



# MTユーザーガイド

機械翻訳で失敗しないための手引き

アジア太平洋機械翻訳協会 AAMT  
MTユーザーガイド委員会  
2022年 Ver. 1.0



<b>MTユーザーガイド</b> .....	3
MTリテラシー .....	3
本書の目的 .....	4
対象読者 .....	4
<b>第1章 技術的要素</b> .....	5
1.1 機械翻訳とは .....	5
1.2 MTの翻訳品質の問題 .....	7
1.3 MTの翻訳品質の評価 .....	8
1.4 MTサービス .....	8
1.5 MT以外の自然言語処理技術 .....	9
1.6 おすすめ文献 .....	9
<b>第2章 翻訳品質</b> .....	10
2.1 誤訳と翻訳品質 .....	10
2.2 正確性 .....	11
2.3 流暢性 .....	12
2.4 用語・専門用語 .....	13
2.5 スタイル .....	14
2.6 JTF翻訳品質評価ガイドライン .....	15
2.7 おすすめ文献 .....	15
<b>第3章 翻訳制作プロセス</b> .....	16
3.1 人手翻訳の制作プロセス .....	16
3.2 制作前プロセスと合意形成 .....	17
3.3 制作プロセス .....	18
3.4 ポストエディット (PE) .....	19
3.5 フルPEとライトPE .....	19
3.6 おすすめ文献 .....	21
<b>第4章 法律</b> .....	22
4.1 MT利用者に関して発生する法律問題 .....	22
4.2 他者から翻訳の依頼を受けてMTを利用する者 .....	23
4.3 MT提供者に関して発生する法律問題 .....	25

# MTユーザーガイド

## ■ MTリテラシー

グローバル化が進み異なる言語の話者間のコミュニケーションの必要性が高まる一方で、AI技術が発達したことにより、機械翻訳(Machine Translation = MT、以下MTと言う)が身近に利用できるようになりました。MTは、瞬時に翻訳を提供してくれます。これにより、私たちは外国語を理解したり、外国語で情報発信したりできます。

しかし、MTには落とし穴があります。

MTの翻訳は完璧ではありません。思わぬ誤訳や訳抜けを起こします。このような翻訳の誤りが、社会問題に発展したケースもあります。たとえば、「ワクチン接種は義務ではありません」という意味の原文が、「ワクチン接種は必要ではありません」と誤訳されていたことがありました<sup>注1</sup>。これはMTの技術的な実力不足による問題ではなく、MTを使う人間の問題です。また勝手に他人の文書を翻訳すると、情報漏えいや著作権違反になることもあります。



これらの問題の原因は、私たちのMTに対する不理解と、「翻訳」と「ことば」についての理解不足にほかなりません。このような理解不足を克服しMTを正しく利用する知識を「MTリテラシー」と本書では呼びます。MTを正しく使い、外国語でのコミュニケーションで失敗しないためのMTリテラシーを読者に提供することが本書の目的です。



## ■ 本書の目的

本ユーザーガイドの目的は、MTを正しく使うために必要な知識(MTリテラシー)を提供することです。MTを正しく使うためには、まず「翻訳とは何か」についての知識が基礎となります。そのため、本書ではMTだけでなく、人間が行う「翻訳」についても説明します。

リテラシーが問われる項目は、大きく分けて次の4つです。

- 技術の知識: MTの技術的概要を理解する
- 翻訳品質の知識: 翻訳品質の評価方法を理解する
- 翻訳制作プロセスの知識: 正しい翻訳が作られる工程を理解する
- 法的知識: MTの利用をめぐる秘密保持、責任、著作権を理解する



## ■ 対象読者

MTユーザーガイドの対象読者は、MTを利用するすべてのユーザーです。その中でも特に、MTを利用して情報を発信しようとする人、および(MTを利用した)人手翻訳(ポストエディット)を翻訳会社に依頼しようとする人には必読です。MTを利用するユーザーである限りにおいては、MT開発者、翻訳発注者、翻訳会社、プロ翻訳者、教育者、研究者にも役立つ情報を提供します。

「MTを利用して情報を発信する」ことについて知ることは、「自らの情報収集のためにMTを利用する」時にも注意すべき点をも包含します。本書では、最終的にできあがる翻訳文書の品質を担保し、MTに特有のリスクを回避するために留意すべき点をまとめました。

注1: <https://multilingual.com/google-translate-causes-vaccine-mishap/>

# 第1章 技術的要素

## 1.1 機械翻訳とは

機械翻訳(Machine Translation; MT)とは、翻訳の原文の単語列を入力とし、訳文の単語列を出力する技術(あるいはアルゴリズム、システム)のことです。

1947年、翻訳は暗号解読の要領で機械的に実現できるのではないか、という発想をもとにMTに関する研究が始められ、これまでに様々な手法が提案されてきました。1990年代以降は、翻訳の事例を集めたデータ(対訳コーパス)に基づいてMTを実現する手法が主流です。ただし、次の図に示すように、2014年に提案されたニューラル機械翻訳(Neural MT; NMT)とそれ以前の非NMTでは、翻訳の実現方法が大きく異なります。

