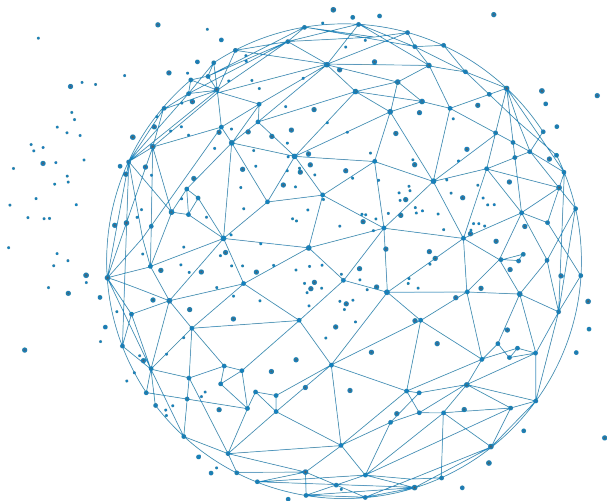First edition 2023 in Fréjus

ÉCOLE THÉMATIQUE
ETBII
BIOINFORMATIQUE INTÉGRATIVE IFB

Making training materials Findable, Accessible, Interoperable and Reusable (FAIR)

Hélène Chiapello, Olivier Sand
& Lucie Khamvongsa-Charbonnier

DOI version final, d'après https://elixir-fair-training.github.io/FAIR-training-handbook/ (in progress)

Hélène Chiapello https://orcid.org/0000-0001-5102-0632
Olivier Sand https://orcid.org/0000-0003-1465-1640
Lucie Khamvongsa Charbonnier https://orcid.org/0000-0002-1194-0546

IFB
INSTITUT FRANÇAIS DE BIOINFORMATIQUE

eliXir
FRANCE

cnrs INRAE Inserm cea INVESTIR L'AVENIR

- Preparing training material is time-consuming => recycling => FAIR
- Share => create a virtuous circle
- FAIR ≠ OPEN (accessibility issues)
- FAIR principles applicable by humans AND machines

- Interoperability / reusability
- clear structure of repository
- self-explicit file / folder names
- Master Doc, README file
- file formats
- metadata
- text heavy slides vs separate extensive narrative (ppt Notes panel…)

Format	Advantages	Disadvantages
PPT and PPTX	• Easily (re)usable • Available to multiple OSs/Software • Widespread	• Limited way to provide detailed training instructions • Not version controlled
Keynote	• Polished overall aesthetic	• Limited to macOS family • Not version controlled
PDF	• Can be displayed identically in any environment	• Not easily editable • Not version controlled
TeX	• Easily editable • Version controlled • Free	• Steep learning curve for trainers
MD, RST, and HTML	• Version controlled Free	• Rendering (need templating to transform into HTML)
Google slides	• Version controlled Free	• Not always possible to use owing to local/institutional policies • Not always accessible (depending on geographic location)

MD, Markdown; PDF, Portable Document Format; PPT, PowerPoint; PPTX, PowerPoint Open XML Presentation; RST, reStructuredText
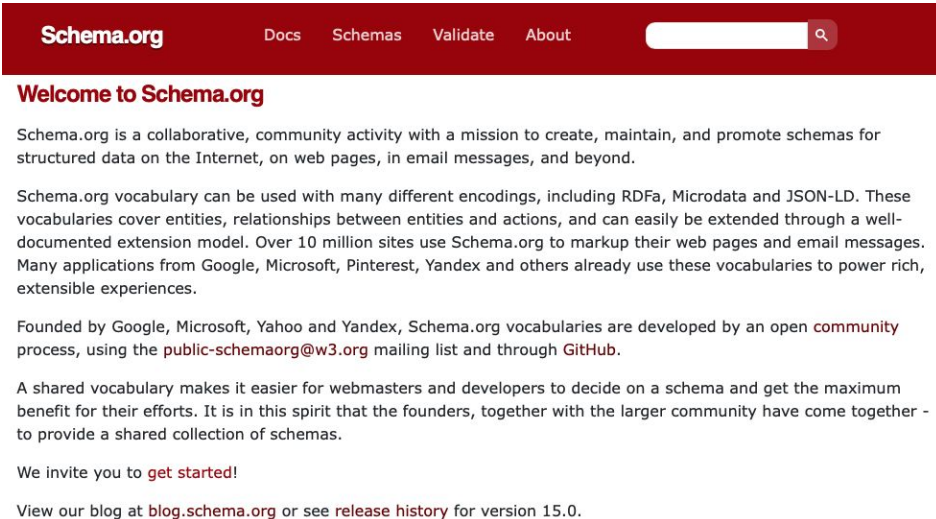
https://doi.org/10.1371/journal.pcbi.1007854.t001

