# Metadata Creation

Brief guide that explores the topic of metadata: What are metadata? What kinds of metadata are there? What is the process of creating metadata?

Source: <https://guides.library.ucsc.edu/c.php?g=618773&p=4306436>

## Introduction

The purpose of this guide is to briefly explore the questions:

* What are metadata?
* What is metadata used for?
* What is the metadata creation process?

Metadata are everywhere. As Jenn Riley observes in her primer, [*Understanding Metadata*](http://www.niso.org/apps/group_public/download.php/17446/understanding%20metadata) (NISO, 2017), "Metadata are pervasive in information systems, and come in many forms. [...] Metadata is key to the functionality of the systems holding the content, enabling users to find items of interest, record essential information about them, and share that information with others."

This guide deals with metadata as it is created within a library, or more broadly, a digital collections setting.

What are metadata?

The most common definition is that metadata are "data about data." A better definition is that metadata are deliberate, structured data about data.

Metadata are used to facilitate and support resource discovery, identification, the organization of resources, and the exchangeability of the data itself as well as the exchangeability of the resource or resources it represents.  Metadata also capture and provide important contextual details, as not all resources are self-describing.

Types of metadata

There are many different kinds of metadata. In the world of digital objects, metadata is usually divided into 3 to 5 categories:

* Administrative metadata, including:
  + **Rights metadata** (i.e., intellectual property rights and use information)
  + **Technical metadata** (i.e., technical details about the object and its instantiation like its file format, file size, and how to open, access and use it
  + **Preservation metadata** (i.e., a log of the series of actions taken against an object in order to ensure it longevity and viability)
* **Descriptive metadata**describes a resource, its content, its identifying characteristics and its "aboutness"
* **Structural metadata** describes how the pieces of a single object fit together and how an object exists in relationship to other objects

## Tools of metadata creation

* Schema

A metadata schema is a list of elements or defined data points that are used to capture information about a resource. Some of these data points might include a title, an identifier, a creator name, or a date.

* Standards

Standards are the tools that tell us how to populate each of the data elements within a schema.  There are three types of standards:

* Content standards

Content standards describe the use of each element or what pieces of types of information go where.  They also give guidance on how to best record or transcribe that information: Where should that information be coming from? What is the best source of information? Which elements should require the use of data values standards and, if so, which value standards should be used?

* Data value standards

Data value standards are lists of standardized subject terms, genre terms, names, etc.  Examples of data value standards include controlled vocabularies like Library of Congress Subject Headings and other discipline-specific thesauri, and encoding or formatting standards like ISO 8601 which prescribes how dates ought to be represented.

* Data structure standards

Data structure standards (XML, RDF, etc.)  exist to tell us how to encode and structure the metadata record so as to ensure its machine readability. It is important to note that there are primary two users of metadata: humans and machines. Any metadata you create must be intelligible to both.

* Best practices

Best practices guide or prescribe the selection and use of a particular metadata schema, which subset of elements within that schema will be used to comprise a metadata profile for use in describing a certain class or collection of objects.  Best practices also serve to: 1) better define or clarify within the context of a given project or organization the use of each metadata element; and 2) more decisively recommend the use of particular content, data value, and data structure standards.

## Quality of metadata

When creating or evaluating metadata it is important to ask:

* **Accuracy:** Is the data recorded correct and factual?
* **Completeness:** Has all relevant data been recorded in full?
* **Consistency:** Has data been entered consistently? Is the same set of metadata elements being used to describe all of the resources in your collection?  Is data being entered in the same format?
* **Interoperability:**Is your data machine readable? Can your metadata be easily migrated to and understood by another system? Can it be aggregated with other metadata sets or collections?
* Background
* The Dublin Core Metadata Element Set is one of the simplest and most widely used metadata schema. Originally developed to describe web resources, Dublin Core has been used to describe a variety of physical and digital resources.
* Dublin Core is comprised of 15 “core” metadata elements; whereas the "qualified" Dublin Core set includes additional metadata elements to provide for greater specificity and granularity.
* Built into the Dublin Core standard are definitions of each metadata element – like native content standard – that state what kinds of information should be recorded where and how.  Associated with many of the data elements are data value standards such as the DCMI Type Vocabulary and ISO 639 language codes, etc. More information can be found on the [Dublin Core Metadata Initiative website](http://dublincore.org/documents/dces/).
* Dublin Core Metadata Element Set

