A Curriculum For Foundational Research Data Science Skills For Early Career Researchers

Hugh Shanahan, Royal Holloway, University of London, UK
Paula Martinez, National Imaging Facility (NIF) / University of Queensland, AU
Jacqueline Stephens, South Australian Health and Medical Research Institute (SAHMRI), AU
This is one of many Outputs produced by an RDA Working or Interest Group. Learn more about RDA Outputs at the next RDA Plenary.

Early Bird Registration closes 14 February 2020, 23:59 UTC
A curriculum for foundational Research Data Science skills for Early Career Researchers

RDA Webinar
February 4 2020

Hugh.Shanahan@rhul.ac.uk
@hughshanahan
Gaps
Gaps

Identified gap in skills for ECR’s
Gaps

Identified gap in skills for ECR’s

Opportunity/Challenge - LMIC’s (not just HIC’s)
Gaps

Identified gap in skills for ECR’s

Opportunity/Challenge - LMIC’s (not just HIC’s)

Practices for what comes next - FAIR, Open ...
Gaps

Identified gap in skills for ECR’s

Opportunity/Challenge - LMIC’s (not just HIC’s)

Practices for what comes next - FAIR, Open …

Isolated ECR’s
Data Science Skills

Responsible Research Practices

Building Communities
Objectives
Objectives

Short-Medium term

Run expanding series of schools in Research Data Science regionally across the world to ECR’s
Objectives

Short-Medium term

Run expanding series of schools in Research Data Science regionally across the world to ECR’s

Long term

Curriculum (or close to) gets taught by HEI’s globally to ECR’s
CODATA-RDA schools

Joint effort between CODATA and RDA
CODATA-RDA schools

Joint effort between CODATA and RDA
CODATA-RDA schools

Joint effort between CODATA and RDA

CODATA - Committee on Data of the International Science Council (ISC).
CODATA-RDA schools

Joint effort between CODATA and RDA

CODATA - Committee on Data of the International Science Council (ISC).

"CODATA exists to promote global collaboration to advance Open Science and to improve the availability and usability of data for all areas of research. “
CODATA-RDA schools
CODATA-RDA schools

Two week event

Focus on ECR’s from LMIC’s

Technical + Responsible Research skills = Research Data Science

NOT another bootcamp!
Curriculum
Curriculum

Open and Responsible Research
Curriculum

- Software Carpentry
- Computational Infrastructures
- Visualisation
- Open and Responsible Research
- Research Data Management
- Author Carpentry
- Information Security
- Analysis
Principles
Principles

Reuse materials - it’s out there, use it.

Materials available under open source licence.

Mix ‘synchronised typing’ with seminars with discussion with team work.

Trieste, Italy is a test bed - other schools are regional

Use as many regional instructors as possible.

Recruit from student body.
State of play

By February 2020 we have delivered ....
State of play

By February 2020 we have delivered ....

9 Two Week Schools on 4 continents
State of play

By February 2020 we have delivered ....

9 Two Week Schools on 4 continents

~400 students taught from 40 countries
State of play

By February 2020 we have delivered ....

9 Two Week Schools on 4 continents

≈400 students taught from 40 countries

In British figures corresponds to ≈£250K of teaching
State of play

By February 2020 we have delivered ....

9 Two Week Schools on 4 continents

~400 students taught from 40 countries

In British figures corresponds to ~£250K of teaching

ALL AS A VOLUNTEER NETWORK
Data Steward strand

**Partnership**

“The FAIRsFAIR project addresses, in a 36 months timeplan, the development and concrete realisation of an overall knowledge infrastructure on academic quality data management, procedures, standards, metrics and related matters, based on the FAIR principles."

Run three instances of the school in Europe

Train-trainer events
Idea

Add a Data Steward strand to school.

Data Steward students work with ECR but have their own modules as well.
Software Carpentry

Information Security

Open and Responsible Research

Research Data Management

Author Carpentry

Information Security
Data Steward specific topics
Data Steward strand
Data Steward strand

First pilot ran August 5-17, Trieste

Five Data Steward students

Further schools will pair Data Stewards with ECR teams
Remember

Most of all what’re doing here is
Remember

Most of all what’re doing here is

Changing a culture

Getting people to work together

Building communities that follow an approach
Remember

Most of all what’re doing here is

Changing a culture

Getting people to work together

Building communities that follow an approach

That is really, really, hard - but more effective in the long run
Recent Sponsors

- Committee on Data
- ICTP
- RDA
- Springer Nature
- H3ABioNet
- TWAS
- CeNAT
- African Open Science Platform
- CONARE
- ITU
- Software Sustainability Institute
- FAIRSFAR
- R Studio
Task Group chairs

Louise Bezuidenhout

Raphael Cóbe

Marcela Alfaro Córdoba

Sara El Jadid

Bianca Peterson

Rob Quick

HS

S. Venkataraman
RDA Output for Curriculum DOI: 10.15497/rda00038

https://github.com/CODATA-RDA-DataScienceSchools/Materials

https://github.com/CODATA-RDA-DataScienceSchools/TaskGroup
The CODATA–RDA Research Data Science Summer School 15–17 August 2018
Miramare, Trieste, Italy

Dr Paula A. Martinez
National Imaging Facility,
The University of Queensland, Australia
ELIXIR is the European organisation coordinating, managing and integrating life sciences data and infrastructure resources to enable research for academic and industry users.

April 2018 Competency Frameworks workshop.
Engagement

It resonated with me because of the audience and the open communities involved.
ELIXIR joined CODATA and RDA to organise the international summer school in research data science #dataTrieste18.

- Software Carpentry R
- Advanced Summer School in Bioinformatics https://doi.org/10.5281/zenodo.3381146
A vibrant experience in class

Rob Quick
@RobQuick5

datastrieste18 is underway! 62 invitees from 32 countries learning #opencode, #opendata, UNIX Shell, git and R. A lot of ground to cover in only two days! I’m already proud of what this group has accomplished and there are 8 more days still ahead of us!!
People coming together to learn from one another

It was an experience to never forget. We are still in touch and that makes it even better :)

![Group photo](image1.jpg)

![Scenic view](image2.jpg)
Coming together is a beginning. Keeping together is progress. Working together is success. – Henry Ford via Mina Zamani

At the end of the first week 😊 with lovely friends from around the world at the CODATA-RDA Research Data science summer school #datatrieste18
More on the RDA and RDA Plenary 15

- RDA homepage: https://www.rd-alliance.org/
Thank you very much!