- phases of training materials life cycle:
    - development
    - running course
    - sharing
    - archiving
- platforms:
    - Google Suite
    - GitHub/GitLab
    - Zenodo
    - a web portal
    - a Learning Management System (LMS)
- depending on the phase, some platforms are more adapted than others

- metadata = data describing *other* data
- metadata improves F, I & R of the *other* data
- quick overview of data
- examples: authors, dates, format, audience…
- metadata standards, controlled vocabularies, ontologies…
- Bioschemas profiles: Course, CourseInstance, TrainingMaterial

Schema.org   Docs   Schemas   Validate   About

Welcome to Schema.org

Schema.org is a collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.

Schema.org vocabulary can be used with many different encodings, including RDFa, Microdata and JSON-LD. These vocabularies cover entities, relationships between entities and actions, and can easily be extended through a well-documented extension model. Over 10 million sites use Schema.org to markup their web pages and email messages. Many applications from Google, Microsoft, Pinterest, Yandex and others already use these vocabularies to power rich, extensible experiences.

Founded by Google, Microsoft, Yahoo and Yandex, Schema.org vocabularies are developed by an open community process, using the public-schemaorg@w3.org mailing list and through GitHub.

A shared vocabulary makes it easier for webmasters and developers to decide on a schema and get the maximum benefit for their efforts. It is in this spirit that the founders, together with the larger community have come together - to provide a shared collection of schemas.

We invite you to get started!

View our blog at blog.schema.org or see release history for version 15.0.

Bioschemas

- Persistent = unique and permanent
- globally unique => namespaces governed by clearly defined authorities
- permanent => maintained for a long period of time
- resolvable => allows humans and machines to access the resource
- PID enhances visibility => F & R
- associated deposit works as a backup
- examples: DOI (Digital Object), ORCID (Open Researcher and Collaborator), PMID…

How to make your training materials accessible ?

> Decide appropriate access rules

> Eventually limit access using authentication and authorization

> Choose a repository to share training material

> Be inclusive regarding accessibility of training material
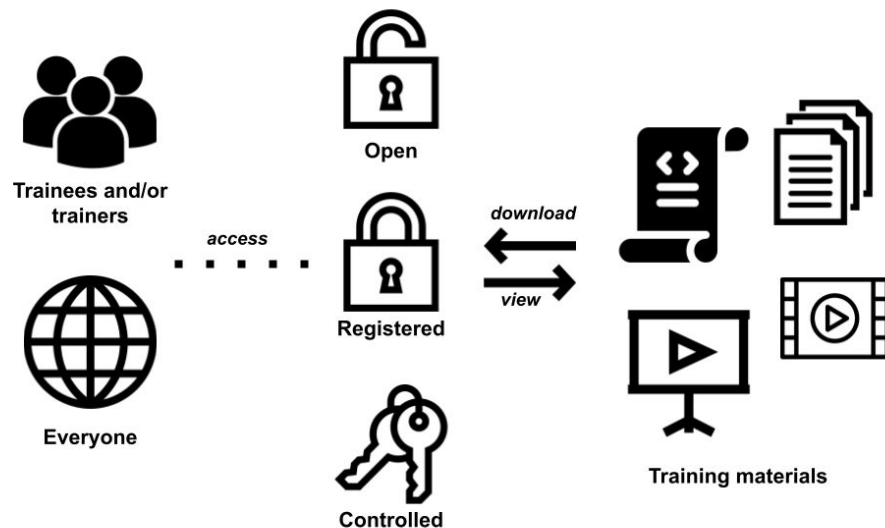
From: flaticon.com, Flaticon license, graphic designer: Freepik, Octopocto, designvector10

- Accessibility defines **who, when, where** (via a web site, a repository) **and how** (using a login, programmatically via an API, etc.) to access training materials
- Recommandations
  - Use standard (http, api, etc.) and (if possible) automated protocols (scrapers, crawlers)
  - Define levels of access : Open, Registered, Controlled
  - Describe the level of access (metadata)
  - Promote inclusiveness (use several languages, adapt to specific need such as people with poor eyesight and colour-blind might, use fonts compatible with dyslexia problems,…)
- Benefits expected for trainees (ex: better learning) and trainers (ex: have feedback from other trainers)

How to make your training materials reusable ?

