Public repositories Episode Notebook

Part of FAIR in (bio) practice, <https://carpentries-incubator.github.io/fair-bio-practice>

Type your name and institution:

-

-

-

**Exercise 1. Public general record**

Have a look at the following record for data set in Zenodo repository:

<https://doi.org/10.5281/zenodo.5045374>

1. What elements make it FAIR?

FINDABLE (persistent identifiers, easy to find data and metadata):

-

-

ACCESSIBLE (The (meta)data retrievable by their identifier using a standard web protocols):

-

-

INTEROPERABLE (The format of the data should be open and interpretable for various tools):

-

-

REUSABLE (data should be well-described so that they can be replicated and/or combined in different settings, reuse states with a clear licence):

-

-

1. Skim through the data set description (HINT there is also a README), try to judge using marks from 0 to 5 (5 best) if

* Is it clear what the content of the data set is:
* Is it clear why (what for) the data could be used:
* Is it well described?
* How confident will you be to work with this data set:
* How easy is it to access the data set content?

1. Give +1 to the statement that the best describes your latest public dataset or a current project folder:

* It does not have a readme like description:
* Its description has only one/two paragraph(s), I have not thought of adding so many details:
* My data set description is similar in details:
* My data /project description is much richer or formalised:

**Exercise 2. Dataset discovery**

Try to find either:

- similar data sets in Zenodo

- data sets interesting for you

Judge using marks from 0 to 5 (5 best)

- how easy is to find similar and RELEVANT dataset:

- It is clear what the content of the data set is:

- It is clear why (what for) the data could be used:

- They are well described:

**Exercise 3. Domain specific repositories (domain == area of knowledge, here data type or study subject).**

Select one of the following repositories based on your expertise/interests:

Have a look at mRNAseq accession 'E-MTAB-7933' in [ArrayExpress](https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-7933/)

* What makes it better than Zenodo:
* What domain specific features can you see:
* Searching:

> - Have a look at microscopy 'project-1101' in [IDR](https://idr.openmicroscopy.org/webclient/?show=project-1101)

* What makes it better than Zenodo:
* What domain specific features can you see:
* Searching:

> - Have a look at the synthethic part record 'SubtilinReceiver\_spaRK\_separated' within the 'bsu' collection in [SynBioHub](https://synbiohub.org/public/bsu/SubtilinReceiver\_spaRK\_separated/1)

* What makes it better than Zenodo:
* What domain specific features can you see:
* Searching:

> - Have a look at the proteomics record 'PXD013039' in [PRIDE](https://www.ebi.ac.uk/pride/archive/projects/PXD013039)

* What makes it better than Zenodo:
* What domain specific features can you see:
* Searching:

> - Have a look at the metabolomics record 'MTBLS2289' in [Metabolights](https://www.ebi.ac.uk/metabolights/MTBLS2289/descriptors)

* What makes it better than Zenodo:
* What domain specific features can you see:
* Searching:

**Exercise 4. Finding a repository**

Check the publisher's / funder' recommended list of repositories, some of which can be found below:

- [BioMed Central / Springer Nature](https://www.springernature.com/gp/authors/research-data-policy/recommended-repositories)

- [eLife](https://submit.elifesciences.org/html/elife\_author\_instructions.html#policies)

- [Elsevier](https://www.elsevier.com/about/policies/research-data)

- [EMBO Press](https://www.embopress.org/page/journal/14602075/authorguide#datadeposition)

- [F1000 Research](https://f1000research.com/for-authors/data-guidelines)

- [GIGAscience - OUP](https://academic.oup.com/gigascience/pages/instructions\_to\_authors)

- [PLoS](https://journals.plos.org/plosbiology/s/recommended-repositories)

- [Scientific Data - Nature](https://www.nature.com/sdata/policies/repositories)

- [Taylor and Francis](https://authorservices.taylorandfrancis.com/data-sharing-policies/repositories/)

- [BBSRC](https://bbsrc.ukri.org/research/resources/)

- [NERC](https://nerc.ukri.org/research/sites/environmental-data-service-eds/policy/)

- [Royal Society](https://royalsociety.org/journals/ethics-policies/data-sharing-mining/)

- [Wellcome Open Research](https://wellcomeopenresearch.org/for-authors/data-guidelines)

Check [Fairsharing recommendations](https://fairsharing.org/recommendations/?q=)

- alternatively, check the [Registry of research data repositories - re3data](https://www.re3data.org/)

1. Find and type a repo for genomics data:
2. Your favourite/chosen data type and recommended repo for it:

-

-

-

1. List repositories you have used (either to download or to share your data):  
   -  
   -  
   -