

## UTX の利点

- 用語集を共有・再利用できる
- 翻訳ソフトの精度を大きく向上
- 単語を調べる時間と労力を削減

## 実例

(IT分野の用語集サンプル)

#src	tgt	src:pos	term status	src:plural
early adopter	アーリー アドプター	noun	approved	early adopters
fast	高速な	adjective	provisional	
optional	省略可能な	adjective	approved	
optional	オプションな	adjective	forbidden	
save	保存する	verb	approved	

#はコメント行を表す。

●1行目(コメント行):用語集の基本情報。各項目はセミコロンと半角スペースで区切る。

#UTX <バージョン番号>; <起点言語>/<目標言語>; <最終更新日付時間>; <著作権者名(権利発生前)>; <license>; <ライセンス>; <追加の情報>(必要に応じて追加できる)

●2行目(コメント行):属性の規定。各項目はタブ区切り。上記の例では

#<原語> <訳語> <原語品詞> <用語ステータス> <原語複数形>

●3行目以降が、実際の項目。各項目はタブ区切り。

## UTX 用語集作りのポイント

- 1つの用語集は1つの特定分野にする。
- 1項目には1つの原語。
- 訳語は、最良のものを1つだけ選ぶ。
- 固有名詞以外は大きく始めない。
- 原形を記載する(市販辞書の見出しの形式)。
- 原語、訳語以外の情報はコメント欄に記す。

詳しくはクイックガイドと仕様書も参照してください。

## 免責事項

UTX、UTX-Simple、UTX-XMLの仕様(以下、これらを総称して「UTX仕様」という)またはUTX仕様に基づいて作成された辞書(以下、「UTX辞書」という)を利用した場合は、以下の事項に同意したものとみなされます。本事項のいずれかが無効または強制不能とされた場合、そのことはいかなる意味でも他の条項の有効性または強制可能性に影響を与えないものとします。

### 1. AAMTおよびAAMT参加者からUTX辞書および関連ツールの作成者へ

(1)UTX仕様は公開されており、どなたでもご使用頂けます。ただし、AAMTおよびAAMT参加者はUTX仕様に関する権利を放棄しておらず、どなたであってもUTX仕様を改変して公開することはできません。

(2)UTX辞書の作成に際して、AAMTおよびAAMT参加者はUTX仕様を現状有姿のまま提供するものであり、UTX仕様に関する一切の事柄を保証しません。UTX仕様およびUTX辞書は、UTX辞書の作成者各位の責任においてご使用ください。

(3)AAMTおよびAAMT参加者は、UTX辞書の作成者がUTX仕様やUTX辞書を使用した結果(権利侵害の有無・訳語の正確性・妥当性・品質を含むがこれに限らない)に関して、一切の責任を負いません。

(4)AAMTおよびAAMT参加者は、UTX辞書の作成者が作成した辞書の著作権の正当性について確認をせず、保証しません。従って、UTX辞書の作成者が当該UTX辞書に関するデータについて適切な著作権を保有していない場合、法的な問題が発生しても、UTX辞書の作成者の責任となりません。

(5)AAMTおよびAAMT参加者は、UTX辞書の作成者に、適切な著作権を行使できる場合に限り、商業使用を含め、UTX辞書の使用者へのUTX辞書のライセンス条件を定めることを認めます。ただし、UTX辞書の作成者は、UTX辞書の基盤となるデータの著作権について、個別にデータの提供元に確認する義務があります。

(6)AAMTおよびAAMT参加者は、UTX辞書に関する各種ツールの作成者に対して、当該ツールの使用結果についてなんらの保証もしません。

### 2. UTX辞書の作成者からUTX辞書の使用者へ

UTX辞書の使用者は、UTX辞書を、UTX辞書の作成者が定めるライセンス条件に応じて使用できません。UTX辞書のライセンスは辞書によってそれぞれ異なりますので、UTX辞書を構成するファイルの先頭部分に含まれるライセンス条件をご確認ください。

### 3. AAMTおよびAAMT参加者からUTX辞書の使用者へ

AAMTおよびAAMT参加者は、AAMTおよびAAMT参加者は、UTX辞書の使用者がUTX仕様やUTX辞書を使用した結果(権利侵害の有無・訳語の正確性・妥当性・品質を含むがこれに限らない)に関して、一切の責任を負いません。UTX辞書の作成者との間で解決をお願いします。

