

Celia van Gelder (ELIXIR-NL), Patricia Palagi (ELIXIR-CH), Gabriella Rustici (ELIXIR-UK), Pascal Kahlem (ELIXIR-Hub) & on behalf of the ELIXIR TrCG group



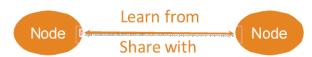


ELIXIR-CZ Annual Conference, 21-22 November 2018

www.elixir-europe.org

# **ELIXIR Training Platform**

- Build a sustainable training infrastructure supported and adopted by all member states
- Deliver a coherent, high-quality and impactful ELIXIR training program
  - addressing training gaps
  - growing bioinformatics capacity and competence in Europe
  - empowering researchers across Europe to use the ELIXIR services and tools
- Builds on and complements trainings in ELIXIR Nodes





























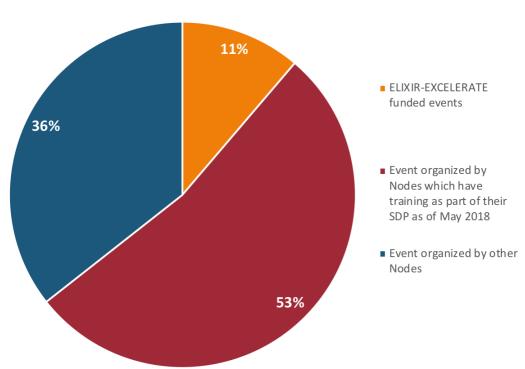


# ELIXIR Training infrastructure - in summary

- Training events for researchers, developers, trainers and infrastructure operators
- Large pool of ELIXIR trainers (Train-the-Trainer) and developed TtT framework
- TeSS portal to share and exchange training content and events <a href="http://tess.elixir-uk.org">http://tess.elixir-uk.org</a>
- Training best practices and guidelines, including standardized metrics for impact assessment - Training Toolkit
- Solutions to ease training and communication: E-learning platform and Virtual Coffee Room (VCR)



# Summary statistics for the ELIXIR training programme 2015-2018

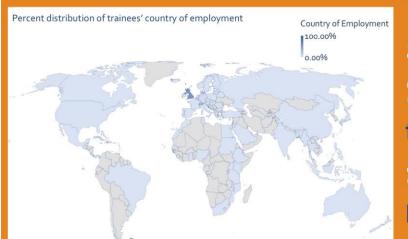


IR-EXCELERATE led events

Int organized by es which have using as part of their as of May 2018

Int organized by other es

"I would recommend the course to others" ~ 89% of survey respondents



"I am now able to use the tools/resources independently or with use of the training materials"

~ 88%\* of survey respondents

570\* training events
9680\* people trained
from 65\* countries

19\* ELIXIR Nodes
providing training

"The training improved my ability to handle data, work more quickly, or communicate with a bioinformatician"

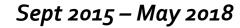
~ 88%\* of survey respondents

### **ELIXIR Training: Quality and Impact**

https://www.elixir-europe.org/platforms/training







N.B. For the node trainings this is an underestimate since not all nodes have provided complete information yet



## Train-the-Researcher (TtR) & Train-the-Developer (TtD)

Since 2015, we run ELIXIR-EXCELERATE-funded TtR and TtD events, covering:

- Workflow tools (Galaxy, Chipster), High Performance Computing, Python, Unix/Linux, AAI Workshop
- Use Case community-driven training: Variant analysis, Genome annotation & assembly, Metagenomics
- e-learning courses ran using ELIXIR-SI EeLP (https://elixir.mf.uni-lj.si), including Unix, Genome Assembly and Annotation, RNA-Seq data analysis using Chipster and HPC.



