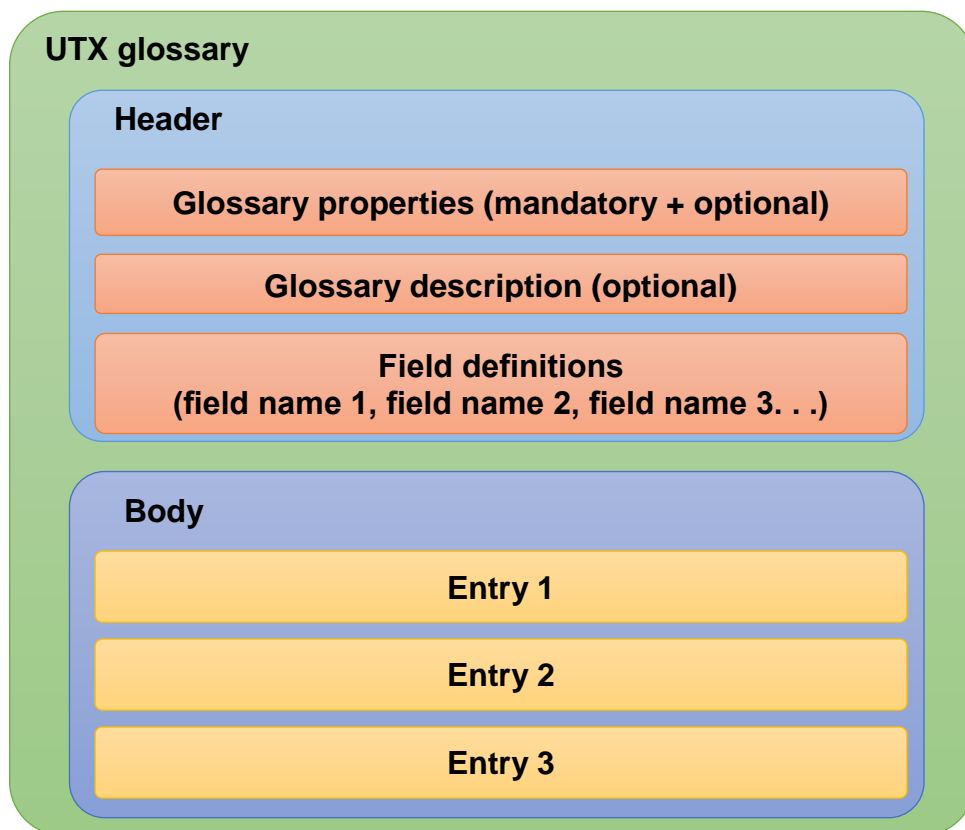# UTX 1.20 Minimal Specification

Table of Contents

0. About this specificaion

This specification explains the minimal information to create a UTX file. For more information, please refer to the full specification.

1. The structure of a UTX file

The following diagram illustrates the structure of a UTX glossary. For optional elements, refer to the full specification.

UTX glossary

Header

Glossary properties (mandatory + optional)

Glossary description (optional)

Field definitions
(field name 1, field name 2, field name 3. . .)

Body

Entry 1

Entry 2

Entry 3

1.1 Core UTX example

The following is an example of a complete UTX glossary with the minimum required information.

#UTX 1.20	
#term:en	term:ja
test	テスト

1.2 File format details

File structure: tab-delimited text

File extension: ".utx"

Character encoding: UTF-8 with BOM.

New line code: "\r\n" (CR+LF). Blank lines are not allowed.

Line comment symbol: # (hash symbol).

2. Header

2.1　UTX header structure

All lines in a UTX header begin with "#," meaning that they are treated as commented-out lines.

A UTX header includes at least two mandatory lines, namely, glossary properties and field definitions. The glossary properties can have two or more lines.

One or more lines can be added between the two mandatory lines for a description of the glossary.

The UTX version is the only mandatory glossary property. It indicates the UTX version of the glossary. It is specified with "UTX", a space, and the version number "1.20."

2.2　Field definitions

The last line of the UTX header (also begins with "#") includes a set of field (or column) definitions.

3. Field definitions and body

The body of a UTX glossary consists of entries in each line. Within a line, elements ("cells" in a spreadsheet) are separated by tab characters.

3.1　Language tags for fields

Use IETF's BCP 47 language tags to indicate languages in any part of a UTX glossary. These are the same as the language tags used in HTML and XML, such as ja, fr, en-US, zh-CN, and zh-TW. See <http://www.w3.org/International/articles/language-tags/Overview.en.php>.

Fields that are specific to a particular language are indicated by the respective language tag. Language tags do not have a space after the colon.

3.2　`term` fields

The term field indicates a field of a term. The languages of these fields are indicated by language tags. See the following table for examples.

Field name	Syntax/field item/field value	Example	
`term`	`term:<language>`	`term:en`	`term:ja`

3.3　Other fields

The full UTX specification includes a complete list of other optional pre-defined fields.

Other than these pre-defined fields, any number of user-defined fields and their field items can be added to a UTX glossary.

For a language-specific field, use a language tag to indicate the language.

Syntax	Example
User-defined field:<language>	`x-termUsage:en`