Workshop on Education in Bioinformatics 2021



Making training materials FAIR experiences, challenges, solutions

Flash talks from: Saskia Hiltemann, Anna Swan, Verena Ras, Maria Doyle, Melissa Burke, Patricia M. Palagi and Celia van Gelder

Flash talks

- Experiences from the Galaxy Training Network Saskia Hiltemann
- Experiences from the EMBL-European Bioinformatics Institute Anna Swan
- Experiences from the H3ABioNet Verena Ras
- Experiences from an Individual Trainer Maria Doyle
- Experiences from the Australian BioCommons Melissa Burke
- Experiences from the SIB Swiss Institute of Bioinformatics Patricia M. Palagi
- Experiences from the OpenAire CoP for Training Coordinators Celia van Gelder





Galaxy Training Network

Saskia Hiltemann





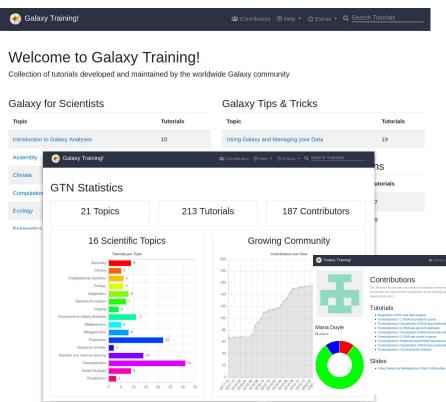




Galaxy Training Network (GTN): Who we are



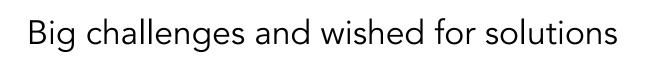
- Data Analysis Training using Galaxy platform
- Community-driven!
- Self-contained materials
 - For self-study by learners
 - For teaching by educators
- Everything Open (from the start)
 - Materials, data, tools, workflows, servers, ...
 - Development, discussions, feedback, usages stats, ...
- Regular Global Training events
- Accessibility for all
 - Alt text for all images, captions on all videos
 - Automated Slide Decks -> Video slides
 - WAVE Web accessibility tool



https://training.galaxyproject.org

How we implement FAIR for training materials

R1	Plan to share online	Webpage, no login required. (via GitHub pages)
R2	Properly describing	BioSchemas markup, EDAM
R3	Unique identity	Internal system for keeping URLs persistent, but no official DOIs
R4	Register online	Automatically scraped by TeSS
R5	Access Rules	No access restrictions, open licences
R6	Interoperable format	Markdown format (converted to HTML by Jekyll framework, PDF export options)
R7	(Re)usable for trainers	MIT/CC-BY licence, speaker notes, instructor tutorials, FAQs, feedback forms, community calls, Gitter, training infrastructure & dashboard.
R8	Usable for trainees	Metadata at top of every tutorial (objectives, time, level, resources, prerequisites,), tutorial search, FAQs, Gitter support
R9	Contribution friendly	GitHub, "Edit me" button, Tutorials for contributors, regular contribution fests, Gitter support, citable materials, acknowledgement of contributions.
R10	Keeping up-to-date	Automated testing, regular contribution fests, topic maintainers, open development on GitHub





Challenges

Balance between ease of contribution and feature-richness of framework

Sustainability; keeping materials up to date with a volunteer community

Wished-for solutions

Obtaining DOIs for training materials (automatable and updateable)

FAIR in practice: more practical, concrete solutions (guides, checklists, tools, badges, ..)



EMBL-EBI

Anna Swan

Who we are



EMBL-EBI Training

Delivering world-class training in data-driven life sciences.

Who our materials are for

Scientists, from all around the world, at all career stages who would like to make the most of biological data.

Scientists who would like to use the wealth of public biodata available, whilst also learning to be more confident users and guardians of their own data.



Live training 🕣

View opportunities to join us, either virtually or in person, to learn with other scientists from around the world



