

Dear Colleague,

One of the biggest challenges in research these days is to ensure that our students and researchers are equipped with skills to handle the ever increasing size of their datasets and use technology efficiently to perform reproducible analysis.

To address this issue several international initiatives have been established over the last few years. The North-West University's eResearch Initiative will be hosting four remarkable organisations who are currently tackling the challenge of computational training for researchers and research support staff.

Our guests include:

- Jonah Duckles – *Executive Director, the Software Carpentry Foundation*
- Tracy Teal – *Executive Director, Data Carpentry*
- James Baker – *Founder, Library Carpentry*
- Simon Hodson, Andrew Harrison, Hugh Shanahan - *CODATA-RDA Working Group on Research Data Science Summer Schools Co-Chairs*

The workshop will be open for online or in-person participation. More information about the various initiatives are summarised below with references.

| | |
|------------------------------|--|
| Date: | 11 April 2016 |
| Time: | 14:30 – 16:30 |
| RSVP: | Please RSVP for physical and virtual participation by 6 April - http://goo.gl/forms/wq5FpKwe4v |
| Physical venue: | Building C1, Room 135, Institutional Office, North-West University, Hoffman Street, Potchefstroom |
| Online participation: | Link to be confirmed |
| Enquiries: | Anelda van der Walt – anelda DOT vanderwalt AT nwu DOT ac DOT za |

THE SOFTWARE CARPENTRY FOUNDATION - TEACHING BASIC LAB SKILLS FOR RESEARCH COMPUTING [1]

Software Carpentry was developed by Dr Greg Wilson and ran for the first time in 1998 at the Los Alamos National Laboratories [2]. Over the course of the last 18 years the organisation has evolved and in 2014 the Software Carpentry Foundation was established [3]. Since 2011 more than 16'000 researchers and postgraduate students from six continents have gone through training and there are now more than 400 qualified instructors worldwide [4]. The Software Carpentry Foundation not only teach researchers tools that enable more efficient and reproducible research, but they have also developed an outstanding instructor training programme based on empirical studies about computational education [5].

Software Carpentry lessons are collaboratively developed and published under the Creative Commons license. Current lessons include material on automation using the Linux shell and Make, tools such as git and GitHub for version control and collaboration, databases with MySQL, and programming with Python, R, and MATLAB [6].

The workshops typically take place over two full days with instructors live coding while participants follow along on their own laptops with pre-installed software (not limited to any specific operating system). More recently attempts has been made to integrate the lesson material and model into university courses and textbooks [7,8].

Since 2013 eight Software Carpentry workshops took place in South Africa. In the twelve months (November 2014 – November 2015) more than 200 researchers and postgraduate students from more than 18 organisations participated in six of these workshops across four provinces [9]. Participants hail from disciplines as diverse as life sciences, medicine, agricultural sciences, economics, environmental sciences, law, libraries, IT, and engineering.

DATA CARPENTRY – MAKING DATA SCIENCE MORE EFFICIENT [10]

Where the Software Carpentry lessons are subject agnostic and workshops typically include participants with various research backgrounds, Data Carpentry, under the leadership of Dr Tracy Teal, specifically focuses on high-quality, domain-specific training for researchers working in data-intensive research areas.

In 2015 Data Carpentry received USD 750'000 in funding from the Gordon and Betty Moore Foundation to support activities such as (i) expanding into new domains, (ii) train more researchers, (iii) increase the number of Data Carpentry instructors, and (iv) to develop strategies for delivering their material online [11].

Current material on offer include workshops with a Biological focus looking at the whole life cycle of data processing and includes tips on organising data properly in Excel to data management in SQL and the use of either R or Python in combination with these tools. A Genomics workshop is also available and include lessons on data organisation, cloud computing, command line usage, and data visualisation in R [12].

Lessons currently under development include material for Geospatial Data Carpentry, Social Sciences Data Carpentry, Data Visualisation, and Working with APIs [13].

[LIBRARY CARPENTRY – SOFTWARE SKILLS FOR LIBRARY FOLKS \[14\]](#)

In February 2015 Dr James Baker, Lecturer in Digital Humanities, asked the question “What would Library Carpentry look like? [15]. Subsequently he joined the Software Sustainability Institute and developed material for the first ever Library Carpentry event held in the UK in November 2015 [16]. The workshop ran as four three-hour session spread over four weeks. Content included basic computational skills (including “jargon busting” and regular expressions), the Linux shell (using commandline tools to mine tabulated data and work with text), version control (git and GitHub for version control, publishing, and collaborating), and OpenRefine (for cleaning data and more) [17].

