Spicy Data Skills Open Science Program With Library Carpentry

***Carlene Barton**, eLearning Technologist, The University of Manchester Library, United Kingdom, carlene.barton@manchester.ac.uk, ORCiD 0009-0006-5698-2447; **Jodie Double**, Head of Digital Production and Access, University of Leeds Libraries, United Kingdom, J.L.Double@leeds.ac.uk, ORCiD 0000-0002-7152-3947; **Nilani Ganeshwaran**, Senior Software Developer (Management), The University of Manchester Library, United Kingdom, nilani.ganeshwaran@manchester.ac.uk; **Ann James**, PhD, Independent Scholar, amyatt.james2@gmail.com, ORCiD 0000-0002-2137-7961; **Phil Reed**, Research Community and Training Manager, The University of Manchester, United Kingdom, phil.reed@manchester.ac.uk, ORCiD 0000-0002-4479-715X; and **Jennifer Stubbs**, Librarian and Assistant Professor, Bradley University, Cullom-Davis Library, jastubbs@fsmail.bradley.edu, ORCiD: 0000-0002-6080-5703*

NUTRITION INFORMATION

This recipe describes how to adapt openly licensed materials from the Library Carpentry (https://librarycarpentry.org/) curriculum to develop a data skills program that works for your individual institution. Librarians may be most comfortable exploring Library Carpentry first before expanding into Data Carpentry (https://datacarpentry.org/) or Software Carpentry (https://software-carpentry.org/), depending on their campus needs. The Library Carpentry curriculum differs from other Carpentries because it caters to the knowledge and skills of librarians. A good example is how to apply Tidy Data principles to historical dates or the lesson about 'Introduction to Data for Archivists'. All instructors receive the same 16-hour train-the-trainer certification, whether Software, Data or Library Carpentry. Each workshop's content is tailored with data files recognizable to that audience (Ecology and farms, MARC or Grateful Dead discographies, Data Management in SQL for Ecologists, R for Reproducible Scientific Analysis). Then learners can see data they are already familiar with to transfer the skills to current problems at their jobs. This recipe focuses on Library Carpentry but can also be applied to Software Carpentry (for research programming skills) and Data Carpentry (for research data structure best practice).

Academic libraries that are members of the Carpentries serve locally customized menus of instruction and professional development on computing and data literacy. The Carpentries support open science initiatives through librarians becoming instructors in a train-the-trainer program. These are often facilitated in collaboration with university-based science disciplines and early career researchers, especially graduate students. By using open-source software, such as OpenRefine, and removing barriers created by proprietary software, the Carpentries are strongly aligned with open science principles. Carpentries' materials are available for use under license (https://carpentries.org/license/), which means that individuals are able to use, share, and adapt material improvements to their own practice and share this learning with colleagues, with attribution (http://creativecommons.org/licenses/by/4.0/).

Carpentries lessons are scalable, customizable, and available to support outreach and collaboration across a range of organizations, including resource-limited institutions. A broader Carpentries community of cooks also provides continued professional development support for instructors and facilitates the creation of new open educational resources (OER) and adaptations on the basis of inclusion, accessibility, and local languages. By hosting the lessons in GitHub under a CC-BY-4.0 license, the materials meet OER definitions. Several are already translated and maintained in Spanish, with other translations in the works. Students and instructors can provide feedback on every lesson page to report errors or suggest improvements.

Cooks may already be familiar with the previous ACRL cookbook recipe "Software Carpentry Al Dente" (Ossom-Williamson et al., 2022), which explains how to cook a single Carpentries workshop. Here, we put greater focus on development at the program level and describe how Library Carpentry can be selectively adapted for developing a spicy

open science data skills program at your academic library.

LEARNING OUTCOMES

After following this recipe cooks will be able to:

- engage Library Carpentry materials to advance an open science program and support professional development for researchers, librarians, and other university staff;
- identify and explain a Library Carpentry workshop and its relevance, describe its available OERs, and find and access its community, including a network of trained volunteer instructors;
- advise colleagues and peers how to leverage existing networks and infrastructure to customize aspects of Library Carpentry (batch cooking vs. cooking for one); and
- adapt existing Library Carpentry lessons and principles to the local context and institution.

