Data Definitions Best Practices
December 12, 2013

Writing a data definition:
The definition should explicitly describe and explain the meaning of the data element, data entity or concept represented and careful to describe context in which the term is used. A well-formed definition should be:

• Concise: Described as succinctly as possible.
• Precise: Described using unambiguous words when possible.
• Non-Circular: The term being described should not be used in the definition.
• Distinct: Described in a way that differentiates this data element, data entity or concept from others.
• Unencumbered: The definition should not refer to physical location or how it is created.

The context should include (when possible):

• Broader Term: The class of the thing being defined, often using the "IS A" relationship. (A school is an organization.)
• Distinguishing Characteristics: Attributes that differentiate the thing being defined from other things in the same class often using the "HAS A" relationship. (A school has an academic program.)
• Function Qualifier: A description of how the thing being defined is used or its core functionality, often using a "USED FOR" relationship. (A school is used for educating students.)

All definitions should start with the restating of the term ("An Employee is...") and be written as complete sentences. These recommendations reduce the risk of ambiguity (especially when the definition is copied to another location) and improve the readability and understandability of the content. Additionally, this practice helps to point out any potential circular definitions.

Sample data definition:
Non-Matriculated Student - A non-matriculated student is one who has not been admitted to GW to pursue a degree program.

Where to display data definitions:
All reports or interfaces that present data or serve as an input for data, should include data definitions that clearly describe that data and the context in which it is used on the report or interface.

Where to store data definitions:
The definition may be documented locally in the report or system or maintained in the Data Governance Center (currently accessible at: http://gwu poc.collibra.com. ). Whether the metadata is stored in the report or in the institutional metadata repository, it should be actively maintained and kept up to date. Because this is easier, more consistent and requires fewer resources to maintain, University Data Governance strongly recommends that all metadata is maintained within the Data Governance Center and links to that store are used in all institutional reports and interfaces.

If an authoritative definition exists for a term in a report or interface, the URL of that definition should be provided at any other location that term or concept is defined (whether this be centrally or within the report or system). If, in addition to the link, the definition is rewritten outside of the authoritative location, the system or report owner is responsible for updating both versions within a reasonable time frame and ensuring they match at all times.