

Data Element Naming Standards

Why are such standards needed?

- ◆ “help foster a [common understanding](#) of data throughout the University,
- ◆ [promote data sharing](#) across systems and among data users by providing [uniformity in how data is defined and named](#),”¹
- ◆ “facilitate the [discovery of data](#) which can be shared throughout the University, and
- ◆ provide uniformity which enables the [control and management](#) of the University data resource.”²

Where are such standards used?

- ◆ **“Data element definitions** – English phrase (or phrases) which describe a data element. Consists of [prime words, optional qualifiers, and class words](#).
- ◆ **Logical data element names** - uniquely identifies a data element within the University's data resource. Derived from the data element definitions.
- ◆ **Physical data element names** – required by the operating software (database) to uniquely identify the data and manage it. Developed from the logical data element names.”³

What are the components of element names?

- ◆ **“Prime words** - describe the subject area of data”,⁴ and “represents the business portion of the name.”⁵ (*See attached list of Prime Words.*)
- ◆ **“Class words** - describe the major classifications or types of data associated with data elements. For each data element defined and named, one class word becomes part of the data element definition and associated logical and physical names.”⁶ (*See attached list of Class Words.*)
 “There are four categories of class words:
 - [Chronology](#) – indicate a point in time, span of time.
 - [Measurement](#) – have dimension, capacity, amount or duration.
 - [Identification](#) – distinguish a person, place or thing.
 - [Text](#) – identify more free form or narrative data.”⁷
- ◆ **“Qualifiers** - further define and distinguish the prime and class words.”⁸ (*See attached list of Qualifiers.*) Qualifiers should be used as necessary to fully identify the data element, and should be listed in an order that has logical English meaning.⁹

¹ Amy Brooks and Judy Smutek, “It’s All in the Name – Implementing Data Element Naming Standards,” CAUSE92 (December 1-4, 1992): 1.

² Brooks and Smutek, 2.

³ Brooks and Smutek, 4.

⁴ Brooks and Smutek, 4.

⁵ University of Illinois Data Management, “Data Naming Standards,” version 2, (January 2002), http://www.ait.s.uillinois.edu/datamngmnt/data_naming_standards.doc (accessed May 21,2003).

⁶ Brooks and Smutek, 4.

⁷ University of Illinois Data Management.

⁸ Brooks and Smutek, 4.

Where are the general rules for the standards on each of the following?

- ◆ **Data element definitions should –**
 - include the classification of the data,
 - identify the prime word,
 - be clear and concise,
 - define what data is not what it is not, and
 - should note if this data element is derived from others.
- ◆ **Logical data element names should-**
 - be of the format *prime word (required) – qualifier (optional) – class (required)*,
 - not use abbreviations unless required to meet a name length limitation in the tool,
 - nouns are singular and verbs in present tense.

Prime Word	Qualifier (or Secondary Prime Word)	Class Word	Logical Name
Account		Balance	Account Balance
Employee	Salary	Amount	Employee Salary Amount
Employee	Prime Department	Code	Employee Prime Department Code
Student	Last	Name	Student Last Name
Student	Cumulative	Hours	Student Cumulative Hours

- ◆ **Physical data element names –**
 - be of the format *prime word (required) – qualifier (optional) – class (required)*,
 - use approved two or three character abbreviations (*see attached list*),
 - contain syntax mandated by database or operating software.¹⁰

Logical Name	Oracle	COBOL
Account Balance	ACCT_BAL	ACCT-BAL
Employee Salary Amount	EMP_SAL_AMT	EMP-SAL-AMT
Employee Prime Department Code	EMP_PRIM_DEPT_CDE	EMP-PRIM-DEPT-CDE
Student Last Name	STU_LST_NAME	STU-LST-NAME
Student Cumulative Hours	STU_CUM_HRS	STU-CUM-HRS

- ◆ **Approved abbreviations should –**
 - make sense to users,
 - as a general rule, only be created for words with five or more characters in length,
 - be associated with only one abbreviated word or forms of that word (i.e. ALLOC for *allocate, allocated and allocation*),
 - not be an English word themselves (i.e. do not use CLASS for *classification*),
 - not be confused with other standards abbreviations (i.e. TX is Texas, not for *taxable*),
 - start with the letter of the word being abbreviated (i.e. use EXCPT not XCPT for *exception*),
 - if one abbreviation is used within a name, then all words should be abbreviated if possible within the above constraints (i.e. UNIV_ID_NBR is better than UNIVERSITY_ID_NUMBER.¹¹ (*See attached list of Abbreviations.*))

⁹ University of California, Davis, "Information Naming Standards," (July 22, 1999), http://www.dataadmin.ucdavis.edu/documents/Information_Naming_Standards_Ver1_2/Information_Naming_Standards_Ver1_2_pdf_link.html (accessed June 2, 2003).

¹⁰ Brooks and Smutek, 5-6.

¹¹ University of Illinois Data Management.