次の図1-1と図1-2をご覧ください。

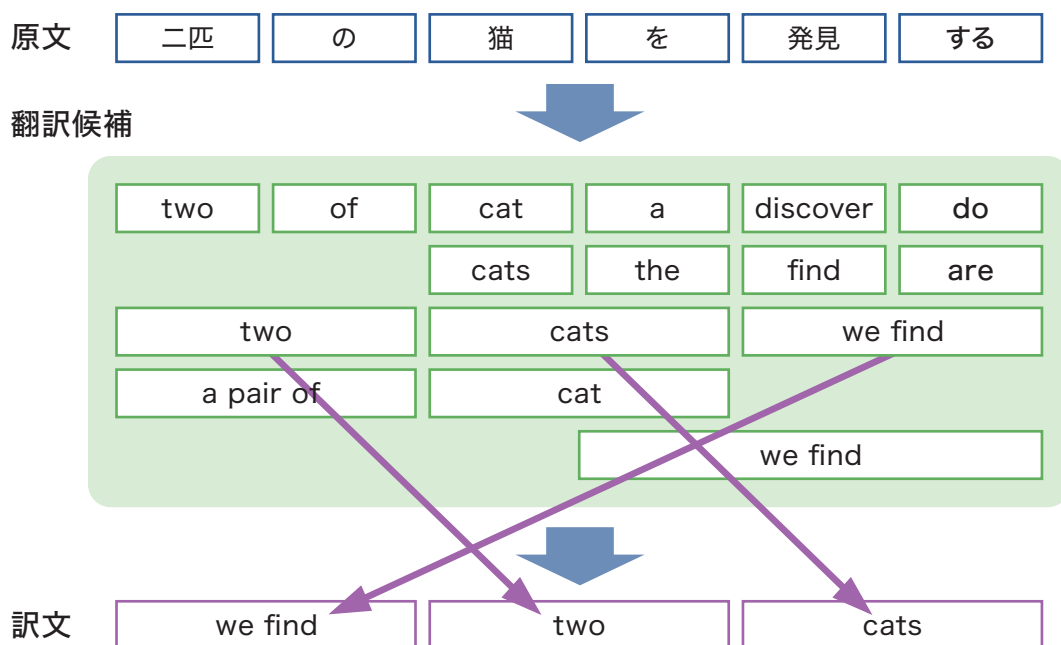


図 1-1 : 非NMT

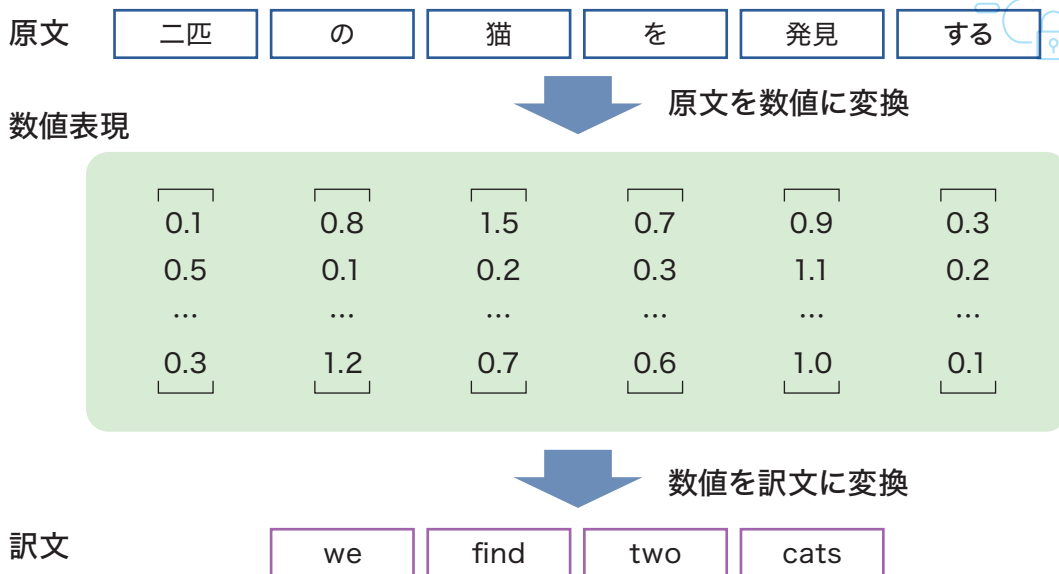


図 1-2 : NMT

非NMTでは、「二匹の猫 / two cats」「二冊の本 / two books」のような原文中の表現と訳文中の表現の対を、対訳コーパスからあらかじめ抽出しておきます。訳出の際には、原文にマッチするそのような対をパズルのように組み合わせて訳文を生成していました。

これに対してNMTは、原文を一旦たくさんの数値に変換し、そこから訳文を生成します。原文の表現を数値の束でソフトに表すとともに、文中の語句の間の関係を参照することによって、より正確で流暢な訳文を生成しようとしています。構造上は、人間が翻訳する際に行っている概念化/脱言語化(deverbalization)に類似していますが、この数値が実際に何を表すのかは明らかになっていません。

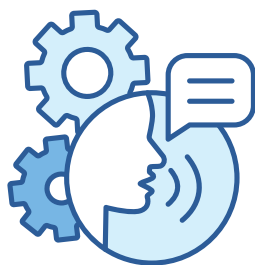
## 1.2 MTの翻訳品質の問題

翻訳の品質評価基準の詳細は2章をご覧ください。ここでは、MTの典型的な誤りをいくつか紹介します。

非NMTでは、原文のどの部分がどう訳されたかは明示されており、訳抜けや過剰生成などの問題は生じにくかったのですが、「take」のように意味や用法が多い表現の訳し分けが苦手で、訳文が流暢ではないという特徴がありました。一方、NMTでは、曖昧な語句の解釈や訳文の流暢性が非NMTと比べて改善されています。しかし、数値表現を経由しているために、訳抜けや余分などの原文との対応上の問題や、訳出途中で一旦誤るとその後ずっと誤り続けるという問題も生じます。

そもそも対訳コーパスから学習できるのは、対訳コーパス中の翻訳の事例という形で「明示的に」与えられている表現間の関係のみですので、うまくいったとしても、MTにできることは言語表現として等価な文の生成に限られます。対訳の事例を学習時に明示的に与えて学習していない限り、書き手や話し手の属性、文脈、翻訳の目的、翻訳成果物の消費者、対象分野の用語集やスタイルなどは考慮されません。たとえば、エスカレーターの乗降時にアナウンスされる「Watch your steps」は、人間であれば「足元にご注意ください」と訳すことができます。一方MTは、このように訳された事例を文脈とともに学習していない限り、「あなたのステップを見てください」としか訳すことができません。対象とする言語対・分野の対訳コーパスが小さくて十分な学習ができない場合も、適切な翻訳結果を得られる可能性が低いです。

近年登場したNMTは非NMTよりも高性能だと言われています。しかし、完璧ではありませんし、常に及第点/最低限の訳出をできるわけでもありません。MTを使用する際は、誤訳によって生じるリスクや、目的に応じて人間が手を加えることの必要性を念頭においておくことが望ましいです。





## 1.3 MTの翻訳品質の評価

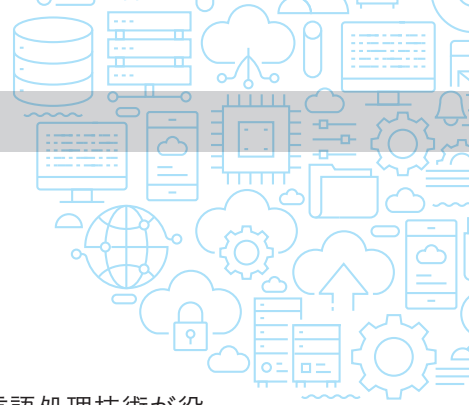
MTの品質も、人間が作成する翻訳と同様に、2章で紹介されているような品質評価基準に基づいて、翻訳成果物の用途や消費者などの翻訳することの目的に沿っているかどうかを鑑みて評価することが望ましいです。ただし、現時点では、このような評価は、人間にしか行うことができません。

