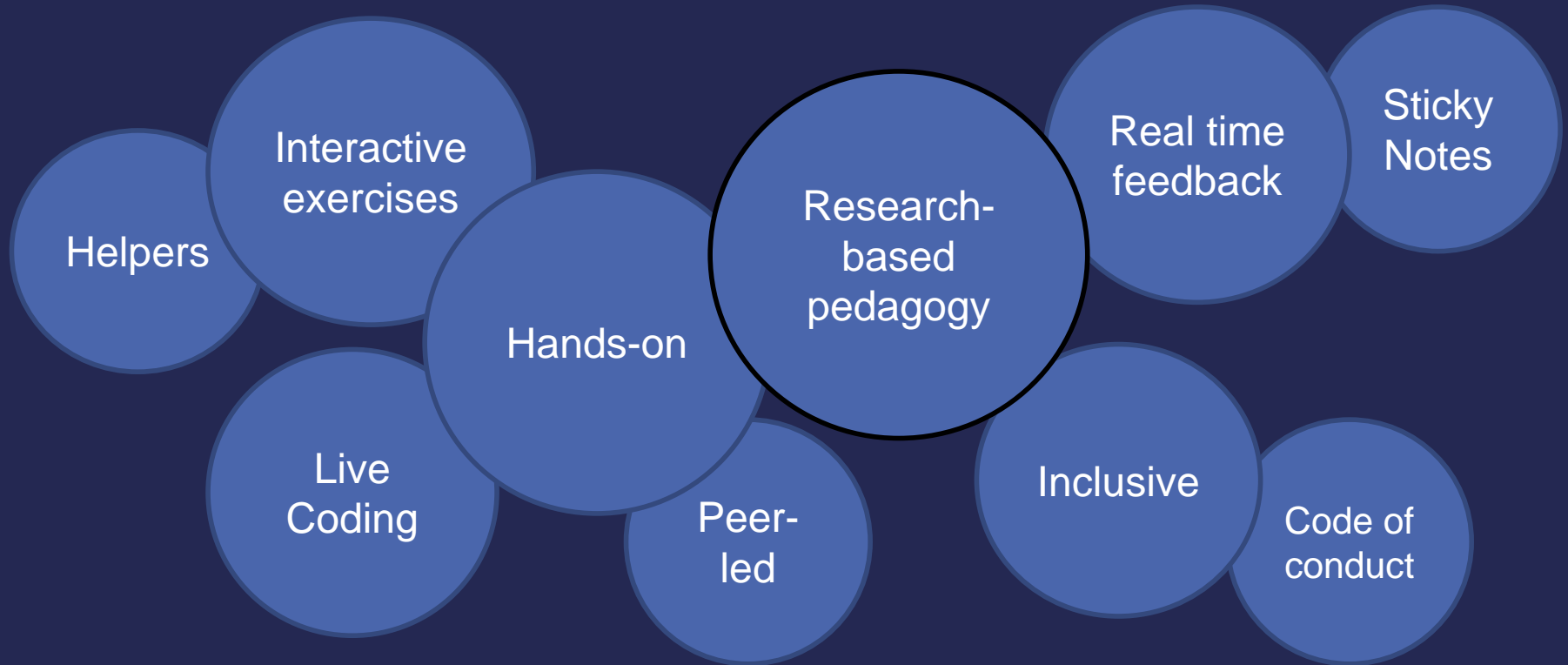
# Filling the Gaps

- Identified needs
  - Researchers writing programs without any training in useful “industry-standard” best practices
  - Researchers struggling with quantity of data and how to organize, analyze, and otherwise manage it
- Solution
  - 2-day, peer-led workshops
  - collaboratively developed lesson modules
  - leveraging the volunteer efforts of community members
  - emphasis on best practices

# Carpentries Philosophy: Content

- Software and Data Carpentry teach **skills**:
  - version control, basic programming, command line
  - data organization, cleaning, analysis and visualization
- Founded on **best practices**
  - building modular code, using data structures, reproducibility
- For a **research** audience
  - Emphasis is not advanced, enterprise workflows or tools, but basic “toolbox” skills for everyday use

# Carpentries Philosophy: Method



# Carpentries: “Broader Impacts”

Build local  
community  
around computing  
and data skills

Promote open,  
reproducible,  
research

Provide  
researchers with  
transferable skills  
in programming,  
data analysis, and  
teaching

Discover job  
opportunities

Meet others in  
similar roles

Adapt lesson  
materials for  
courses or  
trainings; jump  
start course  
creation

Build local  
community of  
well-trained  
instructors

Make oneself  
more attractive  
job (faculty)  
candidate

# The Current Carpentries

- International community with many ways to engage
  - Hosting workshops, teaching, helping, developing lessons, mentoring others, sharing expertise
- Two major lesson stacks + more coming
- Instructor training on how to teach more effectively
- Can get help organizing or self-organize workshops
- In the past 5 years, more than 1,100 workshops of 28,000+ learners in over 40 countries

# HOW COULD “THE CARPENTRIES” HELP YOU AT YOUR INSTITUTION?

---



# Next Steps: Learn More

Look at  
lessons



<https://software-carpentry.org/lessons>

<http://www.datacarpentry.org/lessons/>

Follow



<https://twitter.com/swcarpentry>  
<https://software-carpentry.org/blog/>

<https://twitter.com/datacarpentry>  
<http://www.datacarpentry.org/blog/>

Join  
email lists



<http://eepurl.com/cfODMH>

<http://lists.software-carpentry.org/listinfo/discuss>

# Next Steps: Get Involved

Host a  
workshop



<https://software-carpentry.org/workshops/request/>

<http://www.datacarpentry.org/workshops-host/>

Become an  
instructor



<http://swcarpentry.github.io/instructor-training/>  
[https://amy.software-carpentry.org/forms/request\\_training/](https://amy.software-carpentry.org/forms/request_training/)

Become a  
partner



<https://software-carpentry.org/membership/>

<http://www.datacarpentry.org/partnerships/>

# Next Steps: HPCarpentry?

- What about large-scale computing? Can “Carpentry” principles be applied to create lessons in using high throughput / high performance / large memory systems?
- Unique challenges for this lesson: how to handle a variety of systems, schedulers, and use cases
- Lots of interest and initial outlines here:
  - <https://github.com/swcarpentry/hpc-novice/>
- If interested, join mailing list:
  - <https://groups.google.com/a/carpentries.org/forum/#!forum/hpc-discuss>

# QUESTIONS?

---

For further inquiries, email `ckoch5 (at) wisc (dot) edu`

Slides available at <https://speakerdeck.com/christinalk/>

# Icon credits

<https://thenounproject.com/>

- Book by Pumpkin Juice from the Noun Project
- Mail by Kidiladon from the Noun Project
- Newspaper by unlimicon from the Noun Project
- audience by Gan Khoon Lay from the Noun Project
- Institution by amy morgan from the Noun Project
- Erlenmeyer Flask by Rockicon from the Noun Project
- house plant by Arthur Shlain from the Noun Project
- atom by Awesome from the Noun Project
- Table by [www.mindgraphy.com](http://www.mindgraphy.com) from the Noun Project
- code by Ananth from the Noun Project
- people by Anastasia Latysheva from the Noun Project