NUMBER SERVED

Each meal can yield infinite servings. Chefs can scale up as audience appetite grows and feed entire labs, departments, even colleges (Gofman, 2019). A chef can also produce these meals at a small scale, for a library staff leveraging many classes with limited personnel, for example.

COOKING TIME

Improving your library staff's capacity to support FAIR (findability, accessibility, inclusion, and reproducibility) data principles (OpenAIRE, 2023) and open science is not a quick fix. Major funding bodies increasingly require the research they fund to be findable (with a persistent identifier), accessible (online), interoperable (applying common file format and standards), and reusable (well documented and licensed). It may take months or years to connect and collaborate with your colleagues and peers, apply for and win funding, to advertise and recruit. Like a barbeque, you will need a long, slow heat to get a maximum return.

In the short term, create interest at staff orientations or introductions by combining Carpentries' OER with networking appetizer lessons like Tidy Data (Wickham, 2014). In the medium term, get buy-in by spending time with new chefs to develop recipes and data sets in a local context.

Table 1. *Detailed Cooking Times for Supporting an Open Science Program*

Workshop segment	Steps	Length of time
Start-to-finish	Planning and delivering a workshop, including sourcing guest instructors.	2 months, typically
Installation party (pre-workshop set-up)	It is crucial to have time for participants to install the relevant software they will use in the session if they are to use their own computers	This could be a drop-in hour, ideally face-to-face
The workshop itself	Delivering a single, official Library Carpentry workshop.	4 days of 4 hours each, though you can be flexible
Following up after a workshop	To measure participants' engagement and support them to apply what they have learned to their individual roles and data sets.	A further 3 hours
Putting your staff through the Carpentries instructor training program	They can continue to deliver workshops without using guest instructors	2 full days
Leftovers (follow-up session)	The follow-up session will be incredibly useful for your participants to retain their new skills and improve their workflows so that their outputs are more reproducible, transparent and shareable. Outcomes and feedback about the andragogy, the snacks, next steps for continuing or revising, celebrating good instructors and providing thoughtful training and preparation materials	Half to a full hour

In the long term, individuals developing open science programs at their institutions may want to consider advancing a curriculum which maps to course work expectations across generations. See Table 1 for details about cooking times for supporting an open science program.

DIETARY GUIDELINES

Carpentries' curricula and resources offer pedagogically sound, ready-to-use content on data skills, research computing, and information literacy. These skills may be easily overlooked or rushed in formal classroom environments and self-studies. Information is also provided in an easily digestible format for a novice learner audience. Carpentries resources support recommended practices and save librarians from having to develop custom content or work with researchers to retroactively convert wild data into tidy formats. Librarians deliver Carpentries workshops to broad audiences, including research support staff and administrators, and incorporate relevant, meaningful examples and use cases to customize the content. This may include examples about how technical skills can be applied to systems, workflows, processes, and funding proposals to further the goals of open science.

INGREDIENTS AND EQUIPMENT

Assign the following roles (see figure 1) to your team of kitchen staff:

- one sous chef to coordinate and justify budget and time;
- one short-order cook, often from IT, to troubleshoot in the moment;
- two or more prep cooks who are interested in learning the material;
- one food critic to critique materials and make suggestions for tailoring to local ingredients and context;
- two or more farm-to-table vendors, who represent instructors or faculty members who oversee learners in need of scaffolded and customized coding instruction (like undergraduates) and individuals who may be in need of pedagogical support (the next generation of Carpentry instructors);
- one caterer or delivery driver to host logistics, whether you provide a large-scale, in-person training, a completely remote training, or a range of hybrid alternatives;

Figure 1. Scene from a workshop, modified image from The Carpentries' Instructor Training (CC-BY 4.0), https://carpentries.github.io/instructor-training/.

- one host to oversee everything from registering with the Carpentries to hiring an instructor, arranging space, and managing registration; and
- one busser to clean up outputs and files and pack up leftovers.

Your kitchen will require the following utensils and ingredients:

- a space for learning with minimal distractions
- computers, ideally with a second monitor (one per learner)
- administrator access to install software
- shaped sticky notes or high-contrast color cards and pens
- a projector or large screen for the room to see live coding clearly
- internet access to support downloads and software installation or an offline setup (https://carpentriesoffline.org/)
- electrical power supplies or sufficient battery backups
- printer or other means of providing access to pre- and post-workshop surveys and other materials that need to be delivered in hard copy

PREPARATION

Before cooking this recipe, make sure to read "Software Carpentry Al Dente" (Ossom-Williamson et al., 2022), which introduces readers to the Software Carpentry open source curriculum and defines instructor and participant workflows, including confirmation of availability, outreach and engagement, event preparation, and teaching and learning.