開発の途中段階および研究においては、学習過程にあるMTシステムの翻訳品質を頻繁かつ高速に計測する必要があります。そのため、MTシステムの学習用の対訳コーパスとは別に少量の評価用対訳コーパスを用意し、これを用いて自動評価を行うことが一般的です。具体的には、評価用対訳コーパス中の原文をMTシステムで翻訳した結果と、訳文(人間が作成した参照訳)の「類似度」を計算して、翻訳の品質を近似する手法が用いられています。この手法も、MTと同様に十分に正確ではないものの、研究開発の推進には不可欠です。なお、訳文中の個々の語の意味や、文と文の繋がり良さなどを考慮しようとする試みもありますが、実用化には至っていません。

## 1.4 MTサービス

インターネットを通じて様々なMTサービスが利用可能ですが、特に精度が高いと言われていたものについては、大規模な対訳コーパスに基づいて実現されていると考えて間違いありません。これらのサービスでは、MTサービスのユーザが入力している文章も収集・活用している場合があります。そのため、個人情報や機密情報を含む文章などをMTサービスを使って翻訳すると、法的なリスク(4章を参照)が生じることに繋がります。





## 1.5 MT以外の自然言語処理技術

翻訳制作プロセス(たとえば3章で紹介されているもの)の中には、自然言語処理技術が役立つと考えられる工程があります。翻訳プラットフォーム上で人間による翻訳をサポートする技術として、既存の訳を再利用する翻訳メモリやバイリンガルコンコーダンスがよく用いられています。これらは、表現の完全一致に限れば、電卓による四則演算と同様に100%の精度で実行できます。

人間による翻訳やMT結果の品質を、上述したような参照訳との「類似度」とは異なる方法で自動的に評価する技術もあります。たとえば、人間による参照訳を用いずに、訳文中の誤り箇所や原文中の正しく翻訳されていない箇所を同定する品質推定(Quality Estimation; QE)技術や、訳文が所定のスタイルに従っているかを自動的にチェックする技術などがあります。ただしこれらの技術も完璧ではないので、品質を担保するためには、結局は人間による評価が欠かせません。

## 1.6 おすすめ文献

- ポイボー, T. [著]・中澤 敏明 [解説]・高橋 聡 [訳] (2020)  
『機械翻訳:歴史・技術・産業』森北出版
- Koehn, P. (2020). *Neural Machine Translation*, Cambridge University Press.
- 中澤 敏明(2017)「機械翻訳の新しいパラダイム:ニューラル機械翻訳の原理」,  
情報管理, 60巻, 5号, 299-306.  
[https://www.jstage.jst.go.jp/article/johokanri/60/5/60\\_299/\\_article/-char/ja](https://www.jstage.jst.go.jp/article/johokanri/60/5/60_299/_article/-char/ja)

## 第2章 翻訳品質

### 2.1 誤訳と翻訳品質

1章でみたように、MT（機械翻訳）は誤訳をします。技術が向上し誤訳が起きる頻度が少なくなってきたとはいえ、重大なエラーが文章に1つでも含まれていると、大問題になりかねません。

本章では誤訳の種類について解説します。というのも、「誤訳」という言葉は粒度が粗すぎて、MTの性能や危険性を正しく説明しきれないことがあるからです。誤訳を含む「翻訳品質」は、「正確性」「流暢性」「用語」「スタイル」に細分化されます。これらの内容をすこしでも知っておくことで、MTのどの品質要素が問題なのか、逆にMTが得意とするところは何なのかを、詳細なレベルで理解できるようになります。



図 2-1：翻訳品質の主4要素

#### ■ MTを使う時の注意

プロの人手翻訳者は、翻訳品質に関係する要素を考慮しながら翻訳をします。他方で、MTは、プロの翻訳者と同じように考えて翻訳をしていません。たとえ、MTの翻訳文がたまたま正しかったとしても、最終的にはプロの翻訳者のような人がMTの翻訳文を評価する必要があります。MTが生み出した翻訳文は、決して翻訳品質が保証されたものではないのです。MTの翻訳文の確認を怠った場合、誤訳責任で訴訟、不測の事態、命の危険、人間関係悪化、差別、勘違いなどが発生するおそれがあります。

## 2.2 正確性

正確性とは、原文の情報や意味が訳文に過不足なく盛り込まれているかどうかを示す指標です。正確性エラーには、誤訳、抜けと余分、未翻訳などがあります。

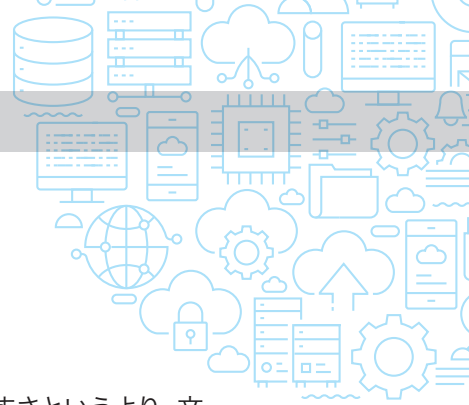
例文として、原文「This user's guide indicates the ways of sound use of machine translation」の正しい訳文が「本ユーザーガイドは、機械翻訳の健全な利用方法を示します」になるものとしたうえで、以下に正確性エラーの詳細例を示します。

- 誤訳：例、「健全な」を「音声の」と訳してしまう。
- 抜けと余分：例、「～は、世界の機械翻訳の利用方法を示します。」と訳してしまい、「健全な」が抜けていて、「世界」が余分。
- 未翻訳：例、「本ユーザーガイドは、機械翻訳のsound useを示します。」と訳してしまい、「sound use」が英語のまま残っている。

### MTを使う時の注意(正確性)

誤訳、抜けと付加、未翻訳はMTに多く見られます。MTの場合、原文に全く関係ない訳文が出現するのも特徴です(余分)。これらは、流暢性に引きずられて見落としやすいため、必ず原文と訳文を比較し、訳文が原文を過不足なく表現していることを、「人間」が確認しなければなりません。確認を行うためには、原文と訳文の両方を比較する必要があります。





