Training and Capacity building in FAIR data stewardship in the Netherlands

Celia van Gelder, Mateusz Kuzak & Christine Staiger (DTL/ELIXIR-NL)

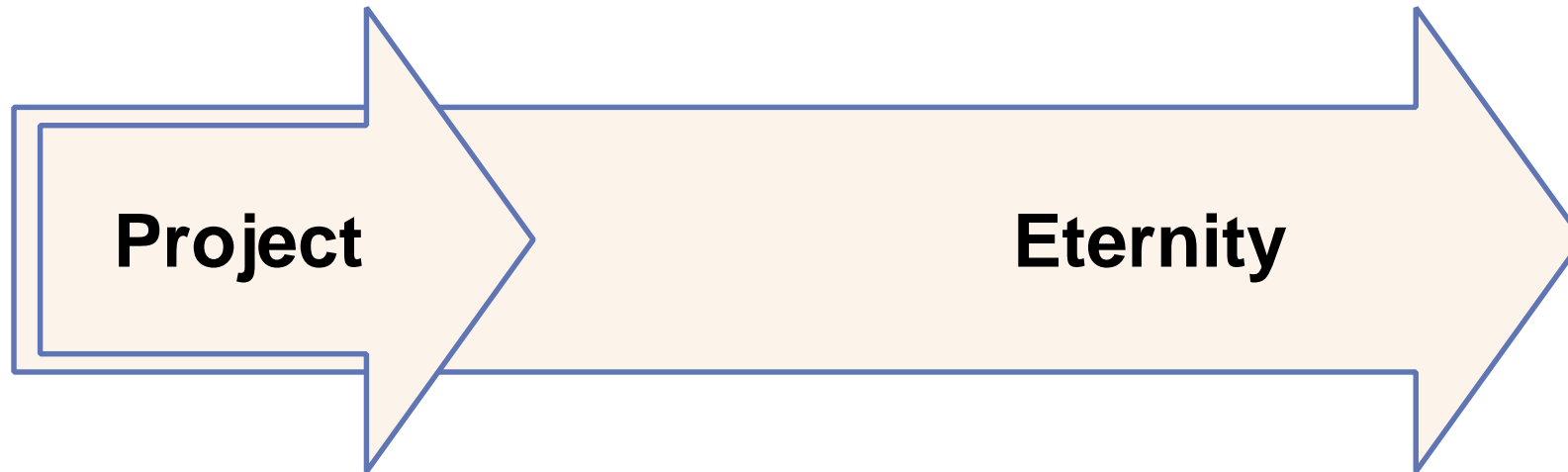*BioSB Conference, 3 April 2019*

www.elixir-europe.org

Data Stewardship (DS)

Responsible planning and executing of all actions on digital data before, during and after a research project, with the aim of optimizing the usability, reusability and reproducibility of the resulting data

Project **Eternity**

Capacity building in data stewardship is needed

From EOSC High Level Expert Group, Oct 2016:

- "There is an alarming shortage of data experts both globally and in the European Union"
- "The lack of core intermediary expertise has created a chasm between e-infrastructure providers and scientific domain specialists"
- "Core data experts need to be trained and their career perspective significantly improved"
- "I3: Fund a concerted effort to develop core data expertise in Europe"

- https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none

FAIR Data Management and Stewardship skills are needed

- Skills in data management (DM) & data stewardship (DS) are crucial to deal properly with the data

- Furthermore DMPs are obligatory for many if not all funding applications

- DS has many many many aspects (nobody can do nor has to do all steps of the DS)

- Who to train? Which target audiences?

- What are the skills needed for each target audience? How to train in these skills?