In addition to carrying out the preparation described by Ossom-Williamson et al., who informed the design, development, and implementation of our recipe, cooks implementing Library Carpentry or similar curriculums in support of open science should also consider the following:

- Identify the learning needs of your audience. What skills do they already have; what do they claim to need; what do you think they need? Use this information to tailor and adapt the lesson materials to your local context.
- Are there any qualified Carpentry instructors at your institution? What about at nearby or partner institutions?
- Discover local assumptions or presumptions that may contribute to barriers. What perennial complaints about knowledge gaps ring in the halls? These can help you to identify which workshop to request and, within that workshop, which elements to emphasize or skip.
- Will sessions be scheduled as two full days or four half days? Learners may prefer a gap between sessions to keep up with their regular work commitments.
- Will instruction be delivered virtually, in-person, or in a hybrid setting? Consider that virtual sessions and hybrid sessions are more labor and time intensive and may not be suitable for novice learners.
- How will faculty or students receive credit for completion? What will qualify as completion?

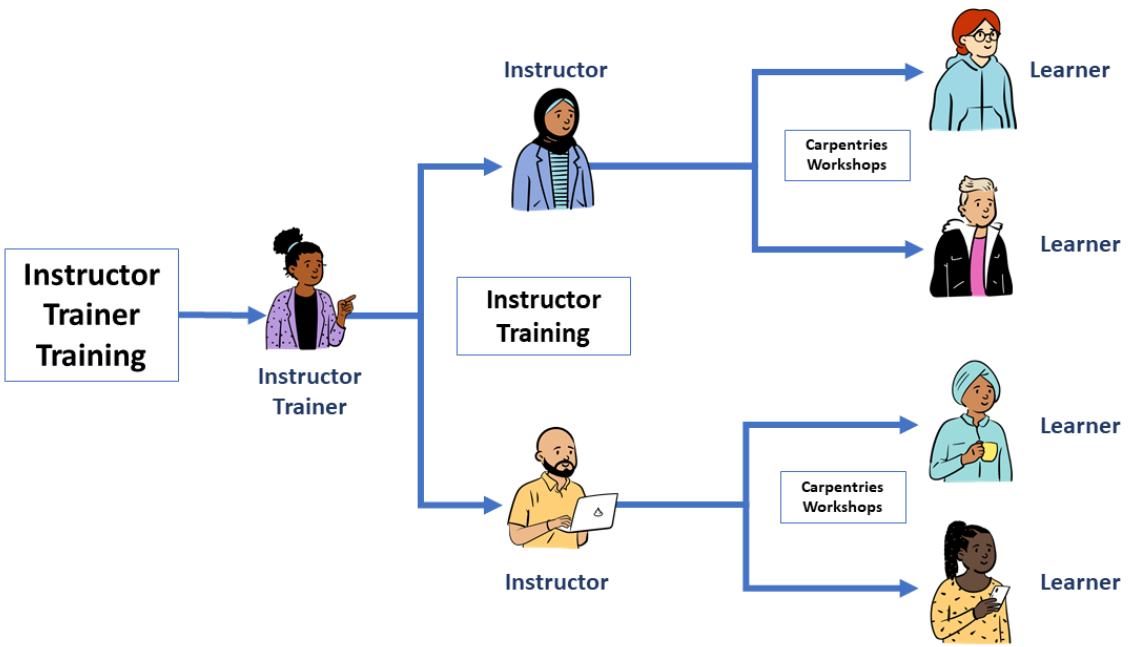

Figure 2. Local communities, from The Carpentries' Intro to Midwest Carpentries Call, September 2023 (CC-BY 4.0), https://docs.google.com/presentation/d/16NlqEKAMfaZgC_ZOGYiD1h6uUdFV1u-1IDi7pbdwHhE/edit#slide=id.ge267fc051b_0_107

