

The Metadata Lifecycle



Introduction

Metadata has always been defined as the "data about the data", however the industry has yet to understand how to effectively collect, manage and use that metadata. The picture below provides a high level concept of how metadata can be used.

The question is how to actually put this picture into practice. In August PHUSE hosted a Single Day Event in Philadelphia followed by a [Metadata Discussion Club at the Annual Conference](#).

This Advance Hub page provides a summary and the beginning of the ongoing discussions around the definition and use of metadata.

Lifecycles

Section content first added by: Kevin Lee

Metadata should be created before datasets and analysis. Then, the datasets will be created and analyses will be done according to metadata. The well-thought, planned process will yield more complete metadata, but in most of cases, metadata will be modified during the process and will be completed almost at the same time, more accurately just before all the analyses and datasets are finished.

Definition of Metadata

Section content first added by: Chris Decker

At the PHUSE annual conference, an engaging discussion was held with the initial goal of defining the the definition of metadata. At the end of this discussion the group came up with the following definition.

Metadata is a collection of information that puts data into context.

Other comments were made to help clarify the definition.

- Metadata can be defined as either metadata or data depending on its use
- Need to work at defining what context means and defining categories of context
- Is a collection of information structured or not structured. The discussion led to collection being structured but the content could be unstructured.

Additionally, the purpose of collecting metadata must be driven by the needs of the business. (Examples to be added).

Do you agree?

Defining Context

Section content first added by: Chris Decker

Within the definition of metadata, the word context was used. Now we must define the different aspects of context that will help better define concrete needs of metadata.

Aspects of Context:

- Medical - the medical meaning of the information
- Technical - how we physical collect, store, maintain, and communicate the metadata
- Structural - What is the structure of metadata model that can support hierarchy and diverse relationships
- Functional - What is the metadata in relation to (e.g. finance, operations, biometrics)

Applications of Context:

- Development of the protocol/SAP
- Collection of clinical data
- Specifications for SDTM/ADaM data
- Metrics to support operational aspects of a study
- Reporting requirements
- Submission
- Reuse of clinical information for future study design or ad-hoc analysis

What else?