On-demand training

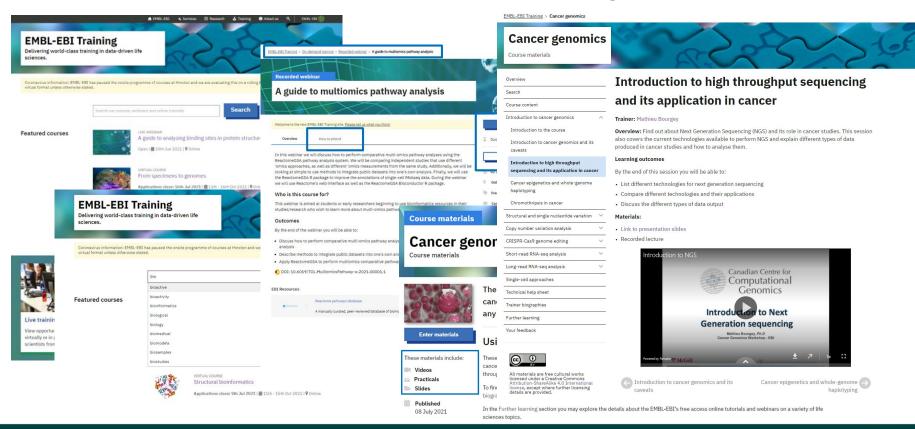
Dive straight into learning through our online tutorials and library of recorded webinars anytime, anywhere.



Support for trainers \varTheta

Looking for training inspiration? Find out how we support trainers and educators in Bioinformatics.

How we implement FAIR for training materials





Big challenges and wished for solutions

- Naming conventions for different types of training and materials that everyone can easily understand
 - Regular user testing to check terminology is understandable
- Keeping materials up to date is time consuming (but very worthwhile)
 - Making our trainers aware of updating requirements when creating materials
- Interoperable can be a challenge for practical materials
 - Virtual machines, the same set up as used in live courses, made available on AWS



H3ABioNet

Verena Ras



Who we are

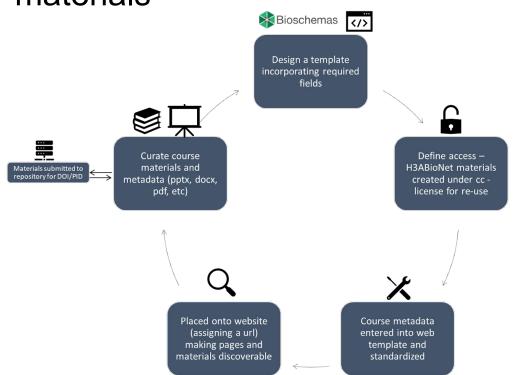
- Pan African Bioinformatics Network as part of the <u>Human Hereditary and Health in Africa Consortium</u> (H3Africa)
- "H3ABioNet was established to develop bioinformatics capacity in Africa and specifically to enable genomics data analysis by H3Africa researchers across the continent. H3ABioNet is developing human capacity through training and support for data analysis, and facilitating access to informatics infrastructure by developing or providing access to pipelines and tools for human, microbiome and pathogen genomic data analysis."
- Major goal is to increase the number of qualified bioinformatics graduates on the continent while creating research opportunities and providing financial support for promising newly-graduated bioinformatics students in Africa, as well as attracting Africans studying abroad back to the continent.
- H3ABioNet delivers high quality training in a variety of formats (see image below).
- Increasing presence across Africa and now globally







How we implement FAIR for training materials

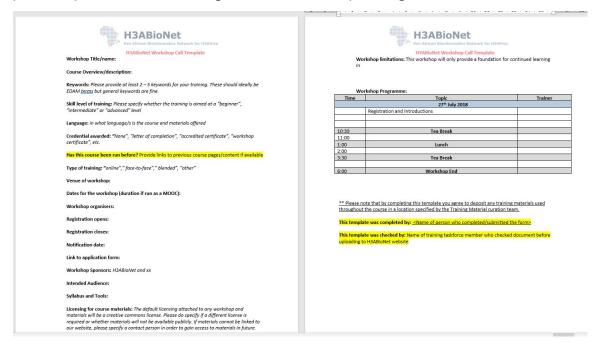


- Developed a repeatable process
- Aligning data to bioschemas
- Submitting data to internal repository (local instance of Figshare) for PIDs
- Adding licensing information to all slides
- Early stages in the process of implementing bioschemas and creating web page templates



How we implement FAIR for training materials

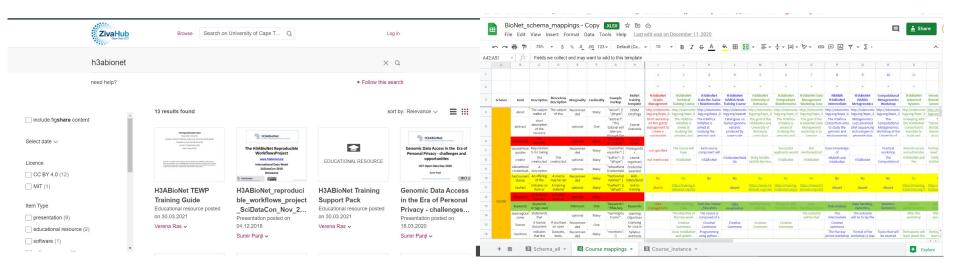
Example template for collecting metadata and placing onto website:





Big challenges and wished for solutions

- Manual process needs automation, especially for sustainability
- Many don't understand FAIR and the benefits so not always easy onboarding people
- Bioschemas not as intuitive to implement when you already have your own framework in place - need for a simple reproducible method for implementation





Individual Trainer

Maria Doyle

Who I am

- Application and Training Specialist for Research Computing at Peter MacCallum Cancer Centre, Melbourne, Australia
- Provide both internal and external training
- External training collaborations
 - Melbourne Bioinformatics
 https://mblue9.github.io/r-intro-biologists/intro_r_biologists.html
 - Galaxy Training Network
 https://training.galaxyproject.org/training-material/hall-of-fame/mblue9/
 - TidyTranscriptomics
 https://github.com/stemangiola/tidytranscriptomics#workshops





How I implemented FAIR for training materials*

R1	GitHub Pages website	
R2	Bioconductor Biocviews	
R3	GitHub release -> Zenodo DOI	
R4	TeSS manual registration (for some)	
R5	Open access website	
R6	R markdown	
R7	CC-BY-4.0 GitHub Licence	
R8	Prerequisites and Learning Outcomes on website	
R9	Contributing.md file in website	
R10	Provide frozen version of past workshops through <u>GitHub website for workshop and Docker</u> , newer workshops use <u>updated material</u>	

^{*}Started thinking how to do as saw <u>FAIR Rules</u> & using Bioc <u>template</u> (now used by <u>>50 workshops</u>). Green: part of Bioc template. White: manual steps.

Big challenges and wished-for solutions





Challenges	Solutions wished
How exactly to make material FAIR?	Checklist and how-tos e.g how to get DOI from Zenodo, what should go in metadata?, contributing file? What I did is here
What does FAIR training material look like?	Examples of training material that is FAIR so can see specifics, describing how they meet FAIR rules
How to make FAIR easily doable?	Automated way to be FAIR. IDEALLY: included in Bioconductor workshop GitHub Actions template here
How to know that material is FAIR?	Badge to add to material. Could then add similar to other badges <u>here</u>

^{*}White is what I've also heard from others, Grey is personal wishes that I think might make FAIR easier for others



Australian BioCommons

Melissa Burke

Who we are

Australian BioCommons is building digital capability for the life science communities



Support volunteer trainers to deliver events nationally





How we FAIRify materials

Review	Assess	Develop	Do
Catalogue materials	How FAIR are we?	Inspiration	Our plan
WhatWhereHowWho	ARDC FAIR assessment tool 10 Simple Rules Quick wins → Improve findability	ARDC Metadata checklist CINECA hackathon Other training providers	Zenodo - Event metadata - Materials - Link to materials shared elsewhere - Crosslink to
https://zenodo.org/communit Australian BioCommons	→ Improve metadata ies/australianbiocommons-train	Who are we FAIRifying for?	TeSS etc Keep learning and adapting!

Big challenges and wished for solutions

Challenge	Solution/wish list	
Translating the principles into practical steps	 Checklists to help make sure everything is done Tools/method for checking FAIRness and finding areas for improvement How to guides/ Case studies 	
Materials are owned by others	Incentives - make it easy, citable DOI, give credit	
Materials shared elsewhere	Crosslink via Zenodo	
Time	 Make it part of the process Templates Automated way of copying info between registries 	