## 2.3 流暢性

流暢性とは、訳文が整っているかどうかを示す指標です。主観的な読みやすさというより、文法など形式面に注目します (well-formedness)。主に、誤字脱字などが、流暢性に関係しません。以下に一例を示します。

- 誤字：例、「ガイド」を「カイド」と訳してしまう。
- 文法誤り：例、「東京で食べる」を「東京に食べる」と訳してしまう。
- コロケーション誤り：例、「約束を保つ」と訳してしまう。
- 待遇表現誤り：例、謙譲語を用いるのが適切なのに使っていない。
- 曖昧性の問題：例、「若い先生の犬」。

上の例のように流暢性の確認は、訳文を読むだけで気付くことができます。しかし、流暢性エラーがないからといって、正確性エラーが無いとは言い切れないのが落とし穴です。たとえば、「ヘリウムは今、利用できません」という訳文には、流暢性エラーがないので一見正しく読めます。しかし、原文が「He is not available now」(彼は今、手がはなせません)だとすると、最初の訳文は致命的な正確性エラー(誤訳)を犯していることになります。

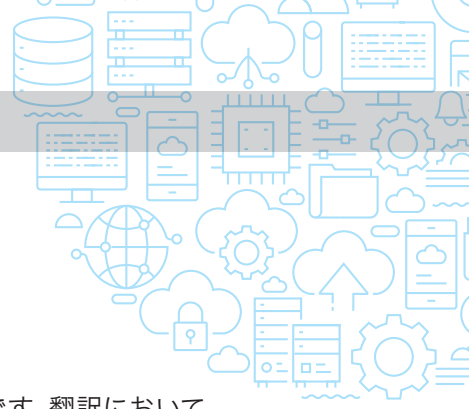
### ■ MTを使う時の注意

「待遇表現誤り」や「曖昧性の問題」はMTに多いですが、一般的にはMTの流暢性の精度は高いと言われます。逆に、その他の流暢性エラーは人手翻訳の方が多い傾向にあります。

MTの訳文について、流暢性エラーがあると、読み手に「プロフェッショナルでない」という印象を与えてしまい、信頼性が下がるおそれがあります。「待遇表現誤り」があると、失礼な言葉遣いとなり、人間関係悪化・差別が発生するおそれがあります。「曖昧性の問題」などは、正確性エラー(誤訳)にも関係するので、致命的な誤訳となり不測の事態、命の危険、勘違いなどを招くおそれがあります。

#### コラム 流暢性エラーをチェックする方法

現時点では、人手翻訳・機械翻訳ともに、翻訳後のチェック段階で補助的な自動ツールを使ってエラーを発見して修正する方法が考えられます。文書の校正支援ツールも役に立ちます。また、MTの場合、カスタムエンジンを使って待遇表現の改善を図る方法や、MTが(ある意味人間も)読み取り易いように留意しつつ、原文を修正してあいまい性の解消や文法誤りの改善を図る方法も考えられます(プリエディット)。



## 2.4 用語・専門用語

用語・専門用語は、特定の分野、企業、または製品の中で用いられる言葉です。翻訳においても、用語・専門用語は、的確かつ一貫していることが求められます。用語集を作成しておくことが多く、翻訳はそれに合わせます。用語・専門用語のエラーの例は以下の通りです。

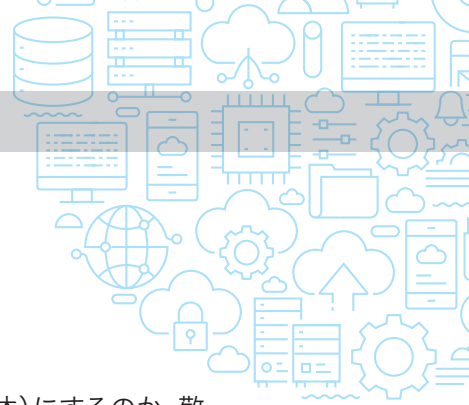
- 指定用語違反: 例、用語集で「use」を「利用」としているのに「使用」と和訳している。
- 特定分野用語違反: 例、自然言語処理分野では、「named entity」は「固有表現」と訳されるが、「名前付きエンティティ」と訳されてしまっている。
- 用語不統一: 例、「犬」と「イヌ」の両方がある。

### MTを使う時の注意

用語・専門用語に関するエラーは、特にMTに多いです。用語に誤りがあると、不測の事態、命の危険、勘違いなどが発生するおそれがあります。

#### コラム 用語・専門用語エラーを低減する方法

MTを使用する時に、用語を用語集に合わせることができる機能を利用することで解決できる場合もあります。たとえば、用語集を反映できる用語集機能があります。ただし、現在のNMTの用語集機能は、コントロールが難しいのが現状で、用語・専門用語エラーを完全になくせる保証はありません。翻訳エンジンによっては、「名詞」の統一には有効である一方で、「他の品詞」の統一には有効ではないものもあります。また、文脈によって訳語が変わることは大いにあるため、限定的な状況以外で、複数の意味を持つ用語に対してこの用語集機能を使うことはお勧めできません。この機能に限りませんが、日進月歩のMTは、各機能の特徴や限界を把握して利用する必要があります。



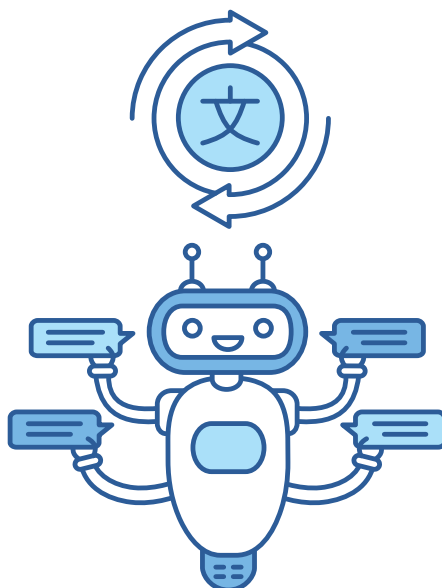
## 2.5 スタイル

スタイル(文体)とは、表記のスタイルに関する問題です。常体(だ・である体)にするのか、敬体(です・ます体)にするのか、数字を半角にするのか全角にするのかなどに関わる決め事です。一般的には、「スタイルガイド」という形で定められています。実務翻訳では、複数の翻訳者が関与したり既存訳と整合を図ったりすることが多いため、スタイルの統一が重要です。スタイル違反の例は以下の通りです。