## AAMT (アジア太平洋機械翻訳協会) 機械翻訳課題調査委員会

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

問い合わせ先: [aamt-info@aamt.info](mailto:aamt-info@aamt.info)

### 共有化・標準化ワーキンググループ メンバー(順不同)

山本 ゆうじ(リーダー)	秋桜舎
伊藤 肇	株式会社インターグループ
村田 稔樹	沖電気工業株式会社
Francis Bond	南洋理工大学(シンガポール)
島津 美和子	東芝ソリューション株式会社
大倉 清司	株式会社富士通研究所
加藤マイケル孝仁	ラーニング コンサルタント
秋元 圭	株式会社クロスランゲージ
目次 由美子	株式会社 シュタール ジャパン

2013年4月版

## AAMT (アジア太平洋機械翻訳協会) 機械翻訳課題調査委員会

## 用語集形式

# UTX

## Universal Terminology eXchange



「翻訳ソフトは役に立たない」  
そう思っていないませんか？

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

# 用語集形式UTXとは

## 概要

UTX(Universal Terminology eXchange)とは、アジア太平洋機械翻訳協会(AAMT)が策定した、シンプルなタブ区切りの用語集形式です。AAMTは、機械翻訳の研究開発者、製造販売者、利用者の三者から構成される団体です(機械翻訳は、翻訳ソフトの核となる技術です)。メンバーはボランティアです。2009年に、翻訳ソフトのユーザー辞書の標準化を目的として最初のUTX仕様が策定され、以後、改訂が重ねられています。この仕様に基づいて、どなたでも自由にUTX用語集(「UTX辞書」と呼ぶこともあります)を作成・公開・共有できます。

## 特徴

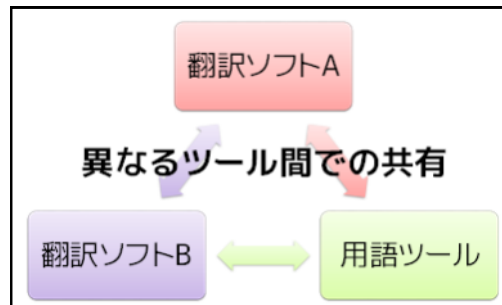
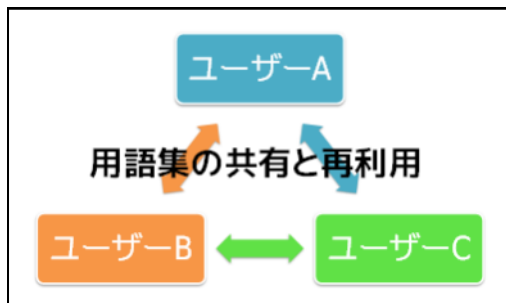
UTXは、**専門用語を共通規格の対訳用語集として共有しやすくするしくみです**。人間翻訳の品質はもちろん、翻訳ソフトの翻訳精度を大きく向上できます。ユーザー視点から、シンプルで**作りやすく、使いやすい**ことを目指しています。UTX用語集は、そのまま人間翻訳での用語集として活用でき、さらに各種形式に変換して、さまざまな翻訳ソフトや翻訳支援ツールで使えます。

## なぜ翻訳で用語集が必要か

企業・大学・組織での専門的な翻訳では、用語集がないと、専門用語が正しく訳されません。初期段階で用語集をしっかり作れば、同じ言葉を何度も調べずすみ、翻訳の費用と労力を減らせます。用語集は訳語統一や誤訳チェックにも役立ちます。高品質な翻訳を正確に行うには、用語集は必須です。

## なぜ UTX を使うのか

UTXを使えば、用語集を簡単に作成・共有・再利用して、翻訳の質を向上できます。「翻訳ソフトは変な訳ばかり出す」と思われていませんか？ 翻訳ソフトがうまく訳せないのは、語句をどう訳すべきかという**翻訳知識が不足している**からです。まず、翻訳知識をUTX用語集として作り、それをユーザー辞書に変換することで、翻訳ソフトの翻訳精度を大きく改善できます。



