There is much discussion about the importance of digital skills requirements across the sector, involving groups such as RLUK Digital Scholarship Network, The National Archives, and Jisc Digital Capabilities. Such matters are particularly relevant for metadata and discovery (Daniels, 2020). The issues include establishing learners’ requirements, developing lessons and competency frameworks, and sourcing an affordable and sustainable pool of instructors. This paper will look at the delivery of Library Carpentry workshops in Manchester and Durham, how Library Carpentry can help to provide an opportunity to develop a community of good practice, and how these are relevant to metadata and cataloguing work. It ends with two case studies: what are people inspired to do after attending Library Carpentry workshops, and the development of a Carpentries-style lesson at Manchester about TEI XML (Text Encoding Initiative Extensible Markup Language).

Introducing Library Carpentry to Manchester

Staff at The University of Manchester Library (UML) have been inspired by the lessons and workshops presented by Library Carpentry (LC). In short, Library Carpentry teaches information professionals how to apply best practice when transforming and working with data sets; it is a non-profit organisation and an international community of volunteers who help people without a background in programming. In the last few years, three UML staff have qualified as LC instructors (Carlene Barton, Nilani Ganeshwaran and Phil Reed). They ran a workshop in November 2018 for their colleagues and those at nearby Manchester Metropolitan University, teaching good practice with tidy data and an introduction to OpenRefine. Feedback was positive and constructive; after one workshop, participants came away with skills and were eager to apply them (although some reported that they would not be able to use the skills right away due to commitments elsewhere).
Taking the show on the road

In December 2019, the Manchester Library Carpentry team were invited to deliver a workshop for North East England. Kindly hosted by Durham University, the syllabus was slightly amended to cover the concept of regular expressions as well as OpenRefine. The icebreaker section was about jargon busting, asking people to share an acronym, name, or concept for the hosts or other participants to explain, such as Unicode or CRIS. Everyone in the room learned something new within the first 15 minutes and could be reassured that nobody knows everything. The workshop was wrapped up by asking learners to set goals for using their new skills; what will they do after two weeks, and after three months (see the case study below for more about these goals).

The feedback was more positive from the Durham session; more people found it immediately useful. This is likely due to the application process and demographic; a greater proportion of people who manipulate data sets very frequently in their job came to the Durham workshop.

“Eye opening! Amazing. This will be really useful for anyone using spreadsheets.”

– feedback from a participant in Durham

Relevance to metadata and cataloguing librarians

There are many immediately relevant opportunities for librarians working with metadata and cataloguing, people for whom cleaning and transforming data are daily tasks. Learning OpenRefine is a way for people to do something immediately; for example, after attending the workshop, someone could open a CSV file and fix inconsistently formatted author names and initials. The Library Carpentry workshop lays a foundation for learning more computational processes; for example, one could learn how to write simple Python scripts to fix inconsistent formatting in a batch of files (other lessons which can be used in a longer Library Carpentry workshop are listed in a later section of this article).

Community is key

Library Carpentry is one of three communities within a parent group called The Carpentries. The other communities of Software Carpentry and Data Carpentry are more mature (dating back to 1998) and are usually taught to academic researchers and students. All lesson content across The Carpentries has been developed by volunteers and is available under a Creative Commons Attribution licence. Workshops are delivered by volunteers and can be arranged centrally (for a fee) or by one of the 80 member organisations from 10 countries around the world. Workshops largely follow a set pattern or syllabus and must be taught by at least one qualified Carpentries instructor. All necessary guidance is provided in the handbook, including the workshop Code of Conduct which ensures a safe and comfortable learning environment (The Carpentries, 2020).

