Challenges in adopting FAIR principles in Training

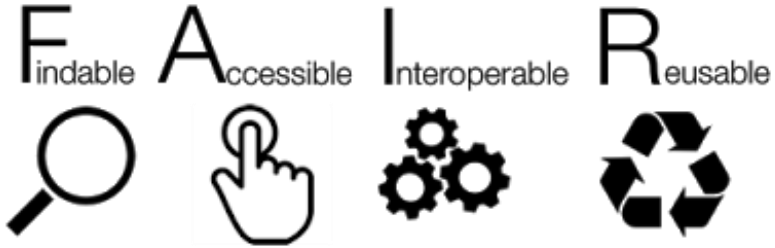
Patricia Palagi
SIB Swiss Institute of Bioinformatics / ELIXIR-CH
Celia van Gelder (ELIXIR-NL), Pascal Kahlem (ELIXIR-Hub) & Gabriella Rustici (ELIXIR-UK)

ELIXIR-EXCELERATE is funded by the European Commission within the Research Infrastructures programme of Horizon 2020, grant agreement number 676559.

www.elixir-europe.org

NETTAB, 22-24 October, 2018, Genova, Italy

Challenges in adopting



TRAINING

Part of the challenge

Part of the success

Content

1. The ELIXIR Training Platform
   - Facts and figures
   - Going FAIR
2. Training in FAIR principles
3. Making training materials FAIR

Content

1. The ELIXIR Training Platform:
   - Facts and figures
   - Going FAIR
2. Training in FAIR principles
3. Making training materials FAIR

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 - in summary

- **Training events** - researchers, developers, infrastructure operators & trainers
- **Train-the-Trainer** - programme & developed framework
- **TeSS** portal - share & exchange training content and events
  https://tess.elixir-europe.org
- **Training Toolkit** - Training best practices, guidelines, standardized metrics for quality & impact assessment
- Solutions to ease training and communication - **E-learning platform** &**VCR**

19 Nodes involved: BE, CH, CZ, DE, EBI, EE, ES, FI, FR, GR, IL, IT, LU, NL, NO, PT, SE, SI, UK

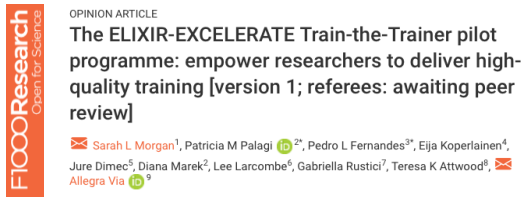
Training for Researchers, Developers, Trainers

Teaming up with The Carpentries: Software & Data

- Pilot in 2015/16 - 278 life sciences researchers and 63 instructors trained
- ELIXIR/Carpentries **agreement** - support the rolling out across Nodes
    - 20 workshops + 2 Instructor trainings

ELIXIR TtT programme – courses, set of materials, course layouts, exchange

- 15 TtT events across 8 Nodes, trained 164 trainers from 12 Nodes

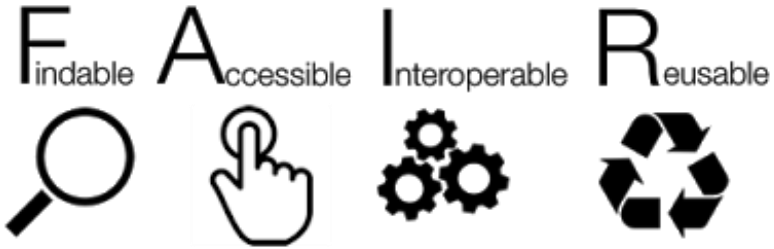


Train-the-Trainer (TtT)

Diffuse

Raise awareness

Part of the best practices



Summary statistics - ELIXIR training programme
Sep *2015* – *May 2018*

Total number of events = 570

Attendance: 9'680 registrations

Trainers: 1'214

Days of training: 1'431

Standardization, Quality & Impact assessment

- Minimal set of descriptors - training events and materials
  - contributed to shape relevant specifications