Train-the-Trainer (TtT) - Building a highly-skilled and coherent community

of trainers

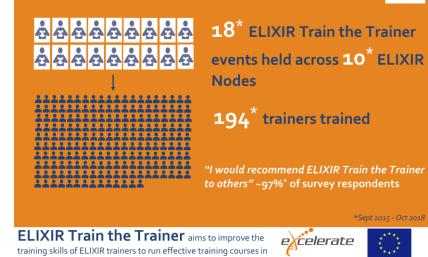
Created an ELIXIR TtT programme with a set of materials and course layouts

 Created a TtT Exchange program – support trainers and Nodes

OPINION ARTICLE

The ELIXIR-EXCELERATE Train-the-Trainer pilot programme: empower researchers to deliver highquality training [version 1; referees: awaiting peer review]

Sarah L Morgan<sup>1</sup>, Patricia M Palagi (b) 2\*, Pedro L Fernandes<sup>3\*</sup>, Eija Koperlainen<sup>4</sup>, Jure Dimec<sup>5</sup>, Diana Marek<sup>2</sup>, Lee Larcombe<sup>6</sup>, Gabriella Rustici<sup>7</sup>, Teresa K Attwood<sup>8</sup>, ≥



bioinformatics and related fields



A new pan-European Train-the-Trainer programme for bioinformatics: pilot results on feasibility, utility and sustainability of learning 3

Allegra Via ™, Teresa K. Attwood, Pedro L. Fernandes, Sarah L. Morgan, Maria Victoria Schneider, Patricia M. Palagi, Gabriella Rustici, Rochelle E. Tractenberg 🗷



### TtT Course setup

### The TtT courses are organized around four sessions:

- Learning principles and how they apply to training
- Training techniques that can be used to enhance learner engagement and participation
- Assessment and feedback in training
- Session, course, and materials design

### The Learning Objectives are:

- To get acquainted with learning principles and how they apply to training
- To be able to select and use training techniques that can enhance learner engagement and participation
- To learn how to use assessment and feedback in training
- To learn how to design a session, a course, and course materials



# Teaming up with The Carpentries

 Pilot Software and Data Carpentry in 2015/16: 278 life sciences researchers and 63 instructors trained OPINION ARTICLE

Developing a strategy for computational lab skills training through Software and Data Carpentry: Experiences from the ELIXIR Pilot action [version referees: 3 approved]

Aleksandra Pawlik<sup>1</sup>, Celia W.G. van Gelder (ip <sup>2</sup>, Aleksandra Nenadic (ip <sup>3</sup>, Patricia M. Palagi<sup>4</sup>, Eija Korpelainen<sup>5</sup>, Philip Lijnzaad<sup>6</sup>, Diana Marek<sup>4</sup>, Susanna-Assunta Sansone<sup>7</sup>, John Hancock<sup>8</sup>, Carole Goble<sup>3</sup>

- Since 2017:
  - > 40 ELIXIR courses have been delivered, many more delivered by the ELIXIR Nodes themselves
  - ELIXIR/Carpentries agreement to support the rolling out across Nodes:
    - 20 Software Carpentry and Data Carpentry workshops
    - 2 Instructor trainings





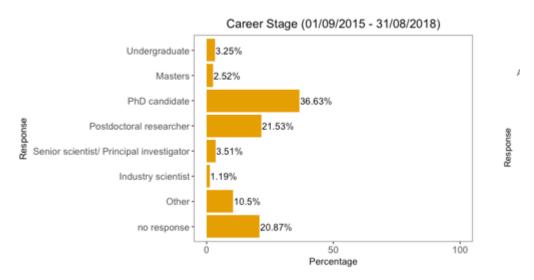
## ELIXIR Training Quality & Impact assessment strategy

- Minimal set of descriptors for training events and materials
- Set of core metrics and feedback surveys, for both short and long term
  - Adopted by 15 ELIXIR Nodes
- Platform for data collection is being set up: First version ready for testing by TrCG, expected roll out early 2019
- Tailored for and adopted by EOSC Life and European Joint Project on Rare Diseases Projects



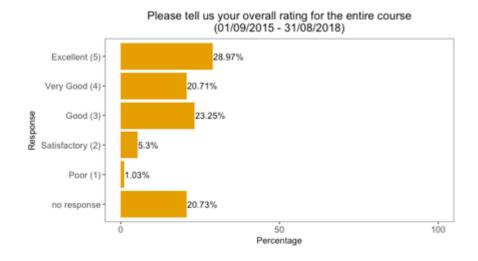
# Quality & Impact assessment – Short term feedback surveys