COOKING METHOD

Our recipe extends and updates Ossom-Williamson et al by providing a quick reheat before deglazing additional details about how to cook up spicy data skills in support of an open science program (see figure 2):

- One, sort beans and remove sand. Instructors should convene to coordinate tasks, assess learners' needs, and make key decisions about the delivery of the workshop's curriculum, relevant topics for instruction, and teaching assignments.
- Two, pre-soak. Sign up with Carpentries' centrally organized workshops (https://amy.carpentries.org/forms/request_workshop/) to recruit instructors and coordinate and recruit additional volunteers to ensure adequate staffing.
- Three, prepare a menu through advertising. Finalize logistics, including dates, times, locations, advance registration requirements. Consider available technologies, and anticipate potential challenges.
- Four, shop for ingredients. Instructors explore the local culture which may include discussing specific learners' needs, preferred styles, languages, and accom-

modations.
- Five, arrange deliveries. Instructors select lesson elements to emphasize or skip.
- Six, receive deliveries. Instructors decide key logistics related to the workshop dates, times, spaces, advance registration requirements. Consider available technologies and anticipate potential challenges.
- Seven, promote the workshop. Coordinate with relevant departments and communications staff members to prepare and distribute customized messages and outreach materials.
- Eight, distribute pre-workshop surveys to learners, collect results, and share with instructors and organizers.
- Nine, host a pre-install party! Meet with learners in advance of the workshop to provide assistance and instruction on installing any required software applications. Make sure you are prepared to support learners of varying technical abilities and experience, both at the pre-install party and during the workshop.
- Ten, host the workshop.
- Eleven, serve the fruit and cheese course: Solicit learners' feedback.
- Twelve, clear the table. Debrief and discuss the workshop implementation process and survey feedback with fellow instructors/organizers. Use feedback to plan for adjustments in future programming, in both workshop content and form.
- Thirteen, host an after-party or picnic! Instructors can maintain momentum for an open science program by coordinating future, related workshops or adjusting ratios and seasonal spices and delivering the same menu to a new audience.

Virtual (online) or hybrid workshops are possible, but require the most of prep cooks. It is much more difficult to diagnose and resolve problems relating to software installations. Learners must be comfortable with webinar software and multitasking windows which can be a significant cognitive load for novice learners who may be intimidated by statistics or code. Virtual sessions require additional set-up and prep-time of about one or two days in advance and/or the use of personal equipment, depending on local IT lead time. Hybrid workshops have been reviewed at the UK Carpentries monthly discussions (https://hackmd.io/@local-uk) with consensus that they are the most challenging for all.

Additional care and consideration may be necessary when working with diverse learners, particularly those from non-STEM fields. For example, arts and humanities researchers may be comfortable working with a diversity of specialized materials and file formats (e.g., maps, digitized papyrus fragments, etc.), which may not be explicitly incorporated in the Carpentries workshops. These researchers may be examining art, artifacts, or objects using estimated date ranges spanning thousands of years and with archaic languages. These diverse learners might use different vocabularies to identify and describe the concepts in the Carpentries materials.

In terms of diversity in technical experience, consider supporting advanced learners with supplemental exercises, workshops, or courses. You can also challenge advanced learners to adapt to the question to their current experience. You can support novices by modifying instructions or vocabulary, using foundational exercises, providing more hands-on support during the workshop, and offering supplemental coaching or consultation sessions.

CHEF'S NOTES

The audience for Carpentries curricula is far wider than librarians alone. Several of this recipe's authors (based at the University of Manchester) have taught Shell and OpenRefine lessons for a graduate program in library and archive studies. Oliver et al. (2021) describe Library Carpentries lessons integrated into health, data, and information science programs in Europe and the United States. Cope et al. (2020) describe how digital skills and development have been improved across the library sector in the United Kingdom. In both cases, personnel included metadata and information specialists, institutional IT professionals, research data management teams, and digital humanists. Consider a wide range of job titles and professional areas when planning Carpentries workshops, including research administrators, grant writers, integrity officers, communications staff, human and animal subjects coordinating staff, adminis-

trative staff, and senior executives.