Training in Data Stewardship – to whom

Skills and competencies for:

- **Researchers**: create, manage and share FAIR data

- **Data Stewards**: provide support for research teams, data repositories, data services

- **Trainers**: teach and apply FAIR principles

- **Reviewers and funders**: assess FAIRification of datasets, data resources in publications, grant proposals and reports

Data Stewardship related activities in ELIXIR-NL/DTL

- Defining the Data Steward profession in the life sciences
- Curriculum Building
- Training courses
- Community building
- Making training FAIR
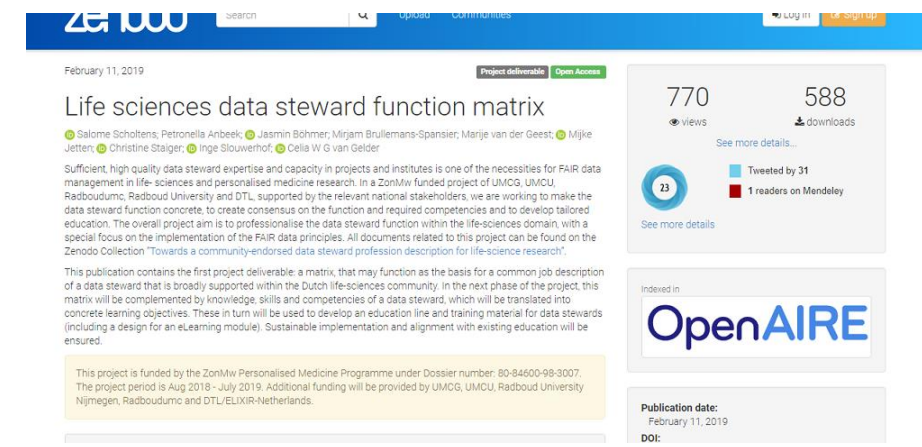- Tool: ELIXIR Data Stewardship Wizard
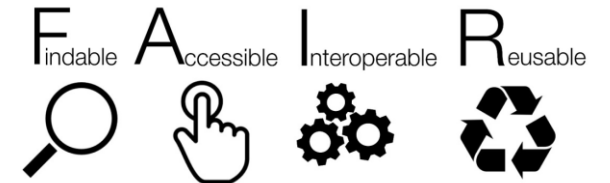
- Most activities are across projects and across the NL border
- Some activities connect to other than Life Science domain

Defining the DS profession:
Towards FAIR Data Steward as profession for the Life Sciences

- A collaborative approach built on existing expertise in order to increase the quality and capacity of data stewards for life sciences.

- Aim: professionalise the data steward function within the life-sciences domain, with a special focus on implementation of the FAIR principles and endorsed by all stakeholders in NL

- By developing training for existing and new data stewards, creating a standardized job description, formalising the function and creating more visibility of the function.

- https://zenodo.org/communities/nl-ds-pd-ls/

- Contact: Salome Scholtens (UMCG), Celia van Gelder (DTL)

First deliverable – Life sciences data steward function matrix

- Presented at Health-RI Conference Feb 2019, *DOI:10.5281/zenodo.2561723*

- Created by a) analyzing ca 45 job advertisements and b) analyzing and mapping to existing data stewardship competency frameworks

- For 2 types of data stewards:

  – Data Steward A is an institutional/department focused and coordinating, policy-oriented function

  – Data steward B is project focused and an operational, supporting-oriented function

- Matrix well received, both nationally and internationally

Life sciences data steward function matrix

Nine different function areas were defined for both data steward A & B:

- 1. policy/strategy,
- 2. compliance,
- 3. alignment with FAIR data principles,
- 4. services,
- 5. infrastructure,
- 6. knowledge,
- 7. network,
- 8. data archiving
- 9. supervision.

