

SPASE Parameter Key Formation for Plain Text Data

May 12, 2011

Contents

Overview.....	1
Applicability	1
Parameter Key Formation	1
History	2

Overview

For some forms of data the most convenient format is as plain text. Often the data has a fixed field layout, but more commonly the data is provided with delimited fields, typically using a comma separated format. The SPASE data model assumes that all data resources are stored in formats which assign names to parameters (or fields) in the data. For data stored in a plain text format this is not true. While there is an existing convention that the first line of a plain text file contain the name assigned to each field, this has not been universally adopted. In order to provide a simple, non-intrusive and universally applicable solution to describing simple tabular plain text data SPASE has defined a Parameter Key Formation Rule for Plain Text Data.

Applicability

SPASE metadata describing a parameter in a data file is linked to the data through the use of the "ParameterKey" element. The value of the "ParameterKey" can be used by data access software to retrieve the parameter. Typically plain text data is read using simple line reading methods and then parsed to retrieve field values from each record. The SPASE Parameter Key formation rules are designed to provide the information necessary to parse simple tabular plain text files which have a homogenous record structure. However, some plain text formats are sufficiently complex that other methods, such as an external format description file or specialized readers, might be necessary.

Parameter Key Formation

Simple tabular plain text file consists of rows and columns. Each row has the same number of columns. Columns can be separated by spaces, tabs or commas. The parameter key consists of the fixed string "column_" with the suffix of the column index. The first column is designated as column 1.

Example: The second column will have the parameter key "column_2".

Some parameters can span multiple, contiguous columns. The parameter key for such columns is the designation for the first column in the span, followed by a dash ("-"), then the last column in the span.

Example: A parameter which spans column 3 to column 6 will have a parameter key "column_3-column_6".

History

Original: March 9, 2011 – Todd King

Simplified: April 7, 2011 – Todd King, per SPASE Working group request.

Augmented: May 11, 2011 – Todd King, to accommodate multi-column spans.