

CODATA-RDA Research Data Science School Curriculum

Hugh Shanahan

Royal Holloway, University of London

Hugh.Shanahan@rhul.ac.uk @hughshanahan

This talk and poster at
<http://github.com/HughShanahan/ElixirCODATARDAschools>



Outline of School

Two week event

Oriented towards Early Career Researchers

“Simple things, explained well”

Show possibilities - not deep dive

Show communities



Open Science

- Louise Bezuidenhout



3 Hours + Exercises

Reflective sessions - focus on Ethics

What does Open Science mean to you?

Is it worth it?

Exercises for each other course



Software Carpentry

- Multiple authors



Introduction to the command line

Git

R - language of choice for School

SQL (deprecated)



Research Data Management

- Kevin Ashley/Sarah Jones/Steve Diggs



Data Life Cycle

Licensing - appropriate use

Data Management Plans

Principles of FAIR data



Author Carpentry

- Gail Peretsman-Clement



Authorship/publishing for the 21st century

Digital Object Identifiers

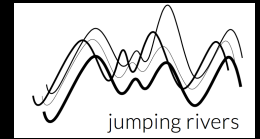
OrCID's

Predatory journals



Visualisation

Colin Gillespie/Jamie Owen



Introduction to ggplot2

Building plots

Faceting



Analysis

Roger Barlow/Ekpe Okorafor



Introduction to Machine Learning

Recommender Systems

Neural Networks / Classification



Computational Infrastructures

Rob Quick/Gergely Sipros



Moving beyond a desktop/laptop

Running jobs on grids/clouds

Using containers



Information Security

Hannah Short/Sebastian Lopienski



New for 2018

Being open != not being secure

File permissions

Outlining attacks in buffer overflow

Basic security protocols



Final thoughts

Perfect is the enemy of good

Iterative improvement is powerful

What's missing? - Experimental Design