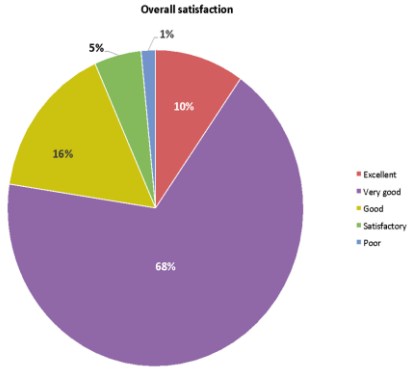
Bioschemas

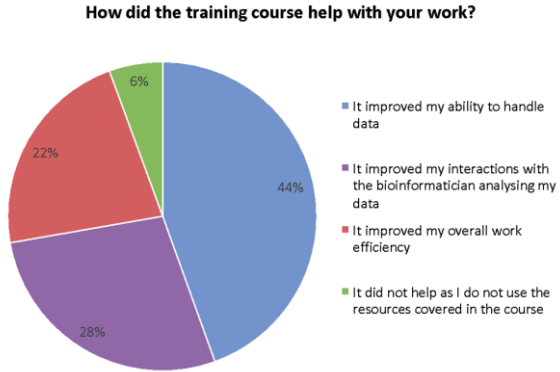
GOBLET
Global Organisation for Bioinformatics Learning, Education & Training

- Set of core metrics and feedback surveys
  - Short and long term
  - Reference in Europe: EOSC Life, EJP Rare Diseases

Findable Accessible Interoperable Reusable

- 333 training events, 15 Nodes, 5310 responses





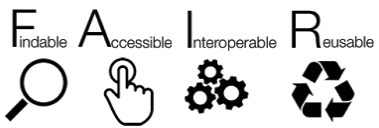
TeSS Training Portal

https://tess.elixir-europe.org

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

- 1023 training materials, 4,934 unique users, 308 upcoming events
- 48 content providers: 21 scraped automatically, 13 ELIXIR Nodes

- Integration AAI and Bioschemas
- Linked to other ELIXIR registries: bio.tools and FAIRsharing
- Events annotated with EDAM topics terms



elixir TeSS elixir

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

FAIR principles

It’s a community effort

ELIXIR FAIR Training – 1st steps

- Workshop - ELIXIR All-Hands meeting, June 2018
- Gathered over 30 participants

ELIXIR FAIR Training – Working group

- Task1: Training on FAIR principles – Mateusz Kuzak & Fotis Psomopoulos

    - 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

ELIXIR FAIR Training – Working group

- Task2: Training events & materials FAIR – Leyla Garcia & Bérénice Batut
  - What FAIR principles mean in this case
  - How to assess FAIRness
  - Metrics to assess FAIRness
  - FAIR challenges and how to overcome them

Content

1. The ELIXIR Training Platform:
   - Facts and figures
   - Going FAIR
2. Training in FAIR principles
3. Making training materials FAIR

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 for FAIR data

Towards Data Stewardship in ELIXIR: Training and Portal - ELIXIR Implementation Study

Data Management & Data Stewardship

- Inventory of the ELIXIR training landscape
- Scoping, training needs and gap analysis
- Development of ELIXIR training
- Run ELIXIR courses

GO FAIR
governance structures
operations
advocacy
International Support and Coordination Office (GFISCO)
training
technique
community coordinators
community coordinators
community coordinators/teams
IN
IN
IN
IN
IN
IN
IN
IN
IN
IN
IN
IN
IN
policy, DS plans, next gen metrics etc.
training, certification, standards etc.
core technology, data, tools, compute.
GO CHANGE
• Culture change
• Open science promotion
• Reward system
GO TRAIN
• Education/training
• Certification
GO BUILD
• Technical implementation
• FAIR data & services
• Technical infrastructure

Content

1. The ELIXIR Training Platform:
   - Facts and figures
   - Going FAIR
2. Training in FAIR principles
3. Making training materials FAIR

Training materials – some examples

1. Lectures: Slides, texts, webpages
2. Exercises
3. Datasets
4. Working environment
    - Software
    - Workflows
    - VM
    - Containers
    - Repositories
    - Standards
5. Etc.

