Using different types of standards

The good practice of using standards (e.g., formats, vocabularies, processes) during research can take many forms, but always leads to consistent and interoperable data. You can use **standards to codify the methodologies** for generating or reusing data, or to **describe data itself**.

You can reuse existing standards or – more sporadically and only if strictly necessary – create new ones, specific to your research.

To make sure that all researchers involved in the project follow a standard **methodology**, for example, you can use the **ISO standards** specific to your disciplinary domain.

Standards in the form of vocabularies, taxonomies and ontologies can be used to univocally structure data and make it interoperable with both existing and future data. Using at least one vocabulary when collecting information can help you to correctly identify data and facilitate their reuse, by making them understandable to others who share the same vocabulary.

'**Vocabularies**' or 'thesauri' are lexical resources comprised of controlled terms used to describe a set of items, knowledge, data, theories from a certain scientific field or disciplinary domain. They provide standard terminology, improving the value of data and making them machine-readable.

Items in '**taxonomies**' are not only described by controlled vocabularies of selected terms, but also ordered in a system, usually a hierarchical one.

In '**ontologies**', the types and therefore the names of the relationships between the items that make up a taxonomy or a vocabulary are made explicit.

On the other hand, specific metadata standards should be used to describe data 🗐 **Metadata and documentation**.

🎬 In the field!

I work on a collaborative project that also involves companies and research centres outside of the University of Bologna. How do I make sure that my results are of high quality and interoperable?

To achieve this goal, common standards need to be shared within the collaborative project.

ISO standards are rules that describe the best way of doing something, as agreed by international experts in the field. Standards cover a wide range of activities – from making a product to managing a process, to providing a service or supplying materials.

For example, quality management standards are designed to help work more efficiently and minimise product defects. Health and safety standards can reduce accidents in the workplace. IT security standards contribute to protecting sensitive information.

I want to use vocabularies and taxonomies to describe the information I gather in my data. What benefits does this bring and where do I find the most suitable ones?

A vocabulary can contain standard terms from a specific disciplinary domain, which can be used to identify the variables in an analysis, to then include them in a tabular file as column headings, for example.

In research with qualitative data, using terms from a standard vocabulary can help express relationships between variables within a text such as a technical report, so that all those involved in the project know exactly what event or type of data is being referred to.

There are both generic schemas, such as the Information Artifact Ontology (that describes information entities, understood as pieces of information encoded in digital or physical entities), as well as extremely specific ones, such as the IUPAC Compendium of Chemical Terminology (which contains some 7,000 chemical concepts derived from the IUPAC Recommendations).

You can use a registry, like FAIRsharing (see box below), to look for schemas, vocabularies and thesauri that suit your needs.

🔗 Useful links

ISO Standards
https://www.iso.org/standards.html

Search tool for metadata schemas, vocabularies and thesauri
https://fairsharing.org/search?fairsharingRegistry=Standard