- 指定スタイル違反: 例、スタイルガイドで常体(である調)と指定があるのに敬体(ですます調)を使っている。
- 特定分野スタイル違反: 例、指定スタイルではないものの、ある専門分野で一般的なスタイルが用いられていない。例、学术论文で敬体(ですます調)が使われている。
- スタイル不統一: 例、指定スタイルではないものの、スタイルが統一されていない。例、箇条書きに常体と敬体の項目が混在、「受入テスト」と「受け入れテスト」が混在している。

### MTを使う時の注意

スタイルエラーは、特にMTに多いエラーです。MTの翻訳文について、スタイルに誤りがあると、文書としての統一性が失われ読みづらくなります。現時点では、翻訳後に人手やツールで修正するしかありません。常体・敬体など限られた項目であれば、それらを選択できるMTを利用して解決を図ることができます。





## 2.6 JTF翻訳品質評価ガイドライン

上述した品質評価の指標は、世界的な共通の品質評価枠組みとして整備されています。日本においては、日本翻訳連盟 (JTF) が、その基準に準じた翻訳品質評価ガイドラインを提示しています。そのため海外企業と取引をする際にも齟齬や混乱を生じさせることなく利用できます。受注者と発注者を中心とした関係者が共通認識を持ち、品質に関する合意を形成しやすくなるはずです。

JTF翻訳品質評価ガイドラインの適用範囲と利用者は、受発注者が存在する翻訳ビジネスの場面での利用を想定していますが、MTの評価や、実務翻訳教育における訳文評価で活用することも可能です。翻訳ビジネスの場面においては、発注者となるクライアント企業や翻訳会社の担当者、受注者となる翻訳会社担当者や翻訳者を主な利用者として想定しています。

『JTF翻訳品質評価ガイドライン』は以下より無料でダウンロードできます。

## 2.7 おすすめ文献

- JTF翻訳品質評価ガイドライン

[https://www.jtf.jp/tips/translation\\_quality\\_guidelines](https://www.jtf.jp/tips/translation_quality_guidelines)

- JTF日本語標準スタイルガイド(翻訳用)

<https://www.jtf.jp/tips/styleguide>

- Multidimensional Quality Metrics (MQM)

Burchardt, A., & Lommel, A. (2014). QT-LaunchPad supplement 1: Practical guidelines for the use of MQM in scientific research on translation quality.

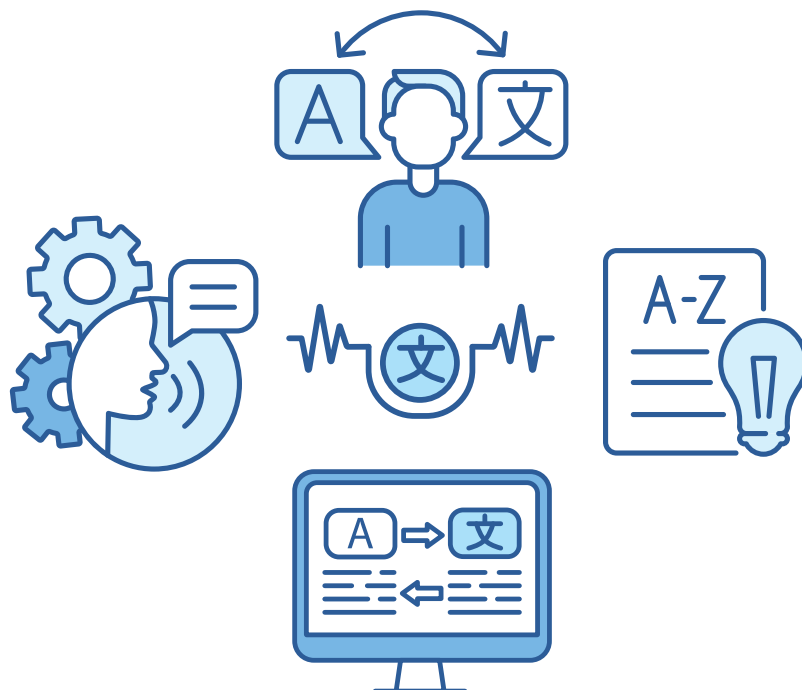
## 第3章 翻訳制作プロセス

### 3.1 人手翻訳の制作プロセス

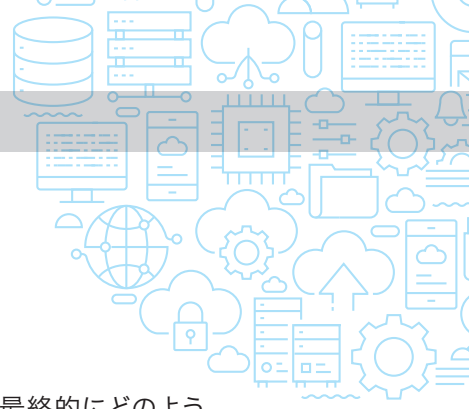
機械翻訳(MT)は誤訳をします。誤訳が原因でトラブルが生じた時の賠償責任もMTは負っていません。これらの問題を避けるためには、翻訳の専門家であるプロ翻訳者や翻訳会社などに翻訳制作を依頼したほうが良いこともあります。

最近では、MTの翻訳結果を人間が修正をするポストエディットという手法も良く耳にしますが、その場合も、素人の翻訳能力では太刀打ちできるようなものではありません。MTの翻訳結果に含まれるエラーを正しく修正するためには、2章で説明したような翻訳品質を理解し、トレーニングを積んだプロ翻訳者やチェッカーが翻訳制作を行うほうが適切です。

本章では、ISO17100・ISO18587等の国際標準に対応した翻訳制作プロセス及びプロ翻訳者の資格について説明します。







## 3.2 制作前プロセスと合意形成

翻訳会社における翻訳制作プロセスの特徴は、翻訳をする前に、発注者が最終的にどのような翻訳に仕上げたいのかを、翻訳会社（翻訳者）が計画して、発注者と合意をしておく「（翻訳）制作前プロセス」を設けていることです。そして、実際に翻訳を行う「（翻訳）制作プロセス」では、合意した仕様や品質要求が満たされているのかを確認するために複数の人たちが作業する体制がとられています。以下では、これら2つのプロセスを説明します。

