## WORKSHOP OUTLINES AND REQUIREMENTS

***NOTE: Please ensure that you can connect to the Eduroam wireless network so you can use your own institutional credentials to authenticate and use the WiFi. Please refer to your university's IT department for assistance in connecting to Eduroam prior to attending this event. If you are not able to connect to Eduroam, a ResBaz WiFi network will be available.***

### Programming with Python: 2-day Software Carpentry

- Course outline: https://goo.gl/W4CD70
- Content includes:
    - Unix/Linux commands
    - Python syntax, loops, if statements, creating functions
    - Version control with Git and Github

### Programming with R: 2-day Software Carpentry

- Course outline: https://goo.gl/W4CD70
- Content includes:
    - Unix/Linux commands
    - R syntax, loops, if statements, creating functions
    - Version control with Git and Github

### Programming with Python - Beginner

- Course outline: https://goo.gl/Y5Wxhr
- No additional pre-course info

### Programming with R - Advanced

- Content outline:
    - Compiling Reports with RMarkdown
    - Applying Statistical Modelling with R
    - Advanced Graphics
    - Scheduling and Sending reports
- Participants are expected to have the most recent version of R (3.4.0) and RStudio (1.0.143) installed on their laptops
- A list of packages we will use are: Dplyr, Ggplot2, Knitr, Markdown, Plotly, Plyr, RColorBrewer, Shiny, Tidyr and more...

### Cleaning and Exploring your data with Open Refine

- Course outline: https://goo.gl/Wq6Vmf
- Please download Open Refine (version 2.7-rc2) on your laptop prior to the course: http://openrefine.org/download.html

## Data Viz with D3.js

- Course outline:
  - D3.js is a popular and powerful toolkit for visualising data on the web, see https://d3js.org/ for many examples.
  - This workshop will give you a hands-on introduction to the basics of d3.js:
    - loading data and joining it to a simple chart
    - interactive charts
    - dynamic graph visualisations
- Assumed skills:
  - familiarity with CSV
  - basic data-cleaning concepts
- Software requirements:
  - The workshop will use the Glitch platform: https://glitch.com/ which runs in any modern web browser – Chrome or Firefox are the preferred options. You can get a head-start by creating an account for yourself on Glitch.

## Community mapping tools

- Course outline: In this workshop participants will try out open source mapping tools that are the foundation for new community mapping initiatives and documentation projects. After playing around with OpenStreetMap and Fieldpapers, we will explore how they fit into the long history of mapping and discuss how we can extend this technology in humanities research and teaching.
- Participants should create an http://openstreetmap.org account prior to the workshop.

## Getting started with Data Visualisation

- Course outline: This information session covers the basics of data visualisation, including what data visualisation is, some things to consider before you start making a visualisation, and underlying principles to help in creating effective visualisations. As the session is intended for participants in any and all disciplines, we will not be focusing on specific tools or niche domain techniques. Instead we will consider the ideas behind visualisations as communication tools, and how we can tailor them to have the most impact and effectiveness in delivering our message.
- Computers will not be used in this session.

## Data Management and Data Publishing

- Course outline: This session provides you with all the info you need to know to get a start on managing and publishing your research data. Come along to learn the basics of research data management and get some hands-on experience using the resources you need to find and publish research data. This session will be non-discipline and non-institution specific.
- Please bring along a computer

## Web scraping

Web scraping is a technique for turning web sites into structured data. This requires instructing the computer how to identify certain parts of a page as targets of extraction, and often following links to spider through a large number of pages where the content resides. We will workshop these general techniques, while discussing issues about building scrapers that are robust to change, the ethics of scraping and the choice of scraping tool.

- Outline:
    - Introduction to web scraping
    - Matching elements of a web page with CSS selectors  (note we also have material on the more expressive alternative, xpath, and can assist participants who are bored of CSS selectors)
    - Visual scraping in the web browser
    - Coding a scraper in Python with requests and lxml
- Knowledge assumed:
    - Basic understanding of the www: knows what a URL is; has seen some HTML before
    - Basic python: loops and functions (though participants will still gain general skills from the class without the Python exercises)
- Tools required:
    - Google Chrome
    - Webscraper.io browser extension for Chrome: https://chrome.google.com/webstore/detail/web-scraper/jnhgnonknehpejjnehehllkliplmbmhn?hl=en
    - Python 3 (at least version 3.4)
    - requests package for python: http://docs.python-requests.org/en/master/user/install/
    - lxml package for python: http://lxml.de/installation.html
    - cssselect package for python: https://pypi.python.org/pypi/cssselect

(One way to get the python packages is by installing Anaconda from www.continuum.io/downloads)

## Excel Fu

- Course outline: https://goo.gl/qk15ZI
- Please make sure you have Excel (2007 or later) installed on your laptop prior to the course.