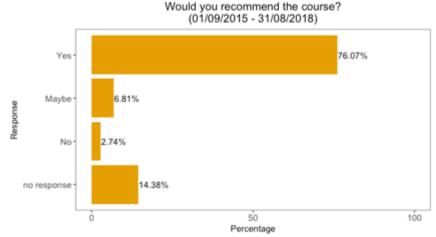
#### Career Stage



In the time period September 2015-August 2018, **short-term feedback** data has been collected for **307** training events run across **18** Nodes with **5040** short term survey responses collected

Data provided by Kim Gurwitz, ELIXIR Training Platform Quality & Impact coordinator



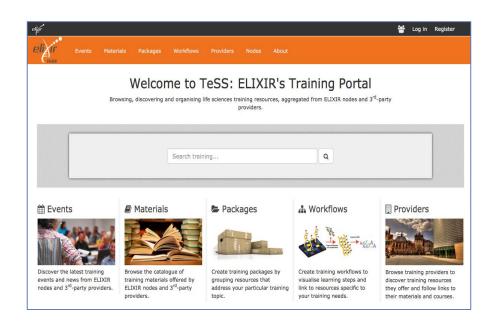




## **ELIXIR Training Portal TeSS**

Aggregate, disseminate and coordinate information on ELIXIR's training activities/material and beyond

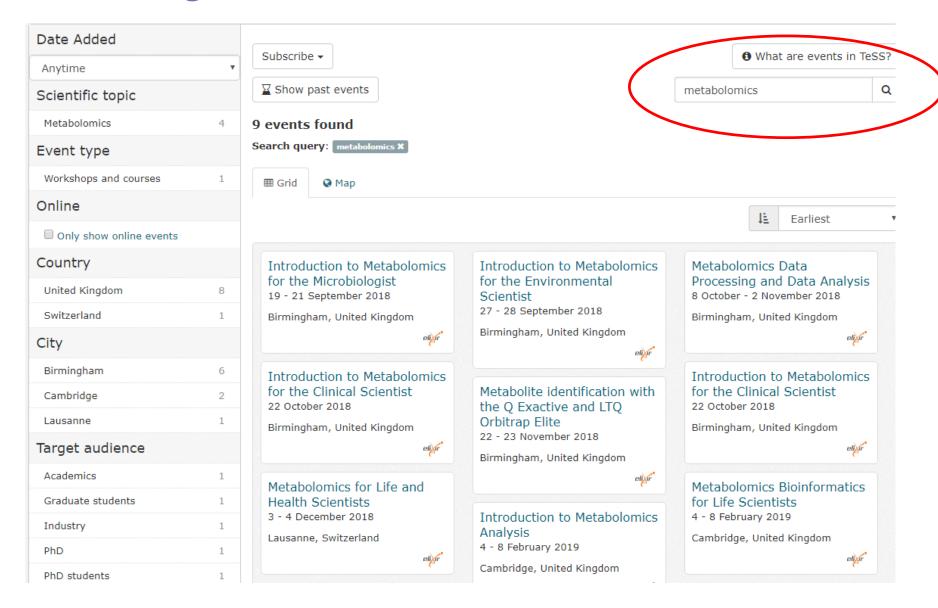
- 1023 training materials
- 48 content providers: 21 scraped automatically, 13 ELIXIR Nodes
- 4,934 unique users (2017-18)
- 308 upcoming events listed
- AAI and Bioschemas integration
- Linked to other ELIXIR registries: bio.tools and
- Events annotated with EDAM topics terms
- https://tess.elixir-europe.org