### ■ 制作前プロセスで決めておく内容の例<sup>注1</sup>

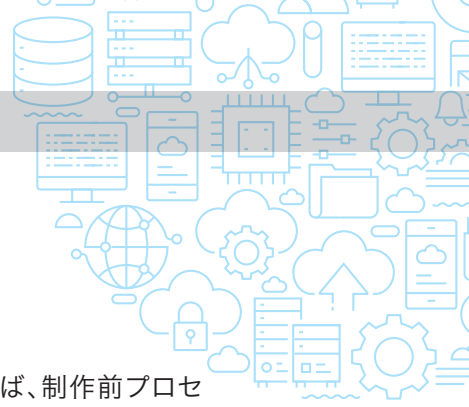
制作前プロセスで計画・合意しておく内容は、1) 翻訳の仕様・目的に関わること、2) 翻訳制作の工程に関わること、3) 作業環境に関わること、4) コスト・法などに関することがあります。

#### 1) 翻訳の仕様・目的に関わること

- 言語ペア：原文が何語で、訳文は何語か。複数の言語に同時に訳す必要はあるのか。
- 分野：原文が扱う分野、専門分野。
- 翻訳の目的：翻訳が使われる目的。たとえば、社内会議で使う資料なのか、それとも一般公開される文書なのか。
- 対象読者：翻訳の読者は誰なのか。原文が専門家のために書かれた物であっても、訳文を一般読者が読む場合がある。

#### 2) 翻訳制作の工程に関わること

- 翻訳者とチェッカーの選定。
- ネイティブチェックの有無：非母語への翻訳を翻訳者が行った（例：日本語母語話者が英語に翻訳をする）場合に、そのあとでネイティブ話者（英語を母語とする人）によるチェックを行うのか。
- 翻訳の依頼者によるチェックの有無：依頼者が、最後に訳文を確認するのか。



### 3) 作業環境に関わること

- 翻訳のための資料の有無：用語集やスタイルガイドはあるのか。なければ、制作前プロセスで用語集を作成する必要があるのか。
- 翻訳ツールの使用の有無：翻訳メモリやMTを使用するのか。
- 参考資料：手本ないし参照すべき翻訳や資料はあるのか。

### 4) コスト・法などに関わること

- 翻訳のコスト、チェックのコスト
- 翻訳作業以外の作業のコスト
- 納期
- 秘密保持契約
- 原文と訳文の著作権(所有権)

## 3.3 制作プロセス

制作プロセスとは、実際に翻訳を制作する工程です。制作プロセスでは最低でも、「翻訳者」が翻訳を行ったあとに、「チェッカー」(翻訳者とは別の人)が訳文の確認をします。確認作業は、バイリンガルチェック(原文と訳文を突き合わせて確認する)が必須です。これはISO17100での標準的な制作プロセスとして定められています。

合意した要件によっては、プルーフリード、編集、クライアントからのフィードバックの反映や修正なども追加で行うことがあります。なお、この工程で翻訳やチェックをする人たちは、専門家としての実力を備えている必要があります。

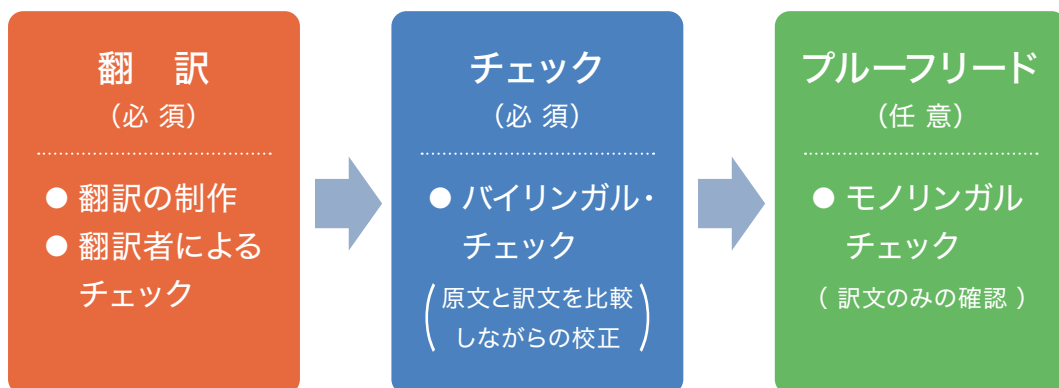
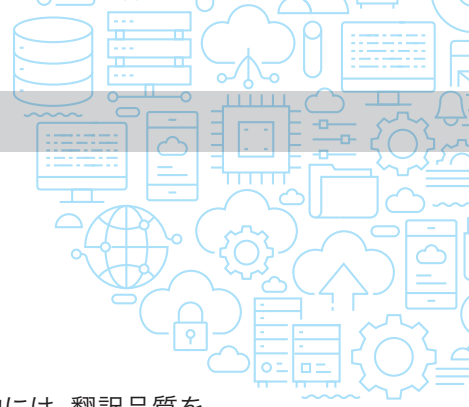


図 3-1 : 制作プロセスの必須/任意工程



## 3.4 ポストエディット (PE)

ポストエディットとは、MTの訳文を人手で修正・改善することです。最終的には、翻訳品質を合意した水準まで引き上げる必要があります。人手翻訳の場合は、翻訳者が、原文をゼロから翻訳しますが、ポストエディットの場合はMT訳からスタートするので、MT訳が良いと判断できる場合は、作業効率が向上すると期待されます。しかし、実際にはMT訳の悪い部分を、どの程度修正すべきなのか、作業員(翻訳者)が迷うことが多々あります。そのために余計に時間がかかってしまう事もあります。

正しいポストエディットの方法はまだ確立されていない状況で、素人がポストエディット行うのは危険です。最終的な翻訳品質は、MTのアウトプットを人間の翻訳者が処理する能力、つまりポストエディット能力に依存します。ポストエディットのプロセスを規定するISO18587もあります。失敗をしないために確認しておくべき項目について説明するTAUSポストエディットガイドラインなどもありますので参考にしてみてください。またポストエディットの経験を多くもつ翻訳会社などの専門家に相談するのも大切です。

## 3.5 フルPEとライトPE

ポストエディットには、大きく分けて「フルポストエディット(フルPE)」と「ライトポストエディット(ライトPE)」の2種類があります。フルPEは「人手翻訳」と同等な品質を目指すものです。ライトPEは原文に書いてある内容が大まかに理解できれば良いレベルにまで修正するものです。本書はMTを使って情報発信する人たちを対象とするため、それに対応する手法はフルPEです。

2章で紹介した翻訳品質評価の項目と対応づけて、フルPEとライトPEの要求品質を考えることもできます。フルPEは「正確性・流暢性・用語・スタイル」のすべての品質項目を十分に満たすように修正します。一方で、ライトPEは「正確性」に関わることのみを修正します。