- Living document ! At the moment we are incorporating all feedback from the community

Data steward Level A - Institutional, coordinating, policy

Ongoing Work

- Definition of the knowledge base, skills and learning objectives ,
  DOI: https://doi.org/10.5281/zenodo.2616477

- Next steps:
  - Education line,
    training material, pilot run,
    script for an eLearning module,
    sustainability plan to ensure
    national embedding

- **Save the date: Workshop 11-12 June 2019, Utrecht**

Table 2: Example of learning objectives following Bloom's Taxonomy

RESPONSIBILITY:
Compliance

KNOWLEDGE:
Knowledge about legislation, ethics, code of conducts and research codes with regard to (research)data, (medical) research and privacy

SKILL/ABILITY:
Translate legislation and codes of conducts with regard to research data to practical implications and guidelines that researchers can understand

LEARNING OBJECTIVES:
-Explain the legislation and codes of conduct with regard to research data to researchers (applying)
-Examine and list the practical implications of legislation and codes of conduct with regard to research data (analyzing)
-Create guidelines based on legislations and codes of conduct with regard to research data (creating)

Figure 1: Venn diagram of different data steward types
Figure acknowledgement: Christine Staiger, DTL
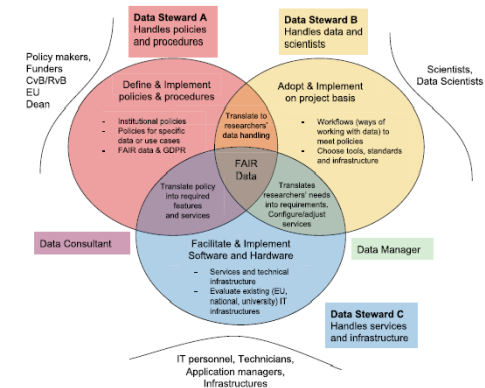
Curriculum building: Fair data stewardship curriculum

- Lessons developed around the FAIRification process
  - Lesson objectives directly connected to the FAIR metrics (and the Community Challenges)
  - Each lesson is an independent module
  - Modelled after the Carpentries, using the best practices by the Carpentries community : The Carpentries Handbook, Ten Simple Rules for Collaborative Lesson, Development Curriculum Development Guide, use Carpentries templates
  - will be linked to/from relevant parts of Data Stewardship Wizard
  - CC BY license
- Certification
  - Open Badges
- **GO FAIR** - GO-TRAIN Implementation Network
- Contact: Mateusz.Kuzak@dtls.nl

Slide: Mateusz Kuzak

DTL | DUTCH TECHCENTRE FOR LIFE SCIENCES

elixir

Data Stewardship Training activities of DTL/ELIXIR-NL

- Helis Academy
  - Postgraduate courses for Data Analysis and Stewardship for Industry and Academia

- Carpentries-NL
  - Software Carpentry Workshops, Data Carpentry Workshops, Instructor Trainings (building training capacity)
  - Contact Mateusz.Kuzak@dtls.nl

- Training for ZonMw Enabling Technology Hotels call 2018
  - Create awareness, assessment of FAIRness, building FAIR expertise in the technology centers

- ELIXIR Implementation Study: Towards Data Stewardship in ELIXIR: Training & Portal
  - Nodes involved: ELIXIR-LU, ELIXIR-CZ, ELIXIR-SI, ELIXIR-NL

Helis Academy

- Bridge gap between industry needs and existing training programs and curricula
- Training program that
  - Improves industry-readiness of graduate students in the Health and Life Sciences
  - Upskill employees (life-long learning) in the Health and Life Sciences sector

Good Manufacturing Practices (GMP)
Clinical Testing
Data Analysis and Stewardship
Business Development
Technology & Life Sciences
Product and Process Design

Partners

- https://www.helisacademy.com/en
- Contact: Christine Staiger , Celia van Gelder (DTL) & Friederike Ehrhart, Lauren Dupuis (Maastricht University)

Helis Academy

Helis Academy Data Analysis and Data Stewardship Learning Blocks

Omics Data Analysis

Transcriptomics data analysis

Multi omics integration

FAIR Data Stewardship

Data management planning

Research life cycle

Statistics

Statistical thinking

Experimental design

Machine Learning

Visual analytics for genomics

Data mining

- Course FAIR Data Stewardship: 27-29 May 2019
- Course Omics Data Analysis: 11-14 June 2019
- Location: Niel, Belgium
- Courses can be followed together or separately
- Courses will be run in Belgium and Netherlands in year 2 and 3 of the project