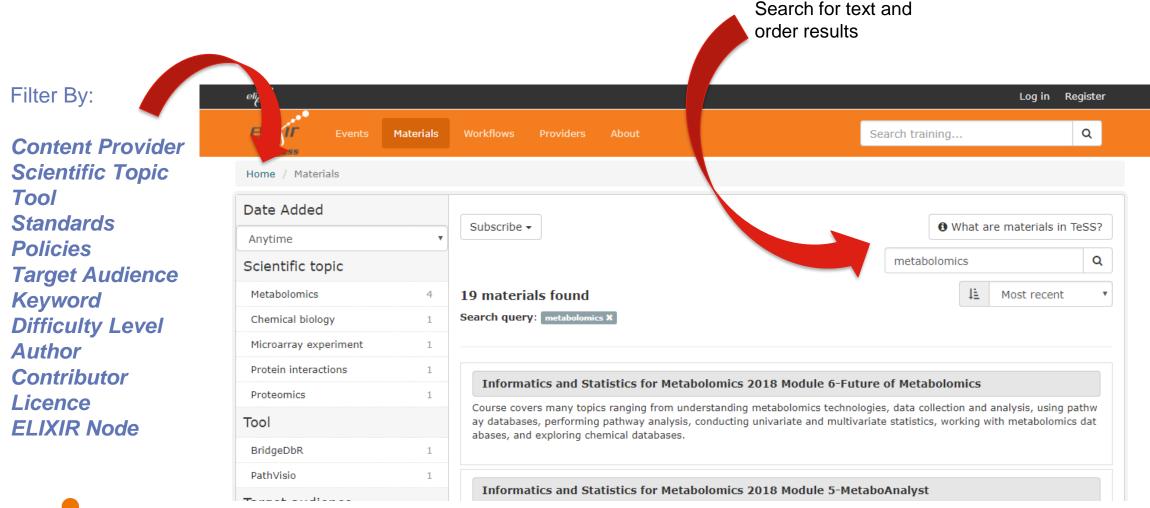
**FAIR**sharing



# TeSS - Training Events



### TeSS - Materials





# Training as a service for Nodes - Training Toolkit

Collection of resources - developing, delivering, and evaluating training

### How to:

- design and implement training courses for different audiences & domains
- make training Findable, Interoperable using the Bioschemas specifications
- design learning paths with competency frameworks & 'Knowledge, Skills and Abilities'
- collect feedback and determine the quality and impact of training activities
- build training capacity through the Train-the-Trainer program
- The training Toolkit is available for all training organizers & coordinators in the Nodes (and beyond)



# Ongoing projects

Integration of ELIXIR Portals from users perspective (TeSS and **\(\beta\)**)



Workshop as a Service, Using clouds and VMs in bioinformatics training



- Developing learning paths for users of ELIXIR services
- Towards Data Stewardship in ELIXIR: Training & Portal
- Training related to ELIXIR services/resources: BioSchemas, Beacons, CWL and biocontainers



- Mapping the landscape of Biocuration in ELIXIR: Practice, capability and training requirements
- 2019: Implementation studies with Galaxy, Rare Diseases, Metabolomics Communities
- 2019: Training contributions in EOSCLife & in EJPRD

# **ELIXIR Industry Strategy: Training**

Ensure visibility and discoverability of ELIXIR training courses to industry, for example:

- Training events communicated through Industry mailing list, ELIXIR Events Portal and TeSS
- Ensure industry usage of TeSS for courses and materials (https://tess.elixir-uk.org)
- Work towards defining an ELIXIR certified training 'Quality stamp'

Perform gap analysis for training needs in industry:

- Survey industry partners in ELIXIR Member countries, using existing 'ELIXIR UK survey' as a basis and expanding to other countries
- Define the training needs (including scope, content, target audience(s) and format) in industry (e.g. using the Training platform survey) and use that to help guide ELIXIR's Training strategy

Improve the communication between training-related experts in industry and ELIXIR Training Platform:

- Communicate scope and vision of ELIXIR's training programme to industry
- Build a liaison system for effective communication including, for example, an e-mail list for industry researchers interested in training

Findability & Visibility:

- TeSS
- ELIXIR Industry Newsletter

Industry training needs Surveys; Learning Path Implementation Study

Liaison Training Platform & Industry representatives



# **ELIXIR & FAIR & Training**















 Workshop - ELIXIR All-Hands meeting, June 2018



























# ELIXIR FAIR Training – Working group

- Subtask: FAIR Training events & materials Leyla Garcia (ELIXIR-UK) and Berenice Batut (ELIXIR-DE)
  - What FAIR principles mean in this case
  - How to assess FAIRness
  - Metrics to assess FAIRness; collaborating with FAIRmetrics team (Michel Dumontier, <u>www.fairmetrics.org</u>)
  - Recent session at BioHackathon Paris Nov 2018:
    - (i) evaluating FAIR metrics in the context of training
    - (ii) evaluating training materials with the metrics
    - (iii) finalised bioschemas descriptions for training



# ELIXIR FAIR Training – Working group

 Subtask: Training on FAIR principles – Mateusz Kuzak (ELIXIR-NL) & Fotis Psomopoulos (ELIXIR-GR)

- Detail current materials and curricula and how to complement it
- Define what is unique to ELIXIR
- Avoid reinventing the wheel collaborate
- How to teach FAIR principles
- Ensure all is Findable in TeSS



# Training in FAIR principles – to whom

### Skills and competencies for:

- Researchers: create, manage and share FAIR data
- Trainers: teach and apply FAIR principles
- Data Stewards: provide support for research teams, data repositories, data services
- Reviewers: assess FAIRification of datasets, data resources in publications, grant proposal and reports



# Training in FAIR principles – some examples

BYOD: FAIRify datasets, combine FAIR datasets





Carpentry style FAIR training – teach recognized
 FAIRification processes – GO Train Implementation network





Defining competencies and building training - Data Steward in Life Sciences

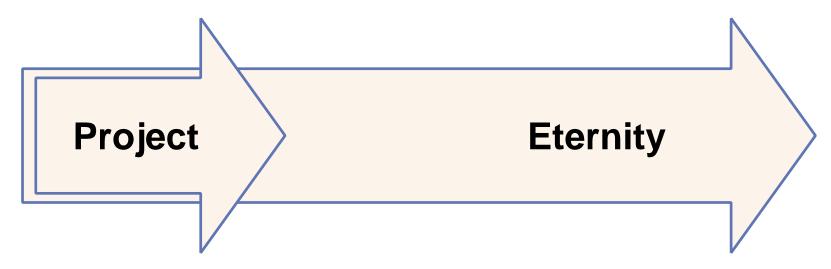




### ELIXIR & Data Stewardship (DS)



**Data Stewardship** is the 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









## DM/DS skills are needed but there are challenges

- Skills in data management (DM) & data stewardship (DS) are crucial to deal properly with the data
- From EOSC High Level Expert Group, Oct 2016:
  - "There is an alarming shortage of data experts both globally and in the European Union"
  - "Core data experts need to be trained and their career perspective significantly improved"

### Challenges:

- Who to train? Which target audiences?
- What are the skills needed for each specific target audience?
  - For researchers: Crucial for research but not everyday practice (yet) and very demanding
  - For Data Stewards: In high demand, Upskilling needed, Skills profile currently being defined



# Towards Data Stewardship in ELIXIR: Training and Portal

### Aim of the Implementation Study

- Bundle and effectively expose existing expertise on DM/DS in ELIXIR
- Formulate, build and implement first solutions for good DM/DS in ELIXIR

### Tasks in the project

- Task 1: ELIXIR Data Management and Data Stewardship (DM/DS) Training Needs and Activities
- Task 2: A Data Stewardship Portal
- In line with ELIXIR strategic objective: "Develop a plan that connect the national data management programmes, advanced user support and training in ELIXIR Nodes to ELIXIR's federated data ecosystem together with monitoring of jointly agreed indicators of use and impact"
- ELIXIR Nodes involved: NL, CZ, LU, SI (and many others in kind)
  - In connection with ELIXIR DMP Working Group



# Task 1: ELIXIR Data Management and Data Stewardship (DM/DS) Training Needs and Activities

- Content of task:
  - Inventory of the ELIXIR DM/DS training landscape
  - Scoping, training needs and gap analysis
  - Development of ELIXIR DM/DS training
  - Run ELIXIR DM/DS courses