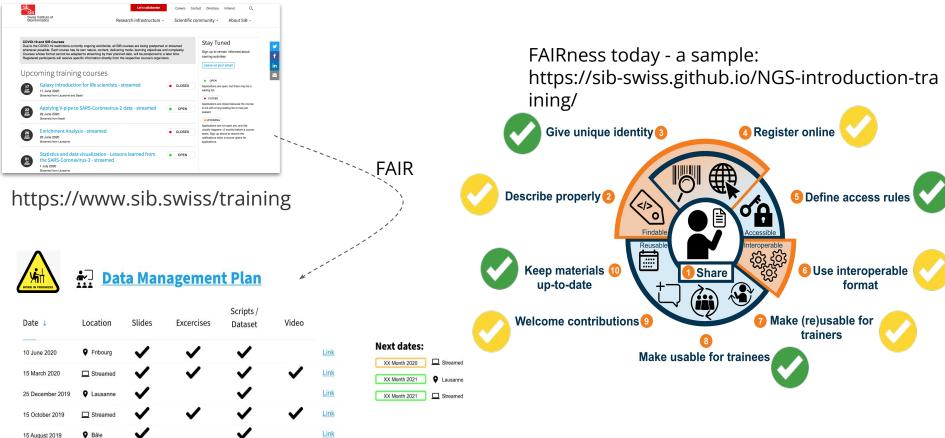


SIB Swiss Institute of Bioinformatics

Patricia M. Palagi

Who we are - SIB Training





How we implement FAIR for training materials*

R1	Not yet systematically, should be automated (or at least partially) with new SIB website	
R2	Bioschemas, EDAM. Internal terminology for terms as materials, levels, etc	
R3	GitHub, DOI with Zenodo, ORCID for trainers and managers	
R4	TeSS scrapper, (GOBLET)	
R5	Moving towards Open work in progress	
R6	MD, PPT, PDF, Jupyter Notebooks	
R7 R8	CC-BY-4.0, copyright guide, Bioschemas metadata, but more data needed to support trainers and trainees (eg lesson time duration)	
R9	GitHub - contribution rules can be improved	
R10	GitHub-Zenodo version control stamp, SIB Training Zenodo Community	





Big challenges and solutions wished

Challenges	Solutions wished
Change in mindset: from authorship to contributorship	Need to continue raising awareness of the benefits - Recognition of trainers is one of them
Collating all pieces of the puzzle together - technically not simple	Checklists and guidelines - no one solution fits all
Common agreement & standardization	Training terminology
It is time consuming	Automation, FAIR training materials checker Learning with and from the community

Life/Success is a journey, not a destination

Open and FAIR is a journey and a destination



Open Aire Community of Practice for Training Coordinators

Celia van Gelder



Open Aire Community of Practice for training coordinators

CoP: Who we are and our training resources

CoP for training coordinators

- Started in 2018; monthly meetings and a Slack channel
- For training coordinators in communities, projects and research infrastructures
- A discipline transcending community of trainers and training organisers
- © Owned and driven by members... sharing, collaborating, contributing, coordinating
- Strong link with themes Open Science, Research Data Management (RDM)
 and Data Stewardship
- Examplesof activities:
 - Organizing workshops
 - Webinar/writing sprint about GDPR and training
 - Several CoP members were part of the EOSC Working Group Skills & Training (2020), <u>Digital Skills for FAIR and Open Science report</u>

CoP Training resources

- Individual trainings/courses and their materials
- Course Catalogues/Registries
 - ELIXIR TeSS, ENVRI-FAIR Training Catalogue, SSHOC training toolkit, etc, etc
- EOSC-Catalogue (under development)
 - Catalogue of catalogues







www.openaire.eu/cop-training



Activities related to FAIR training resources by CoP & CoP members

www.openaire.eu/cop-training

- Working on findability and interoperability for training resources & catalogues
 - **RDA ETHRD-IG Task force** minimal metadata for Training resources
 - Catalogues Task force in the EOSC Skills & Training Working Group (2020)
 - FAIRSFAIR D6.3: checklist (built on/added on the 10SR paper)
 - ELIXIR Training: bioschemas to annotate Training resources in <u>ELIXIR Training Portal TeSS</u>
 - o <u>Terms4fairskills</u>: a terminology for the skills necessary to make data FAIR and to keep it FAIR
- Metadata & Training Catalogues related workshops by/with CoP members
 - Workshop Training in the EOSC (Feb 2020, the Hague)
 - <u>Training Resource Catalogue Interoperability Workshop</u> (October 2020, virtual)
 - Workshop on <u>Harmonising Training Resource Metadata for</u> <u>EOSC Communities</u> (April 2021, virtual)
 - Have informed EOSC Training Catalgue efforts

Workshop report & recommendations





Our challenges & wish list related to FAIR training resources

www.openaire.eu/cop-training

- Main challenges:
 - Findability of training resources
 - Interoperability of training catalogues/registries
 - Sustainability
- CoP wants to bundle expertise and provide guidance to training coordinators, and to the EOSC Board/Governance
- July 2021 CoP started a Task Force about Making training materials FAIR
 - Aims:
 - Review of current relevant projects and activities and to what extent they implemented FAIR for their training materials
 - Identify implementation studies/examples
 - Outputs by Jan 2022 Best practice guidance, article, possible assessment tool
- Join us for the next round of discussions at the Open Science Fair on September 20 23, 2021: https://www.opensciencefair.eu - Fostering local and global open science communities