# Workshop Biometadata-05 - Data deposition walkthrough[[1]](#footnote-20)

#### Select your repository

* Check your data management plan (DMP), if any
* Ask where colleagues and PI prefer to upload/look-up data and metadata
* Otherwise, use the following resources to find a suitable repository:
  + [with the NFDI4Microbiota Knowledge Base](https://knowledgebase.nfdi4microbiota.de/Research-Data-Management/22-data-repositories.html)
  + [with the EMBL-EBI data submission wizard](https://www.ebi.ac.uk/submission/)

#### Check for mandatory criteria or mandatory metadata fields

* Look for mandatory fields in your selected repository
  + Otherwise, use [ENA’s checklists](https://www.ebi.ac.uk/ena/browser/checklists), or [NFDI4Microbiota’s minimal metadata suggestions](https://github.com/NFDI4Microbiota/MetadataStandards)
* Watch for suggested metadata templates in your selected repository
  + Otherwise, browse and generate your metadata template via [QIIMP](https://qiita.ucsd.edu/qiimp/) or use [NMDC’s submission sandbox](https://data-sandbox.microbiomedata.org/submission/) to pre-fill and validate your metadata
* Note whether specific units or value types are expected
* Check if mandatory criteria include selection of terms from existing ontologies
  + If yes, use the filters of the [OLS](https://www.ebi.ac.uk/ols4) to refine your search and find relevant terms that apply to your dataset. For instance, for sequencing data, ENA requires terms from ENVO and/or UBERON ontologies for the metadata fields env\_broad\_scale, env\_local\_scale, env\_medium
  + Use [ontology annotators to look-up multiple terms](https://clavellab.pages.rwth-aachen.de/workshop-biometadata-05/lectures/ontologies-for-biological-data.html#/annotate-list-of-terms)

#### Assess the quality of ontology and their terms

* Use the [workshop’s criteria](https://clavellab.pages.rwth-aachen.de/workshop-biometadata-05/lectures/ontologies-for-biological-data.html#/selecting-ontology) to ensure the ontology:
  + has sound terms definitions that you agree with
  + is actively developed and updated to reflect the scientific practices
  + has a community space to exchange with the community (tickets, issues)
  + has an adequate license and appropriate funding (for maintenance)

#### Fill out your metadata while you research

* Use the generated metadata template and regularly update the metadata fields regarding your samples and research techniques

#### Upload your (meta)data before your desired publication date

* Check if you can create a project in your desired repository **in advance**
* Upload your metadata template or pre-fill the fields
* Create an embargo if you need to delay the access to the data and metadata

1. This document is available at <https://doi.org/10.5281/zenodo.10605803> and is a supplement to the [Workshop Biometadata-05](https://clavellab.pages.rwth-aachen.de/workshop-biometadata-05/index.html) with examples from sequencing data. [↑](#footnote-ref-20)