	フルPE	ライトPE
正確性	○	○
流暢性	○	×
用語	○	×
スタイル	○	×

表 3-2 : フルPEとライトPE

## コラム ISO17100とISO18587

翻訳業務に関連した国際規格であるISO17100とISO18587を紹介します。両者に共通する特徴としては、翻訳品質そのものを保証する規格ではなく、高い翻訳品質を提供するために翻訳サービス提供者(TSP)が備えておくべき必要事項、翻訳工程、品質管理、リソース管理などから構成された翻訳ワークフローを定義する規格であるといった点が挙げられます。

たとえば、ISO17100には、下の工程表で示されるような翻訳ワークフローの流れが定義されています。

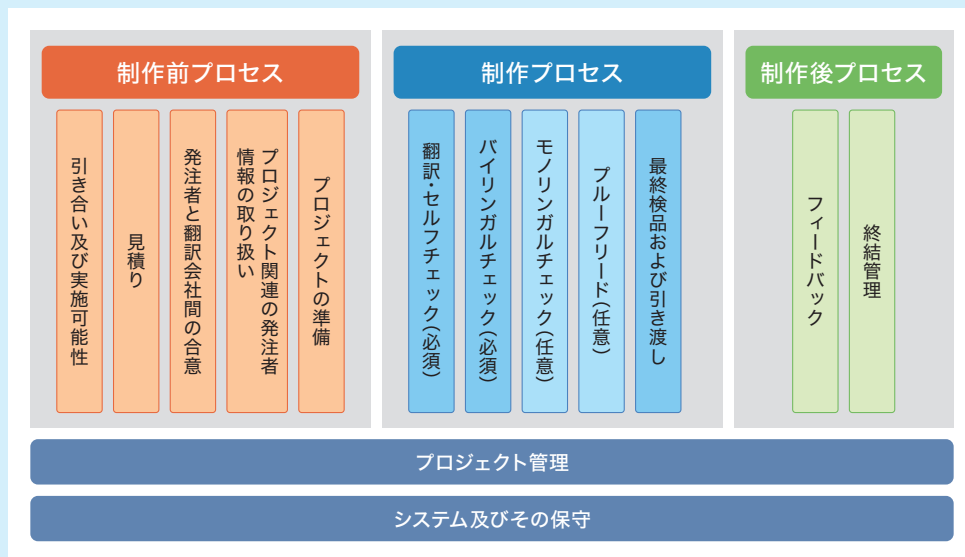


図 3-3 : ISO17100 翻訳制作プロセス<sup>注2</sup>

その後、MTの利用が広がったことを受けて、ISO18587という新たな規格が作成されました。この規格の特徴は、MTの出力を基にポストエディットを行う制作プロセスを翻訳ワークフローに位置付けた点です。

制作プロセスはISO17100と異なるものの、制作前プロセスについてはISO17100のものがほぼそのまま援用されています。たとえば、クライアントの合意を得ることやクライアント情報に留意してフローを設計することなどは、ISO17100と同様です。

ISO18587が想定しているのは「フルPE」です。ただし、最終的なテキストの公開を目的としていない場合は、情報の要点、テキストの趣旨などの表現に問題が生じないレベルのポストエディットでよしとする「ライトPE」も提唱されています(ただし、ISO18587の規格対象ではありません)。この場合、最終成果物としての翻訳文に対して、簡潔明瞭性は要求されていますが、スタイル的な要件は特に問われません。

## 3.6 おすすめ文献

- ISO (2012). ISO/TS 11669. Translation projects – General guidance. International Organization for Standardization.
- ISO (2015). ISO 17100. Translation services – Requirements for translation services. International Organization for Standardization.
- ISO (2017). ISO 18587. Translation services – Post-editing of machine translation output – Requirements. International Organization for Standardization.
- TAUS (2016). MT post-editing guidelines.  
<https://info.taus.net/mt-post-editing-guidelines>



注1：執筆にあたり以下を参照した。<http://www.ttt.org/specs/>

注2：ISO (2015), p.12のAnnex 1を基に作図した。

# 第4章 法律

## 4.1 MT利用者に関して発生する法律問題

### ■ みずからの情報発信のためにMTを利用する者

みずからの情報発信のためにMTを利用する者に関して生じる法律問題は、「MTを利用して翻訳を行い、翻訳結果を用いて情報発信を行ったところ、その翻訳結果に誤りがあり、その誤りに起因して第三者が損害を被った場合に、MT利用者(情報発信者)はどのような責任を負担するか」というものがあります。

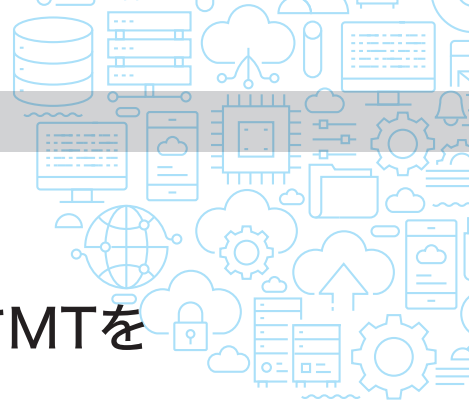
たとえば、地方公共団体が多言語にて情報発信する際にMTを利用して翻訳を行って情報発信を行ったが、当該翻訳結果に誤りがあり、当該翻訳結果を信用して行動した住民が損害を被った、というような例が考えられます。

この場合においては、単純にその情報発信者(MT利用者)が誤った情報を発信したということが問題にあり、それにより第三者が損害を被ったのであれば、その情報発信者(MT利用者)は法的責任を負うことになります。MTの技術特性(ブラックボックス性や翻訳結果の誤りが一定程度含まれる)は、当該情報発信者(MT利用者)の責任の有無や内容に全く影響を及ぼしません。



#### ❗ ポイント

- MTを使って情報発信をした場合において、当該発信情報の誤りに起因して第三者が損害を被った場合、情報発信者に責任が生じます。



## 4.2 他者から翻訳の依頼を受けてMTを利用する者

他者(以下「クライアント」という)から有償で翻訳の依頼を受け、当該翻訳に際してMTを利用する者(主として翻訳会社を想定している)に関して生じる法律問題は、①翻訳に際してMTを利用する行為がクライアントとの関係で問題ないか、②MTを利用した翻訳結果に誤りがあった場合にMT利用者はクライアントに対してどのような責任を負担するか、の2点です。

