

CODATA-RDA Research Data Science School Curriculum

Applying bioinformatics and data science competency
frameworks to Elixir Training, Amsterdam, April 4-6, 2018

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 - focus on LMIC's

Data Science skills for Early Career Researchers

“Simple things, explained well”

Show possibilities - not deep dive

Show communities



Next schools

Foundational School ICTP, Trieste August 6-17

Apply via <http://indico.ictp.it/event/8329/>

Advanced school ICTP, Trieste August 20-24

Apply via <http://indico.ictp.it/event/8561/>

Foundational school Kigali, Rwanda October 22-November 2



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

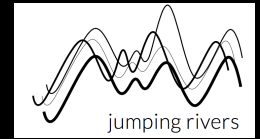
OrCID's

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



Acknowledgements

Many funders, many people...

CODATA Task Group chairs

Marcela Córdoba, Universidad de Costa Rica

Sarah Jones, DCC

Ciira Maina, DKUAT, Kenya

Rob Quick, Indiana University

CODATA

Simon Hodson

ICTP Schools

Clement Onime, ICTP

Ivan Girotto, ICTP

RDA Interest Group

Andrew Harrison, University of Essex