> What information allow other trainers to reuse training materials ?

> How to write learning outcomes

> How to add useful contextual information

> How to construct a detailed description which allow reusability of training material ?

- Use **rich and standard Metadata** to describe training materials and training sessions
  - look at Bioschema and RDA standard terms for trainings
  - add keywords relevant to your field
- Write **learning outcomes** in a SMART format
  - it can help you define the purpose, intended audience of your training session and what participants need to know (the prerequisites) before joining it
- Provide additional **contextual information (instructor notes)**
  - Ex: *required resources (hardware, softwares, datasets,...)*, *structure and duration* and *additional notes*

8. Licensing

Licences are a standardised way of giving others specific permissions to use work you created
The most common type of licence used for training materials is **Creative Commons**

Checklist

- Clearly label your materials with the chosen licence
- Include your contact information in case anyone needs to ask you any questions about your materials or their use
- Cite any other creators' materials that you include in yours - do not reuse others' works if you are not sure about their permissions
- Check with your institute/company if there are any specific licence rules or guidance

https://creativecommons.org/choose/
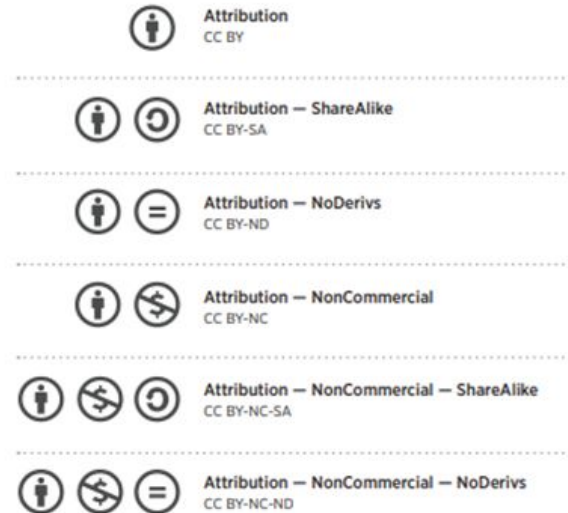
Image taken from https://wiki.creativecommons.org/images/6/6d/6licenses-flat.pdf. The image is in the public domain, CC0.

Let learners, trainers and anyone interested know about your training materials

Share training materials using

- Training registries:
    - TESS (ELIXIR Training e-Support System)
    - GOBLET (Global Organization for Bioinformatics Learning, Education and Training)
- General online repositories: Zenodo, Github, Gitlab
- Social media: Twitter, Facebook, LinkedIn, Instagram, YouTube,..
- Blogs
- Congresses or other events

Interested in giving/receiving feedback on materials/content

- Add contribution to your training material
  - The rules for participation and contribution need to be defined
  - Contributing file, open source, interoperable with repositories (GitHub)
- Importance of attribution of credit
  - Attributing credit is as important for training material as it is for scientific literature
  - roles: author, trainer, contributor, funder, sponsor (matrix)
- Using material of others
  - Academic citations/reference style (ORCID)
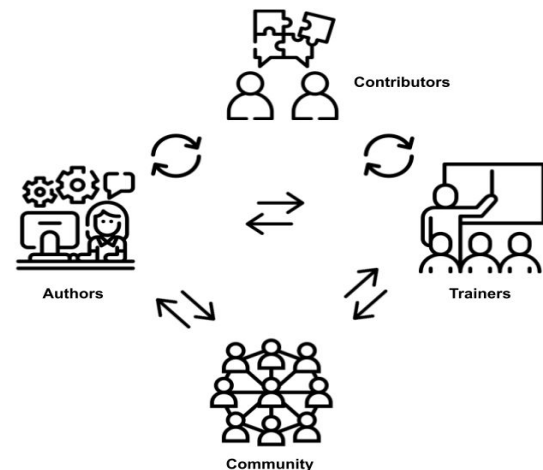
From: flaticon.com, Flaticon license, graphic designer: wanicon, rddrt, Eucalyp, Freepik, Vectors Market, Kharisma

Creating training materials

1. provide guidelines on how to cite them
2. credit the authors and other contributors