Guidance for Data Dictionaries

|  |  |
| --- | --- |
|  |  |
| Data Dictionary Defined | A data dictionary is a shared repository that defines the meaning, composition, allows values, and other information for all data elements used in an application. Begin collecting data definitions as you encounter them during requirements elicitation. These definitions will eventually feed into database schema and program variables during design and construction. The data dictionary can be stored as an appendix to the SRS or as a separate document. |
|  |  |
|  |  |
| Setting Up a Data Dictionary | A separate data dictionary makes it easy to find the data information readers need, and it avoids redundancy and inconsistencies. Some analysis and design tools include a data dictionary component.  If you set up the data dictionary manually, consider using hypertext to facilitate a user’s navigation throughout the data dictionary. Clicking on a data item that is part of a data structure would jump to the definition of that individual data item. In the sample data dictionary fragment at the end of this document, you can click on “Quantity” in the Requested Chemical entry to jump to the definition of the Quantity data element. Organize the data dictionary entries alphabetically. |
|  |  |
|  |  |
| Data Dictionary Template | The table format shown in the [Data Dictionary Template](#data_dictionary_template) below is a convenient way to record data definitions. The notations used to define items in the data dictionary are described in the following sections. The last page of this document illustrates a fragment of a data dictionary using that template. |
|  |  |
|  |  |
| Primitive Data Elements | A primitive data element is one for which no further decomposition is logical or necessary. The “Composition or Data Type” column in the template describes the data type for each primitive. If the primitive can take on a limited number of discrete values, list them in the “Values” column. |
|  |  |
|  |  |
| Data Structure | A data structure, or record, is composed of multiple data items, separated by a plus sign. In the sample data dictionary fragment, Requested Chemical is a structure made up of six data elements. Leave the “Length” and “Values” columns in the template blank for data structures.  If an element in the data structure is optional, enclose it in parentheses. The Vendor data element is optional in the Requested Chemical structure; that data value could be omitted in a given record.  Elements within the data structure definition could be literal text strings, enclosed in quotation marks, such as characters to be used when formatting a multipart data object like a telephone number. Structures can incorporate other structures: the Requester data structure includes the Delivery Location structure. Each data item that appears in a structure must itself be defined in the data dictionary. |
|  |  |
|  |  |
| Repeating Group | If multiple instances of a data item can appear in a data structure, enclose that item in curly braces. This is a “repeating group.” Show the allowed number of possible repeats in the form *min:max* in front of the opening brace. If the maximum number of instances is unlimited, use “n” to indicate this. For example, “3:n{something}” means that the data object must contain at least 3 instances of “something” and there is no upper limit on the number of instances of that “something” that can be present.  In the sample data dictionary fragment below, a Chemical Request must contain at least 1 requested chemical but may not contain more than 10 chemicals. Each request has the additional attributes of a unique identifier, the date the request was created, the person who requested it, and a company charge number to be billed. |
|  |  |

**Data Dictionary Template:**

Data Dictionary for <Project>

*Last Updated <date> by <updater>*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data Element | Description | Composition or Data Type | Length | Values |
| *<name of the data item being defined>* | *<textual description of the business meaning of the data element>* | *<for primitive data elements: state the data type (integer, floating point, alphabetic, date, etc.) and, as appropriate, format (e.g., date as MM/DD/YYYY)*  *for data structures show the components that comprise the structure>* | *<maximum number of characters for primitives; blank for structures>* | *<list of allowed values, default, rules governing legal values, and any other description of the data values>* |

**Sample Data Dictionary Fragment:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Data Element | Description | Composition or Data Type | Length | Values |
| Chemical Request | request for a new chemical from either the Chemical Stockroom or a vendor | Request ID  + Requester  + Request Date  + Charge Number  + 1:10{Requested Chemical} |  |  |
| Delivery Location | the place to which requested chemicals are to be delivered | Building  + Lab Number  + Lab Partition |  |  |
| Number of Containers | number of containers of a given chemical and size being requested | integer | 3 | default value is 1 |
| Quantity | amount of chemical in the requested container | numeric | 4 | no default value |
| Quantity Units | units associated with the quantity of chemical requested | alphabetic characters | 10 | grams, kilograms, milligrams, each; no default value |
| Request Date | date the chemical request was placed | MM/DD/YYYY | 10 |  |
| Request ID | unique identifier for a request | integer | 8 | system-generated sequential integer, beginning with 1 |
| Requested Chemical | description of the chemical being requested | Chemical ID  + Number of Containers  + Grade  + [Quantity](#quantity)  + Quantity Units  + (Vendor) |  |  |
| Requester | information about the individual who placed a chemical request | Requester Name  + Employee Number  + Department  + Delivery Location |  |  |
| Requester Name | name of the employee who submitted the request | alphabetic characters | 40 | may contain blanks, hyphens, periods, apostrophes |