Robinson and Lowndes (2022) say that one critical and overlooked aspect in the journey to open science is to gain support from managers and leaders. They do not need to be experts in open science but leaders who support their teams and make strategic decisions about investment, training and innovation. In collaboration with NASA, Robinson and Lowndes developed the Openscapes Flywheel, which is a "framework for managers to facilitate and scale inclusive [o]pen science practices" (2022). They have already seen the ways in which adopting open science strategies have increased findability, accessibility, inclusion, and reproducibility in research at government, academic and non-profit organizations.

HOW TO ADAPT THE RECIPE FOR REPRODUCTION

Reflective chefs will review individual elements' contributions to local program replication. Collecting diners' feedback can refine and facilitate localization of later reproductions. Here are some suggestions for how chefs may wish to adjust the recipe on subsequent bakes.

Depending on the feedback received in workshops throughout the program, emphasize certain botanicals through relevant examples or real-world applications of materials to evidence the value of learning tools and methods. Skip the soup: when undertaking data cleaning for individual, stand-alone projects skip the version control episode. Adjust to a picnic because of a rainy day: switch workshop duration/format. Extend the timeline to accommodate different learning styles or family/life balance needs; adapt content to hybrid or more accessible venues.

To support advanced learners, add a souffle course of extra exercises, workshops, or courses (challenge diners to adapt the question to their current research); extend the menu by developing intermediate or advanced Carpentries lessons or additional supplemental open educational content (see Additional Resources for details). Bring your own (spices) data to bring the experience home within and across existing workshops.

To support novices across the program's menu, add foundational-level exercises (pureed food, applesauce, creamed peas). Offer supplemental coaching or consultation sessions to support data literacy (chopstick holders, dietary or nutritional counseling between workshops). Give extra explanation, paraphrase (alternative condiments). Give diners an opportunity to complete the exercises / work with them (fork vs chopsticks).

We leave you with a few tips to make #Open a daily staple of your diet:

- Write email content using Markdown.
- Save spreadsheets as CSV files.
- Store data in institutional repositories.
- Use open educational resources.
- Make data tidy; use a codebook.
- Highlight incremental changes to existing workflows, curricula, and engagements for learners to apply open science skills daily.
- Maximize opportunities for individuals across the institution to adopt and reinforce open science principles and practices—to minimize allergic reactions.

ADDITIONAL RESOURCES

- Multiple ways to follow up after Carpentries training, https://carpentries.org/blog/2023/01/multiple-ways-to-follow-up-after-carpentry-trainings-benefits-impacts-and-librarians-wishes/
- How to create a carpentry lesson, https://carpentries.org/blog/2023/10/cldt-curriculum-and-trainers/
- The Carpentries instructor training curriculum, https://carpentries.github.io/instructor-training/
- The Carpentries Handbook, https://docs.carpentries.org/
- Join the conversation, https://carpentries.topicbox.com/groups/community-development
- Library Carpentry lessons, https://library-carpentry.org/lessons/

REFERENCES

Cope, J., Ganeshwaran, N., & Reed, P. (2020) Building a Library Carpentry community in the UK. *UKSG eNews,* 478. https://www.uksg.org/newsletter/uksg-enews-478/building-library-carpentry-community-uk

Crawl, A;, Eliott, B., & Trusler, A. (2023). The Carpentries membership tip sheet. *The Carpentries*. https://doi.org/10.5281/

zenodo.8125209

Gofman, A. (2019). A Carpentries approach to ACRL Framework instruction. *JeSLIB*, e1173.

Oliver, J., Goldman, J., & Förster, K. (2021). *Integration and reuse of Library Carpentry content into curricula*. The Carpentries. https://librarycarpentry.org/blog/2021/03/integration_and_reuse_of_lc_content_into_curricula/

OpenAIRE. (2023). *How to make your data FAIR*. https://www.openaire.eu/how-to-make-your-data-fair

Ossom-Williamson, P., Williams, S., & Khan, H. R. (2022). Software Carpentry al dente: Rendering tech training for online artisans. In Kelly Getz and Meryl Brodsky (Eds.), *The Data Literacy Cookbook*, 70-72. ALA Editions.

Robinson, E., & Lowndes, J. S. (2022). The Openscapes Flywheel: A framework for managers to facilitate and scale inclusive open science practices. *Earth ArXiv*. https://doi.org/10.31223/X5CQ02

Wickham, H. (2014). Tidy data. *Journal of Statistical Software*, *59*(10), 1–23. https://doi.org/10.18637/jss.v059.i10