- We started with looking at competences needed to acquire
  - Organised around research data life cycle
  - Build on existing work (inside and outside Life Sciences & inside and outside ELIXIR)
  - After this: build modular and reusable training material and run courses

### Data Steward Competencies - Generic vs Life Sciences (LS) Domain

- Generic
  - EU:
    - EDISON (<a href="http://edison-project.eu/">http://edison-project.eu/</a>)
    - EOSCPilot WP7 (<a href="https://eoscpilot.eu/media/deliverables">https://eoscpilot.eu/media/deliverables</a>)
  - NL: LCRDM
    - Working alongside the LS domain work (ZonMw project)







- Life sciences domain
  - ELIXIR:
    - Implementation study DM/DS: ELIXIR-NL, -CZ,- LU,- SI
    - Implementation study Learning Paths: ELIXIR- BE is leading the DM/DS use case
  - NL: ZonMw funded project "Towards FAIR Data Steward as profession for the Life Sciences"









### 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
- Funded by ZonMw plus inkind contribution of the project partners

# Radboudumc Radboud University

















### **Towards FAIR Data Steward as profession for the Life Sciences**

#### Deliverable 1:

- Job profile/Role description for two levels of Data Stewards;
  - Institute /organization level
  - Project level
- endorsed by all stakeholders in NL
- ultimate goal: establishing function profiles in the NFU HR framework

### **Activities:**

- We analysed job vacancies and mapped with various existing frameworks (EDISON, DAMA, Purdue, EOSCpilot)
- Version 1 of the role descriptions have been reviewed by stakeholders in Nov 2018.



### **Towards FAIR Data Steward as profession for the Life Sciences**

Next steps:

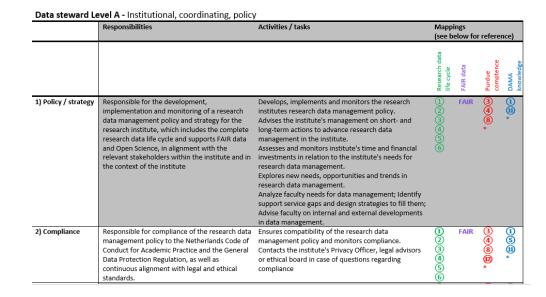
Next version of descriptions will be available for external use and review

January 2019

 Attaching knowledge base and skills to these role descriptions for the two DS roles

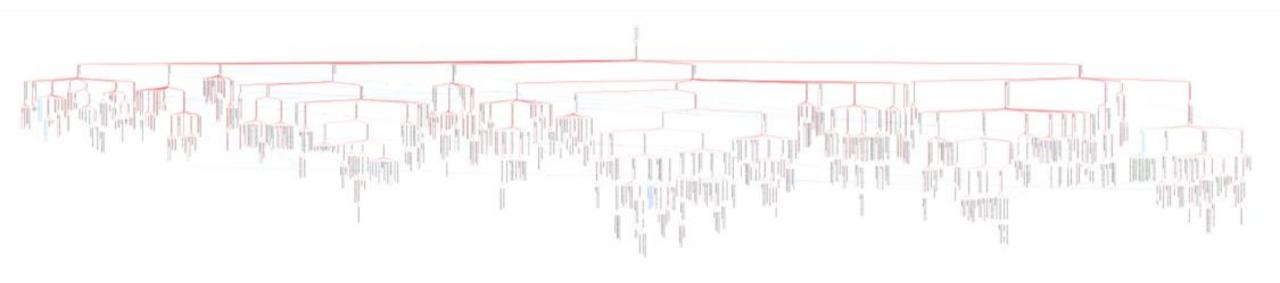
Design dedicated trainings & roll out
 2 f2f trainings

- Make script for elearning trajectory
- Align and join forces with ELXIR Implementation Study activities





# Data Stewardship Mind Map (Rob Hooft, DTL)



Mind Map is capturing the questions of all experts involved in DS; it is formalized into a **knowledge model** and is implemented in the DSW



During measurement phase

Analyze chance and effect size

In Analysis

Do people store data on computers in the

Who has access to the labs?

lab?