Training materials

=

digital objects

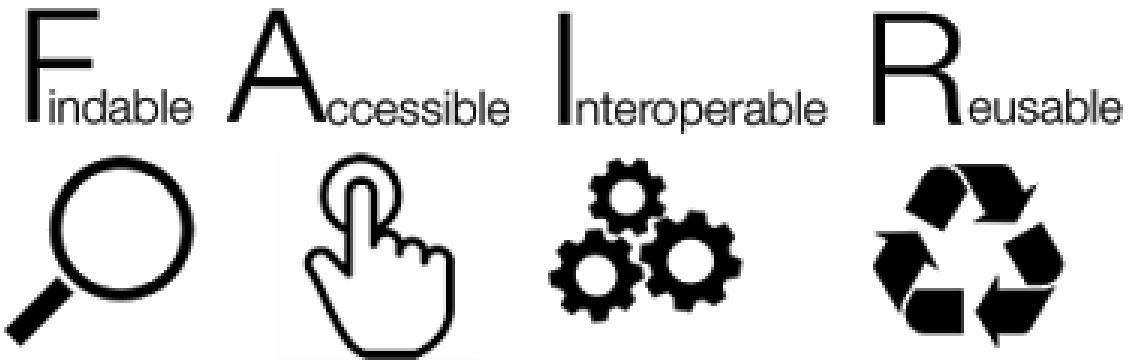
FAIR Training materials – to whom and why

- For Computers
- For Human beings – trainers and trainees
  - Find the materials
  - Reuse/replicate/remix training materials

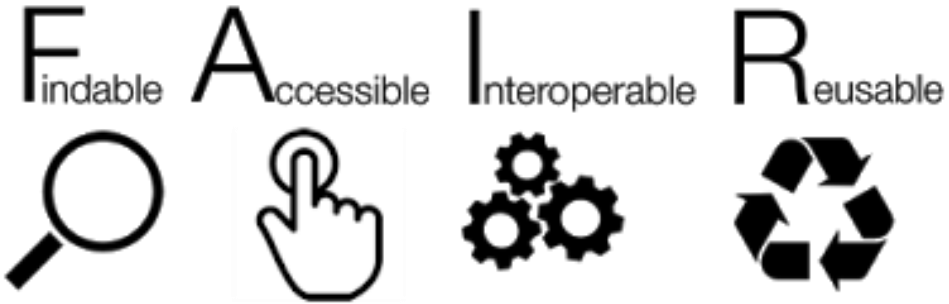
Training materials – some examples

1. Lectures: Slides, texts, webpages
2. Exercises
3. Datasets
4. Working environment
    - Software
    - Workflows
    - VM
    - Containers
    - Repositories
    - Standards
5. Etc.

Question for you



To facilitate adoption



We need to have a common understanding

Meaning - FAIR principles - Training materials

Findable

- Most proper identifier for each object with versioning
  - **Zenodo, Figshare** DOI - documents, slides
  - **FAIRsharing** DOI – resources
  - **Datacite** - datasets
- Data described with rich metadata
  - **Bioschemas** and **schema.org**
  - **EDAM**
- Findable in searchable resources
  - **TeSS, Datacite, FAIRsharing, BioContainers**

Meaning - FAIR principles - Training materials

Accessible

- Retrievable by ID using standard
  - Accessible via http/https or ftp
  - Accessible vs Open: may require a login, sign up for an account
- Metadata available even when data are no longer available
  - Sustainability of metadata
- Language barrier (metadata)
- Accessible for disabled people (metadata)
- Jargons

Meaning - FAIR principles - Training materials

Interoperability

- Use vocabularies that follow FAIR principles
  - Community standards for data format, following widely adopted guidelines
  - Widely-implemented ontologies/vocabularies to describe subject of courses – **EDAM**
- Include references to other (meta)data
  - Integrating all materials: datasets, Vms, environment to run the course
  - Registering metadata in an ID repository (e.g. Datacite)
- Very useful to build learning paths: matching requirements with learning outcomes