また、これまで翻訳ソフトのユーザーは、細切れのユーザー辞書を個人で作成しても、有効活用できませんでした。Excelやテキスト ファイルでも、各項目の形式が共通化されていないと共有や再利用は困難です。ウェブサイトでもさまざまな用語集が公開されていますが、実際にはすぐに活用できず、手間のかかる修正と調整が必要です。しかし、UTXのような標準規格を使えば、さまざまなツールで用語集を広く共有し、すばやく再利用できるようになります。

## だれが作り、使うのか

主に翻訳者や翻訳ソフトのユーザーが、作り、使うことを想定しています。作成や使用にあたって、文法や言語学などの高度な専門知識は不要です。UTX用語集は、複数形、名詞や動詞など品詞の区別など、最低限の情報のみで作れます。

## どのような分野で使うのか

IT、医療、法律、工学など、専門用語が多い翻訳であれば、どのような分野でも使えます。

## どんな語を含むのか

UTX用語集は、製品・部品名、病名、薬品名、法律名など、**特定分野の専門用語**や、人名、地名、施設名などの**固有名詞**のみを含みます。多くの場合は名詞、特に複合名詞がほとんどです。たとえば「XML declaration」のような語は、ユーザー辞書に登録することではじめて「XML宣言」などと正しく訳せます。「window=窓」のような基本的な語彙は、翻訳ソフトのシステム辞書に最初から含まれているため含めません。翻訳ソフトの購入時に付属しない、きめ細かい対訳の情報を集約して共有・再利用することにより、翻訳精度を向上できます。

文(センテンス)は、一種の「単語」として扱うのが適切な場合にのみ含めることができます。ただし、原則としてUTXは、翻訳メモリー(文単位の対訳データベース)とは区別して使います。

## 多言語対応と用語管理

UTXの文字コードは、Unicodeなので、ほとんどの言語を扱えます。基本的には、1つのUTX用語集に、起点言語(原文)から目標言語(訳文)への方向の訳が含まれます。UTX 1.20から、複数の目標言語を指定できるようになりました。またUTXでは、**用語の品質管理ができ、正**

**しい用語を確実に使用できます**。各用語には、「暫定」「禁止」「承認」「非標準」(provisional, forbidden, approved, non-standard)のいずれかの用語ステータスを指定できます。複数ユーザーが用語の情報提供をするとき、当初の用語ステータスは「暫定」(または空白)です。その後、用語管理者が、各用語を確認し、適切であれば用語ステータスを「承認」にします。「承認」された語は逆方向の翻訳でも使用できます。また「禁止」ステータスを付けて、特定の用語の使用を禁止することもできます。最良の訳語ではないが、処理上含める必要がある語は「非標準」ステータスになります。

## どのようなシナリオで作るのか

1. 用語集をゼロから作成する
2. 翻訳作業で発生する訳語の用語を収集する
3. 各種用語データ間での変換で、中間変換形式として

## どうやって作るのか

UTX用語集は、Excelなどの**表計算ソフト**や**テキスト エディター**で簡単に作成・編集・表示できます。開発中のものを含め、各種形式とUTX形式の相互変換を行うツールもあります。

## どうやって使うのか

UTX用語集は、シンプルな形式なので、簡単に変換して各種ツールにインポートできます。翻訳メモリー ツール OmegaT、用語ツール ApSIC Xbenchなどのツールではほぼそのまま使えます。

## 用語集作成に費用は掛かるか

UTXの仕様書は無料でダウンロードできます。公開されているUTX用語集が無料で利用できるかは、作成者の使用許諾に基づきますが、多くは無料です。

## オープンソース開発者の方へ UTX 形式で用語を公開・共有しませんか？

UI文字列や専門用語を、UTX形式にして、相互に公開・共有することで、ソフトウェアやマニュアルをすばやく正確に多言語化し、世界中の人に使ってもらうことができます。

## UTX メーリング リスト

どなたでもUTXに関する議論にご参加いただけます。「UTXメーリングリスト」でウェブを検索してください。

## もっと詳しく知りたい

<<http://www.aamt.info/japanese/utx/>>から、UTXの仕様書、サンプル用語集、UTXクイック ガイドが無料でダウンロードできます。