Who has access to those computers?

Do people have data on external hard drives?

Is there any data on USB sticks people carry?

Encrypted?

Is the Risk of information loss / theft acceptable?

Are they encrypted?

Is any data on laptops?

Is the laptop password protected?

Are researchers using cloud accounts?

Consider encryption

Are reports sent over e-mail?

Consider encryption

Do people share accounts?

Do the data centers where data is stored have Certifications?

Security for hosting

Are all internal web services used via https?

Did you run hilote to see what the canacity

# The ELIXIR Data Stewardship Wizard (DSW)

# "Smart Data Management for FAIR Open Science For serious researchers and data stewards"

### The Wizard is for:

- Researchers:
  - Support in DS Planning
  - Raise awareness
- Data Stewards:
  - Checklist DS Planning
- Experts:
  - Their contact details can be linked;
  - Allows tailor made DSW instances for an organisation, a Node etc.
- Funders / Institutes:
  - Customize, Domain specific



https://dsw.fairdata.solutions



# The ELIXIR Data Stewardship Wizard (DSW)

- The knowledge model is used by the DSW to display a questionnaire
- It opens relevant questions (only)
- It contains more detailed guidance on some topics
- It contains links to Barend Mons book pages(Data Stewardship for Discovery)
- It measures the FAIRness of the result per chapter

### Ongoing activities:

- It will contain links to course information / TeSS / EeLP / FAIRsharing/ bio.tools (hackathon Dec 2018)
- Collaborations
  - with funders (ZonMw) they want to easily and objectively quantify the FAIR metrics of DMPs.
  - with company DSM wan to run their local copy of DSW

# https://dsw.fairdata.solutions

### **DSW Team:**

- Robert Pergl, Marek Suchánek,
   Jan Slifka, Vojtěch Knaisl (ELIXIR-CZ)
- Rob Hooft, Mateusz Kuzak (ELIXIR-NL)





Today there will be a training wokshop about the DSW.
Please try the DSW and give us your feedback!!







#### Design of experiment

Data design and planning

Data Capture/Measurement

Data processing and curation

Data integration

Data interpretation

Information and insight

### Design of experiment

Before you decide to embark on any new study, it is nowadays good practice to consider all options to keep the data generation part of your study as limited as possible. It is not because we can generate massive amounts of data that we always need to do so. Creating data with public money is bringing with it the responsibility to treat those data well and (if potentially useful) make them available for re-use by others.

#### Is there any pre-existing data?

Are there any data sets available in the world that are relevant to your planned research?

- No
- Yes

#### Will you be using any pre-existing data (including other people's data)?

Will you be referring to any earlier measured data, reference data, or data that should be mined from existing literature? Your own data as well as data from others?

- No
- Yes \equiv \equiv

# D

#### What data formats/types will you be using?

Have you identified types of data that you will use that are used by others too? Some types of data (e.g. genetic variants in the life sciences) are used by many different projects. For such data, often common standards exist that help to make these data reusable. Are you using such common data formats?

0

○ No

Yes

Item	
Is this a standard data format used by others too?	
○ No	
○ Yes	
Does this data format enable sharing and long term archiving?	
Complicated (binary) file formats tend to change over time, and software may not stay compatible with older versions. Also, some formats hamper long term usability by making use of patents or being hampered by restrictive licensing	
○ No <b>=</b>	
○ Yes	

# Complicated (binary) file formats tend to change over time, and software may not stay compatible with older versions. Also, some formats hamper long term usability by making use of patents or being hampered by restrictive licensing No Yes

Will you convert to a file format more suitable for archiving later?

You may need to reserve time and budget for this

Does this data format enable sharing and long term archiving?

