Stony Brook University

Data Dictionary Standards

Prepared by the Stony Brook Data Governance Council, June 8, 2017
Approved Nov. 9, 2017, updated March 28, 2018

Following a guidance from the International Standards Organization (ISO) that effective use of data requires a shared "common understanding of the meaning and descriptive characteristics of that data" (2004), the Stony Brook University Data Governance Council (DGC) has set standards for data dictionaries. A data dictionary is a compendium of data definitions for multiple data elements that exist in a data store. For systems providing data, data dictionaries must be maintained to guide data users in the meaning and proper usage of data.

Principles

1. Data dictionaries are designed to promote communication and production of meaning; as such dictionaries document the existence, meaning, and use of data elements
2. Data dictionaries must be accessible to all users who enter and extract data from a data store
3. Data stewards must actively maintain data dictionary contents, including definitions, values, and other metadata
4. Data caretakers and users are responsible for actively using data dictionaries to correctly enter, select, and analyze data elements
5. Data dictionaries should be reviewed on a regular schedule to ensure currency

Required Elements

Dictionaries must include the following required elements:

Dictionary Element	Description	Examples
Data Store	The name of the database	PeopleSoft - csprod
Table name	The name of the table housing the element	PS_STDNT_CAR_TERM
Data element	The alphanumeric sequence used to identify the field.	CUM_GPA ACAD_CAREER
Data element name	The name of the field or term in English. No abbreviations or acronyms allowed.	Cumulative Grade Point Average Academic Career
Definition	A long description of what the field means, a verbal explanation of how it is calculated, how it should be used, and how it should not be used. The definition should not simply re-state the name of the data element.	Computed by the Consolidate Academic Statistics process (SRPCCONS). Total grade points divided by total units. If the student is active in multiple programs and careers with different grading bases (for example, 100-point scale and 4-point scale), then the program logic reports only the student's primary career and program GPA. See the logic description for the OVRD_CUM_GPA field. (from PeopleBook 9.0) Academic career is a concept used in Campus Solutions to designate all course work undertaken by a

Dictionary Element	Description	Examples
		student at an academic institution; you group this course work in a single student record. For example, a university that has an undergraduate school, a graduate school, and several professional schools can define an undergraduate career, graduate career, and a separate career for each professional school (for example, law, medical or dental). You might also make extended education or continuing education its own academic career, or make separate academic careers for every school or college at the undergraduate level. (from PeopleBook 9.0)
Source and Data Logic	Source system, table, and field(s) and data logic, including equations, if applicable. List all sources if multiple sources are used to populate	PS_STDNT_CAR_TERM (see definition for data logic)
Data Type and Length	Numeric, Character; length and decimal format, if appropriate	Numeric (x.xxx) Character (5)
Allowable Values / Parameters	Provides a description of what data may be entered into a field. For numeric fields, this would be minimum and maximum values. For character fields, list allowable values and labels if fewer than 10 or provide name of set-up or dimension table in which values and labels are stored	CUM_GPA range 0 – 4.0 ACAD_CAREER UGRD = Undergraduate GRAD = Graduate DEN = Dental MEDS = Medical PHARM = Pharmacy
Semantic Rules	Explain any patterns and meaning that may be encoded in the code	Term is arranged by century-year-term. 1st character is the century, 2nd & 3rd characters are year, 4th character is a 1, 4, 6 or 8 for winter, spring, summer, or fall terms.
Data steward	Identifies the office of the steward(s) responsible for the data. May indicate groups of individuals if steward is unit dependent, e.g. the Steward for an employee address field would differ from the Steward for a student address field	Registrar's Office
Date created	Lists the date the dictionary entry was created	mm/dd/yyyy
Date updated	Lists the date the dictionary entry was modified. Leave null for newly created entries	mm/dd/yyyy

Recommended or optional elements

Dictionary Element	Description	Examples
Alias names	Lists common terms substituted for the data element	GPA Student Level, Type
Data sensitivity classification	Lists security and access controls per University policy 302: Sensitive Information Classification Policy	High Moderate Low
Usage	Describes typical usage or reports where the data element frequently appears	
Notes		