Meaning - FAIR principles - Training materials

Reusability

- Released with clear and accessible data usage licensing
  - Should be compatible with Reuse
- Detailed provenance - Attribution
- Meet domain-relevant community standards
  - Annotation of the pedagogical scenarios
  - Editable: text better than slides/pdf – slides more common
  - Good documentation on how to setup teaching infrastructure, computing resources, etc
  - Adaptability vs reusability (versioning)

Data Life Cycle – Training materials

Good stewardship of training materials

- Make sure are reusable
  - Once obsolete (format, tools) should be archived
  - Plan for how long they should be 'kept' FAIR
- Plans for sustainability - persistent
  - meta(data)
  - Identifiers
  - Repositories
- Question: add a new dataset or new slide, will this be a new course (DOI) or a new version of the same course

THE CARPENTRIES

We teach foundational coding and data science skills to researchers worldwide.

What we do

The Carpentries teach foundational coding, and data science skills to researchers worldwide. Software Carpentry and Data Carpentry workshops are based on our lessons. Workshop hosts, Instructors, and learners must be prepared to follow our **Code of Conduct**.

More ›

Who we are

We are a diverse, global community of **volunteers**. Our community includes **Instructors**, helpers, **Trainers**, **Maintainers**, **Mentors**, community champions, **member organisations**, supporters, workshop organisers, **staff** and a whole lot **more**.

More ›

Get involved

See all the **ways you can engage** with the Carpentries. Get information about upcoming events such as workshops, meetups, and discussions from our **community calendar**, or from our twice-monthly **newsletter**, *Carpentry Clippings*. Follow us on **Twitter**, **Facebook**, and **Slack**.

More ›

FAIR training – Use case - Carpentries

Orange= I could verify/assess FAIRness

FINDABLE

- Globally unique/persistent ID
- Standard rich-metadata
- Metadata Include ID of data described
- Registered/indexed in searchable resource

ACCESSIBLE

- Retrievable by ID using std protocol
- Protocol is open, free, universally implementable
- Metadata longevity

INTEROPERABLE

- Language for knowledge representation
- Use standards, ID, formats, guidelines, controlled vocabularies and ontologies
- Include references to other data

REUSABLE

- Clear and accessible data usage license
- Detailed provenance
- Meet community standards

FAIR training – Use case - TtT

ELIXIR-EXCELERATE Train the Trainer

The course is made up of four sessions, covering:

- 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

Learning Objectives:

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

FAIR training – Use case - TtT

Orange= I could verify/assess FAIRness

FINDABLE

- Globally unique/persistent ID
- Standard rich-metadata
- Metadata Include ID of data described
- Registered/indexed in searchable resource

ACCESSIBLE

- Retrievable by ID using std protocol
- Protocol is open, free, universally implementable
- Metadata longevity

INTEROPERABLE

- Language for knowledge representation
- Use standards, ID, formats, guidelines, controlled vocabularies and ontologies
- Include references to other data

REUSABLE

- Clear and accessible data usage license
- Detailed provenance
- Meet community standards

FAIR Training – future actions/events

- Paris BioHackathon – November 2018 - FAIRness assessment for training materials. Hackathon/Curatathon
  - TtT
  - Galaxy
  - 4OSS : github code, version control, documentation, metadata with EDAM ontologies – FAIR from the start

Final thoughts

- Becoming FAIR takes time, but it increases visibility of courses & research
- Once you start being FAIR, FAIR becomes natural
- Being part of the community which is defining FAIR training is enlightening , stimulating, rich
- The community will grow and encourage adoption of standards, workflows - accelerate adoption
- Training can support adoption and the communities
- Training - Investment in the future, we are training the future generations

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

Celia van Gelder, Gabriella Rustici, Pascal Kahlem

SIB Training Group

Geoffrey Fucile, Frédéric Schütz, Walid Gharib, Diana Marek and Grégoire Rossier

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, **Celia van Gelder,** Gabriella Rustici

Thank you!

ELIXIR-EXCELERATE is funded by the European Commission within the Research Infrastructures programme of Horizon 2020, grant agreement number 676559.

@ELIXIREurope

/company/elixir-europe