| Dublin Core Element | Use | Possible Data Value Standards |
| --- | --- | --- |
| Title | A name given to the resource. |  |
| Subject | The topic of the resource. | [Library of Congress Subject Headings (LCSH)](http://authorities.loc.gov/) |
| Description | An account of the resource. |  |
| Creator | An entity primarily responsible for making the resource. | [Library of Congress Name Authority File (LCNAF)](http://authorities.loc.gov/) |
| Publisher | An entity responsible for making the resource available. |  |
| Contributor | An entity responsible for making contributions to the resource. | [Library of Congress Name Authority File (LCNAF)](http://authorities.loc.gov/) |
| Date | A point or period of time associated with an event in the lifecycle of the resource. | [W3CDTF](http://www.w3.org/TR/NOTE-datetime) |
| Type | The nature or genre of the resource. | [DCMI Type Vocabulary](http://dublincore.org/documents/dcmi-type-vocabulary/) |
| Format | The file format, physical medium, or dimensions of the resource. | [Internet Media Types (MIME)](http://www.iana.org/assignments/media-types/) |
| Identifier | An unambiguous reference to the resource within a given context. |  |
| Source | A related resource from which the described resource is derived. |  |
| Language | A language of the resource. | [ISO 639](https://www.loc.gov/standards/iso639-2/php/code_list.php) |
| Relation | A related resource. |  |
| Coverage | The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. | [Thesaurus of Geographic Names (TGN)](http://www.getty.edu/research/tools/vocabulary/tgn/index.html) |
| Rights | Information about rights held in and over the resource. |  |

Step One

When preparing to describe your resources, there are a number of questions that you will want to consider:

* What you are describing?

Are you describing a physical object, a digital object, a digital representation of a physical object?

* What kind of information do you want to record?

Who is my audience? What information is needed to identify the resource? What information is needed to properly contextualize it? How do I want people to find it or interact with it? How do I expect them to search for or discover it? How do I expect to use it? How do I expect others to use it now and in the future? What information is required to communicate who owns it, who can use it and to what extent?

Step Two

As you begin to answer the questions presented in Step One, list out the information that you would like to include as data points, e.g., title, subject, access rights, etc. For example, if you are wanting to overlay images onto a map, you will want to record coordinate data. This is your metadata wish list.

Step Three

Consider the descriptive information or metadata that you may already have: Which elements or what kind of information are recorded or represented there? Is information missing about your resources? Is there information that would challenging to find or create?

Step Four

Find your "golden minimum."  Determine what information is essential to facilitate discovery, identification, and to give sufficient context, but no more. What exactly is the golden minimum in the space of your project depends on your project goals and available resources.

Step Five

Finalize your list of data points. Choose to codify this list as your own metadata schema or map it to an existing schema, such as Dublin Core.

Step Six

Decide whether you want to make use of data value standards (controlled vocabularies, thesauri, encoding or formatting standards). If so, which standards would apply to which fields? Alternatively, you can create your own data value standards, such as, a subject vocabulary specific to your topic or collection of resources or a controlled list of names. Document your decisions as your best practices.

Resources

Baca, M., & Getty Research Institute. (2008). *Introduction to metadata*. Los Angeles, CA: Getty Research Institute. Link: <http://d2aohiyo3d3idm.cloudfront.net/publications/virtuallibrary/0892368969.pdf>

* Setting the stage / Anne J. Gilliland
* Practical principles for metadata creation and maintenance

Dublin Core Metadata Initiative. (2012). *Dublin Core Metadata Element Set, Version 1.1*. Link: <http://dublincore.org/documents/dces/>

National Information Standards Organization (U.S.). (2004). *Understanding metadata*. Bethesda, MD: NISO Press. Link: <http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>

NISO Framework Advisory Group. (2007). *A framework of guidance for building good digital collections*. Bethesda, MD: NISO Press. Link: <http://www.niso.org/publications/rp/framework3.pdf>

* Metadata (pages 63-85)

Riley, J. (2017). Understanding metadata: What is metadata and what is it for? Baltimore, MD: National Information Standards Organization (NISO). Link: <http://www.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf>

Take-home assignment

* [Take-home assignment handout](https://guides.library.ucsc.edu/ld.php?content_id=28741917)