# DSW related Trainings & Workshops - Past

- **ELIXIR AHM2017** ELIXIR-CHARME Workshop 'FAIR Data and Data Stewardship in ELIXIR: How to write your own FAIRy tale (by Interoperability & Training Platform)
- **EBI webinar** Rob Hooft July 2018 "Tools for data management planning" (in EBI series): <a href="https://www.ebi.ac.uk/training/events/2018/bringing-data-life-data-management-biomolecular-sciences">https://www.ebi.ac.uk/training/events/2018/bringing-data-life-data-management-biomolecular-sciences</a>
- Workshop/hands-on **ECCB 2018 Athens** "Implementation of Data Management Plans & Data Stewardship in practice"; Materials: <a href="https://elixir.mf.uni-lj.si/course/view.php?id=40">https://elixir.mf.uni-lj.si/course/view.php?id=40</a>
- NETTAB 2018 Genova "The ELIXIR Data Stewardship Wizard how to guide life science researchers towards FAIR data management"; workshop/hackaton on DSW contents



# DSW & DS/DM Trainings & Workshops - Upcoming

- ELIXIR-CZ Annual Conference , Nov 2018
  - Training DSW by Robert Pergl (ELIXIR-CZ)
- DSW: specialised workshops:
  - Workshops/hackatons for power users and contributors;
    - Connecting TeSS, FAIRsharing and EeLP to DSW. December 2018 in CZ.
  - Workshop for service providers; for those who want to deploy their local instance of Wizard e.g. for their own organization or institute; In 2019
- DS/DM Training workshops
  - 3 workshops; in 2019
  - Aligning with ZonMw trainings



# Final thoughts

- ELIXIR Training Programme is shaped and implemented by and for the ELIXIR Nodes
- It is a community effort, we have to build and sustain the training portfolio together by combining expertise and manpower
  - Join ELIXIR Training projects, implementation studies, working groups!
- Nodes can benefit from the ELIXIR Training Infrastructure and the Training Toolkit
  - Please add your trainings to TeSS. This can be automated!
  - Make sure for the trainings you do to capture the feedback through the ELIXIR surveys
- Contact the Training Platform for more information: <u>training-exco@elixir-europe.org</u>



# Acknowledgements



#### ELIXIR Training Coordinator Group (TrCG)

In the picture: Pedro Fernandes (PT), Eva Alloza (ES), Roland Krause (LU), Alexander Botzki (BE), Daniel Wibberg (DE), Brane Leskosek (SI), Allegra Via (IT), Jure Dimec (SI), Loredana Le Pera (IT), Bálint L. Bálint (HU), Dietlind Gerloff (LU), Niall Beard (UK), Mateusz Kusak (NL), Eija Korpelainen (FI), Pascal Khalem (Hub), Patricia Palagi (CH), Jessica Lindvall (SE), Kim Gurwith (UK), Celia van Gelder (NL), Hedi Peterson (EE), Gabry Rustici (UK), Victoria Dominguez del Angel (FR)



#### **ELIXIR Training Executive Committee**

Patricia Palagi, Celia van Gelder, Gabriella Rustici, Pascal Kahlem

#### ZonMw project Towards FAIR Data Steward as a profession in the life sciences

Salome Scholtens & Marije van der Geest (UMCG), Nelly Anbeek (UMCU), Mirjam Brullemans (Radboudumc), Mijke Jetten (Radboud University), Celia van Gelder (DTL), DTL Data Stewards Interest Group

#### ELIXIR Implementation Study Towards Data Stewardship in ELIXIR

Roland Krause & Dietlind Gerloff (LU), Robert Pergl, Marek Suchánek, Jan Slifka and Vojtěch Knaisl (CZ), Brane Leskosek & Marko Videk (SI), Paula Andrea Martinez (BE), Mateusz Kuzak & Celia van Gelder (ELIXIR-NL)

#### **ELIXIR FAIR Training Working Group**

Mateusz Kuzak, Leyla Garcia, Pete McQuilton, Kim Gurwitz, Melissa Burke, Sarah Morgan, Bérénice Batut, Ricardo Arcila, Fotis Psomopoulos, Victoria Dominguez Del Angel, Patricia Palagi, Celia van Gelder, Gabriella Rustici



# Thank you!