59 participants from 14 organisations participated in the seminal Library Carpentry series of workshops and the feedback was overwhelmingly positive [18]

[THE CODATA-RDA SCHOOL OF RESEARCH DATA SCIENCE – TEACHING DATA SCIENCE APPROACHES AND SKILLS THAT ARE ESSENTIAL FOR 21ST CENTURY RESEARCH \[19\]](#)

CODATA, the Committee on Data for Science and Technology, is an interdisciplinary Scientific Committee of the International Council for Science (ICSU) [20] with the aim of improving the quality, reliability, management and accessibility of data of importance to all fields of science and technology. CODATA was established in 1966. One of the organisation’s working groups, the *CODATA-RDA Working Group on Research Data Science Summer Schools*, specifically aims to develop an accredited and recognised curriculum to address the need address recognised need for Research Data Science skills across disciplines. They also work towards developing a programme to train trainers to ensure the sustainability of the program [21].

The first CODATA-RDA Research Data Science Summer School will take place from 1 -12 August 2016 in Trieste, Italy [22]. The school is mainly targeted at early-stage post graduate students from developing countries. The school will include material on [23]:

- The basic Software Carpentry curriculum
- Aspects of data handling from the Data Carpentry curriculum
- Open Research Data
- Research Data Analysis
- Research Data Visualisation
- Computational Infrastructures for Research Data Analysis

References

1. Software Carpentry – <http://www.software-carpentry.org>
2. Software Carpentry: lessons learned version 1 (Feb 2014) - <http://f1000research.com/articles/3-62/v1>
3. Software Carpentry: lessons learned version 2 (Jan 2016) - <http://f1000research.com/articles/3-62/v2>
4. Recent Statistics - <http://software-carpentry.org/blog/2015/10/recent-stats.html>
5. Our New Instructor Pipeline - <http://software-carpentry.org/blog/2016/02/new-instructor-pipeline.html>
6. Software Carpentry Lessons - <http://software-carpentry.org/lessons/>
7. Software Carpentry as a University Course - <http://software-carpentry.org/blog/2016/02/swc-as-a-university-course.html>
8. Software-Carpentry-Inspired Book: Released and On Sale! - <http://software-carpentry.org/blog/2015/07/effective-physics-book.html>
9. "A Year Of Software Carpentry in South Africa" - <http://software-carpentry.org/blog/2016/01/a-year-of-swc-in-south-africa.html>
10. Data Carpentry - <http://www.datacarpentry.org/>
11. Data Carpentry receives grant from Moore Foundation - <http://www.datacarpentry.org/blog/moore/>
12. Data Carpentry Existing Lessons - <http://www.datacarpentry.org/workshops-offered/>
13. Data Carpentry Workshops Under Development - <http://www.datacarpentry.org/workshops-development/>
14. Library Carpentry - <http://librarycarpentry.github.io/>
15. What would Library Carpentry Look Like? - <http://britishlibrary.typepad.co.uk/digital-scholarship/2015/02/what-would-library-carpentry-look-like.html>
16. Library Carpentry – Where and When - <https://librarycarpentry.github.io/timeplace/>
17. Software Skills Training for Librarians - Programme outline and FAQ - <http://librarycarpentry.github.io/outline/>
18. Library Carpentry in Words and Numbers: All Code No Woodwork - <http://cradledincaricature.com/2015/12/01/library-carpentry-in-words-and-numbers-all-code-no-woodwork/>
19. The CODATA-RDA School of Research Data Science - <http://indico.ictp.it/event/7658/>
20. CODATA Mission - <http://www.codata.org/about-codata/our-mission>
21. CODATA-RDA Working Group on Research Data Science Summer Schools - <http://www.codata.org/working-groups/research-data-science-summer-schools>
22. First CODATA-RDA Research Data Science Summer School - <http://indico.ictp.it/event/7658/>
23. The CODATA-RDA School of Research Data Science Poster - <http://indico.ictp.it/event/7658/material/poster/0.pdf>