There are sound, evidence-based pedagogical practices applied to all Carpentries lessons, taught to all instructors before they qualify (Allen et al., 2019). For example, formative assessment is used frequently, dismissive language is not permitted, and mistakes are used as learning opportunities. This quality foundation has helped in the delivery of 2,300 workshops run in 61 countries between 2012 and 2019.
Working with data is important in fields such as biology, where researchers and support staff are increasingly required to handle “big data” and apply “FAIR Principles” for data management\(^1\) (Wilkinson et al., 2016). It has been shown that sharing good practices using the Carpentries model will “foster open, transparent, and sound scientific results beneficial to society” and that developing a community of practice “will bring many benefits to its members and to their institution” (Stevens et al., 2018). The benefits of applying FAIR Data Principles when working with metadata have been clearly established (Ball, 2020). It follows, therefore, that developing a community of practice for metadata specialists using Library Carpentry is an opportunity well worth exploring.

**Looking into the future**

The Manchester Library Carpentry team would like to run more Library Carpentry workshops and inspire more people to join as instructors. In the short term, any workshops might have to run online instead of in person, so more time will be required to investigate how that could work most effectively.

The team would also like to encourage people to contribute to lesson development. Anybody can suggest edits to the lessons hosted on GitHub; see the ‘Repository’ links on the lessons page. Each repository has a list of ‘Issues’ which are discussions about any changes between maintainers and with the public.

In addition to the tidy data, regular expressions, and OpenRefine lessons, Library Carpentry has a core curriculum designed to run over two days (though you can split these into four parts or half-days). These lessons include the UNIX shell and an introduction to Git. The curriculum can be extended to include SQL, webscraping, Python and data for archivists, although some of this content is in the early stages of development. There is a standalone lesson titled ‘Top 10 FAIR Data & Software Things’ which may be of interest to those working in research libraries and many others. Further lessons at the earliest stages of development include MarcEdit, Wikidata, R and XML.

Finally, you may wish to use the principles of Carpentries lessons, or indeed the lesson templates, following the Creative Commons Attribution licence. There are examples of lesson templates being adapted and reused from Neurohackweek, 23 (Research Data) Things and The University of Manchester Library’s Introduction to TEI (see case study 2 below).

**Case study 1: What can people do after Library Carpentry?**

At the end of the Durham workshop, learners were asked “Based on what you have learned today, what will you do differently in the next two weeks? Three months?”. Some of the responses received were:

- “Process chain of Excel commands but in OpenRefine.”
- “Catalogue data, although getting it in and out of the catalogue is a factor.”
- “Harmonising the way data is presented across data sets, e.g. author names.”
- “Find faulty terms in thesaurus that broke. Data recovery after corruption.”
- “Learn more about APIs.”
- “SCONUL stats, renewals.”
- “RegEx to collate reports, who has been using services.”
- “Survey data, cleaning up.”

Learning how to automate or articulate one’s work using scripts may save time in the long run, and also increase the transparency and accountability of workflows.

\(^1\)“FAIR Principles” mean that digital assets are findable, accessible, interoperable, and reusable.
Case study 2: Introduction to TEI

‘Introduction to TEI XML’ was a series of face-to-face workshops at The University of Manchester Library delivered in November and December 2019. It was designed by Jane Gallagher, Elizabeth Gow, Dr Jo Edge, and Phil Reed, with special thanks to Professor David Denison, Nuria Yáñez-Bouza, Dr Giles Bergel, and Dr Christopher Ohge. It was aimed at Special Collections staff, cataloguers, and other library and information professionals. There was formative assessment throughout, aligning with the learning outcomes, and opportunities for collaborative learning.

After the workshops completed, the materials were stored on the local intranet, and were thus unavailable to those outside the initial group. To adapt the workshop into an open educational resource, the author followed the open practice exemplified by The Carpentries and turned the materials into a lesson hosted on GitHub. The resource was released with a Creative Commons Attribution Non-commercial licence. Some adaptation was required; however, the initial approach was a close match pedagogically to the Carpentries format so this was quite straightforward.

It is currently a functional ‘alpha’ release with various small adjustments required. One day, it could inspire the development of the suggested Library Carpentry XML workshop. If the release is shown to effective, the approach of building Carpentries-style lessons may be suggested for other forms of staff development at Manchester. Meanwhile, Software Carpentries lessons continue to be taught at Manchester about Python and the Unix shell. These workshops have been running for three years, showing the ongoing value of the Carpentries approach to researchers.

References