Community Building –
The Data Stewards Interest Group

- June 2017 initial meeting of the Interest Group

- Leiden UMC, DTL, and UMC Utrecht founding members; mailing list has 90 members

  (NL, BE, LU, UK, SE, FR, DE)

- 9x meetings up until now;

  next meeting June 14, 2019 (remote attendance possible)

- Chair: Jasmin Böhmer, UMCU

- https://www.dtls.nl/community/interest-groups/data-stewards-interest-group/

Slide: Jasmin K. Böhmer

Mission & Activities of the Data Stewards Interest Group

Providing	Fostering	Joining
... platform for practical experience exchange	...the national implementation of Data Stewards(hip)	... efforts to produce hands-on solutions

The Input... experience, guidance, agreements, anecdotes, definitions

The Output ... use cases, documentation, case studies

So far:
- Collection of 24 stories and anecdotes for the promotion of good data stewardship
- Collaborative Paper on FAIR readiness for organizations
- Started inventory of training materials and courses

Slide: Jasmin K. Böhmer

Making Training FAIR – ELIXIR FAIR Training Working Group

- chair: Mateusz Kuzak (ELIXIR-NL)

- Task force 1: **Training about FAIR data**
  - making training resources findable via TeSS, the ELIXIR Training Portal
  - Defining necessary skills for implementing FAIR data
  - developing a set of structured terms to classify the skills necessary to make data FAIR and keep it FAIR (with GO-TRAIN)
  - lead: Mateusz Kuzak, Fotis Psomopoulos
- Task force 2: **Making training resources FAIR**
  - developing and implementing FAIR metrics for training resources ((in collaboration with Michel Dumontier and the FAIR metrics group)
  - lead: Leyla Garcia, Bérénice Batut

Tool: The ELIXIR Data Stewardship Wizard

- The Data Stewardship Wizard (DSW) is a web-based application for preparing smart Data Management Plans.

- It connects to additional learning resources and helps raise awareness about DS challenges.

- *https://ds-wizard.org/*

The Wizard is for:

- **Researchers:** Support in DS Planning, Raise awareness

- **Data Stewards:** Check list DS Planning

- **Experts:** Their contact details can be linked

- **Funders / Institutes:** Customize, Domain specific

- Allows tailor made DSW instances for an organisation, a Node, a company etc.

- Contact: rob.hooft@dtls

Summary

- **FAIR data treatment** and **training of researchers and data experts** are core themes of DTL and ELIXIR-NL

- A significant part of our training activities can be categorized as **FAIR Datastewardship Training & Capacity Building:**
  - Defining the Data Steward profession in the life sciences
  - Curriculum Building
  - Training courses
  - Community building
  - Making training FAIR
  - Tool: ELIXIR Data Stewardship Wizard

- Please join our activities if you are interested!
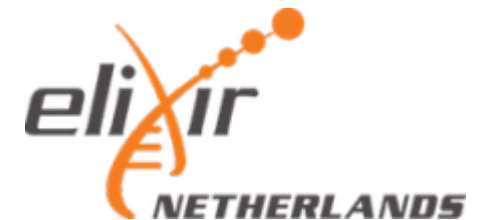
Acknowledgements

- UMCG: Salome Scholtens, Marije van der Geest
- UMCU: Nelly Anbeek, Jasmin Böhmer
- Radboudumc: Mirjam Brullemans
- Radboud University: Mijke Jetten, Inge Slouwerhof
- Maastricht University: Friederike Ehrhart, Lauren Dupuis, Chris Evelo
- ELIXIR-Luxembourg: Roland Krause, Dietlind Gerloff
- ELIXIR-Belgium/VIB: Paula Andrea Martinez, Alexander Botzki
- Data Stewardship Wizard: Rob Hooft (DTL), Robert Pergl, Marek Suchánek (ELIXIR-CZ)
- DTL Data Stewards Interest Group
- Our funders: ELIXIR-EXCELERATE & ZonMw

Thank you!