

# UTX, a simple glossary format

## What is UTX glossary format?

UTX is a simple glossary format established by AAMT

Rules of how to make glossaries for technical terms

Tab-delimited format that can be edited in Excel etc.

## Advantages of UTX

You can easily compile and manage glossaries containing technical terms

Common format makes it easy to share and reuse terminology data

Improves the efficiency and accuracy of both human translation and machine translation

## UTX glossary sample

Information about the glossary (creation date, license, etc.)			
#UTX 1.11; en-US/zh-CH; 2014-09-25; copyright: AAMT (2012); license: CC BY 3.0			
#src	tgt	src:pos	term status
Asia-Pacific Association for Machine Translation	亚洲太平洋机器翻译协会	properNoun	approved
contributor	用语提交者	noun	provisional
optional	省略可能	adjective	approved
optional	可选	adjective	forbidden
provisional term	临时词	noun	approved
provisional word	临时词	noun	non-standard
unidirectional	一方向	adjective	approved
monodirectional	一方向	adjective	non-standard
Source term (American English)	Target term (Chinese)	Part of speech	Term status

Term status provides **reliability**

## Introduction

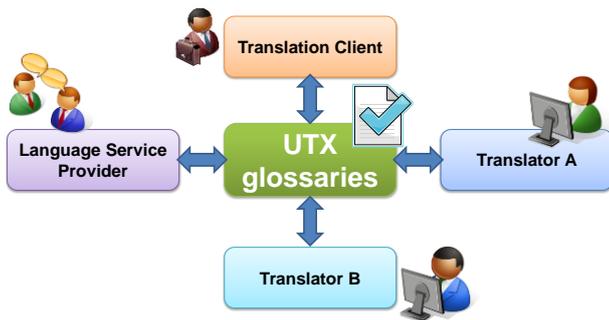
UTX (universal terminology eXchange) is a simple, tab-delimited terminology format, which is established by AAMT (Asia-Pacific Association for Machine Translation). A glossary created in the UTX format is useful for translators as it is, but it can also be easily converted into various formats and used for a variety of translation tools. For example, UTX glossaries can greatly improve the accuracy of translation software by providing terminological data. With the UTX format, any user can **quickly create a simple, reusable, easy-to-use glossary**.

AAMT is an organization consisting of researchers, developers, manufacturers, and users of machine translation (MT).

## Why do you need glossaries for translation?

In specialized translations for enterprises, universities, or organizations, technical terms cannot be translated correctly if there is no glossary. If a glossary is created in an early stage, translators don't have to research and check the same technical terms, and it can greatly reduce the cost and effort of translation. A glossary is also essential for term consistency and proofing. In short, a glossary is indispensable to perform an accurate, quality translation.

## UTX helps sharing and reusing of glossaries



### Why UTX?

With UTX, a user can easily create, share, and reuse glossaries to improve the translation quality. Have you ever thought that translation software can produce only strange translations? The problem is often that it doesn't have sufficient glossary data of certain words and phrases. You can greatly improve the accuracy of translation software (machine translation) by accumulating translation knowledge as a UTX glossary, and then converting it into a user dictionary of the translation software.

An Excel glossary or a simple plain text file is difficult to share or to reuse because their format is not standardized. Many glossaries are available on the Internet, but their formats are not readily usable. Time-consuming corrections and fine-tuning are required before you can use them in translation tools. However, if you use a standardized format such as UTX, you can share a glossary among various tools, and quickly reuse it. UTX can also be converted to/from different glossary formats.

### Who creates and uses UTX?

UTX is specifically designed to be created and used by translators and end users of translation software. It can be created from minimum data such as source and target terms. You can always add extra information, if it is required.

### In which domains?

UTX can be used in any specific domain that has technical terms, such as ICT, medicine, legal, engineering translation.

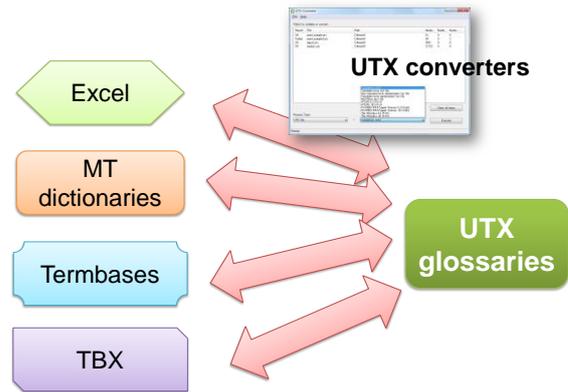
### What kind of words should we include?

A UTX glossary contains only **technical terms of specific domains**, such as names of products, parts, diseases, medicines, and laws. It also contains **proper nouns**, such as names of people, places, and facilities. General words should not be included unless they have specific meaning in a particular domain.

### In what kind of scenarios do we use UTX?

1. When you create a glossary from scratch.
2. When you collect target terms proposed in a translation project by multiple translators.
3. When you need an intermediate conversion format for the conversion between various terminological formats.

## Conversion to/from UTX



### How do we make and use a UTX glossary?

A UTX glossary can be easily created, edited, and viewed with any spreadsheet applications or text editors. In tools such as OmegaT (a translation memory tool) and ApSIC Xbench (terminology reference tool), UTX can be used with only a few modifications.

In addition, UTX can be used with various tools by converting to various formats. Converters such as the official UTX Converter and Glossary Converter are available. With the official UTX Converter, the conversion is currently possible between UTX and ATLAS, The Honyaku, and Transer series (such as PC-Transer).

More info: <http://www.aamt.info/english/utx/tools.htm>

### Where is it used?

Aside from the uses in global companies and organizations, the Japan Patent Office created UTX glossary data with 2.2 million entries.

### More information

The UTX specification, sample glossaries, and the converter are available for free. Based on the specification, anyone can create, publish, and share UTX glossaries. If you are an organization or company user and wish to create a UTX glossary (or terminology data), please contact us. We might be able to give you some advice.

**Sharing/Standardization Working Group (UTX team),  
MT Research Committee,  
AAMT (Asia-Pacific Association for Machine Translation)**

<http://www.aamt.info/english/utx/>

Contact: [aamt-info@aamt.info](mailto:aamt-info@aamt.info)

### UTX team members (in no particular order)

YAMAMOTO Yuji (leader)	CosmosHouse
MURATA Toshiki	Oki Electric Industry Co., Ltd.
Francis Bond	Nanyang Technological University
OKURA Seiji	Fujitsu Laboratories Limited
Michael Konin Kato	Japanese Greats Co., Ltd.
AKIMOTO Kei	Kotobaya Inc.
TAKAHASHI Hiroyuki	Cross Language Inc.
KAMEYA Hiroshi	SunFlare Co., Ltd.