### 4.2.1 翻訳に際してMTを利用する行為がクライアントとの関係で問題ないか

この点については、翻訳に際してMTを利用する行為がクライアントとMT利用者との間の翻訳契約(特に当該翻訳契約内の秘密保持条項)に違反しないかが問題となります。

秘密保持条項は、端的に言うと「秘密情報」を「第三者に開示」する際には事前にクライアントの承諾を必要とする、という条項ですが、MT利用者がクライアントから受領した翻訳対象文書は通常「秘密情報」に該当します。

したがって、MT利用者が翻訳対象文書をMTサービスを利用して翻訳する行為が、秘密情報(翻訳対象文書)を「第三者に開示」したことになるかが問題です。

この点については、当該MTサービスにおけるセキュリティの有無や、有償無償は無関係で、実質的に当該MTサービスの利用が「第三者に開示」することに該当するか否かによります。

具体的には、(Google翻訳、DeepL、みんなの自動翻訳などの)MTサービスの提供者が、入力された翻訳対象文書にアクセスしないことがMTサービスの利用規約上明確化されており、かつ、MTサービス内で物理的にもMTサービス提供者によるアクセスが制限されている場合には、「第三者に開示」したことはありません。

次に、MTサービスの提供者が、入力された翻訳対象文書にアクセスする場面が保守・メンテナンスの場合などに限定されている場合も、「第三者に開示」として解釈できます。

一方、MTサービス提供者が、入力された翻訳対象文書を同サービス提供者自身のために



利用する(例:翻訳コーパスとして保存する、翻訳エンジンの学習のために利用する)場合には、MT利用者は、翻訳対象文書をMTサービス提供者という第三者に開示していると解さざるを得ないため、MTサービスの利用がクライアント・MT利用者との間の翻訳契約上許容されていないと見做されれば、秘密保持義務違反に該当する可能性が高いこととなります。

### ① ポイント

- 翻訳の依頼を受けてその翻訳対象文書をMTを使って翻訳をする場合は、翻訳の依頼者との間にどのような秘密保持契約が結ばれているのかを確認してください。
- MTサービスを利用する行為は、当該MTサービス提供者が、入力された翻訳対象文書を同サービス提供者自身のために利用する(例:翻訳コーパスとして保存する、翻訳エンジンの学習のために利用する)場合には、原文の情報を第三者に開示していると見做され秘密保持義務違反になる可能性があります。

## ■ 4.2.2 MTを利用した翻訳結果に誤りがあった場合に MT利用者がクライアントに対して負う責任

この点は、翻訳契約において翻訳結果に誤りがあった場合に翻訳者はどのような責任を負うかという一般的な問題であり、MTを利用したか否かは結論に影響を及ぼしません。したがって、本ユーザーガイドにおいては最低限の言及だけに留めます。

一般的に翻訳契約の法的性質は、請負契約(「翻訳」という仕事を完成させることが義務となっている契約)であり、仮に翻訳結果に誤りがあった場合、翻訳者は契約不適合責任(民法559条、同562条など)を負うこととなります。

具体的には、翻訳結果に誤りがあった場合には、翻訳者は再翻訳や翻訳報酬の減額、誤訳が原因で生じた損害についての賠償責任を負うこととなります。特に技術分野の翻訳については、誤訳により生じる損害額も多額に上る可能性があります。

一方、これら翻訳者(MT利用者)の責任は、翻訳契約において翻訳者の責任を制限(たとえば、損害賠償額について翻訳料金相当額を上限とするなど)することでコントロールすることができるため、翻訳者(MT利用者を含む)にとって合理的な内容の翻訳契約を締結することは非常に重要となります。



## 4.3 MT提供者に関して発生する法律問題

MT提供者に関して生じる重要な法律問題は、「MT提供者は、MT利用者(翻訳会社)から提供された翻訳対象文書を自らのMTエンジンの学習のために利用できるか」という問題です。



図 4-1: クライアント・MT利用者・MTサービス提供者の関係

この点に関しては、当該利用行為が「MT提供者とMT利用者との間のMTサービス利用契約に違反しないか」という問題と、「著作権法違反にならないか(著作権侵害にならないか)」の2つに分けて検討をする必要があります。

### 4.3.1 契約上の問題

まず当該利用行為が「MT提供者とMT利用者との間のMTサービス利用契約に違反しないか」ですが、この点については、MT提供者・MT利用者間のMTサービス利用契約において「翻訳対象文書をエンジン性能向上のために利用することが可能」と定められていれば、MTサービス利用契約違反の問題は生じません。

ただし、これは「MT提供者・MT利用者の間の契約」の問題であり、「MT利用者・クライアント間の翻訳契約」の問題ではありません。「MT利用者・クライアント間の翻訳契約」については、先ほど説明したとおりであり、MT利用者がクライアントの承諾なく翻訳対象文書をMT提供者に提供し、同データを学習に利用させた場合、当該行為は、「MT利用者・クライアント間の翻訳契約」の秘密保持条項に違反する可能性があります。

### 4.3.2 著作権法上の問題

次に「著作権法違反にならないか(著作権侵害にならないか)」の問題ですが、この点については結論から言うと著作権侵害にはなりません。

翻訳対象文書は著作物であるため、翻訳対象文書を学習に適した形に加工し(翻案)、学習に用いるためには、原則として当該文書の著作権者の許諾が必要です。



しかし、日本著作権法においては、翻訳エンジンのようなAIソフトウェアの生成に必要な著作物の利用行為（データの複製や翻案）については、原則として著作権者の承諾を行わなくても可能であるという権利制限規定が存在しています。それが、平成30年改正著作権法によって導入された著作権法30条の4第2号です。

（著作物に表現された思想又は感情の享受を目的としない利用）

第三十条の四 著作物は、次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合には、その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる。ただし、当該著作物の種類及び用途並びに当該利用の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

一 略

二 情報解析（多数の著作物その他の大量の情報から、当該情報を構成する言語、音、映像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うことをいう。第四十七条の五第一項第二号において同じ。）の用に供する場合

三 略

条文は上記のとおりですが、要するに「『情報解析』に必要な限度においては原則として著作物を自由に利用できる」という内容の権利制限規定です。

そして、翻訳エンジンのようなAIソフトウェアの開発は、当該「情報解析」に該当するため、結果として「翻訳エンジンの生成に必要な限度においては原則として著作物を自由に利用できる」ということとなります。

具体的には、30条の4の柱書には「その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる」と規定されているため、「自らが翻訳エンジンの開発を行うために、著作物を収集・複製・改変などする行為」「翻訳エンジンの開発を行う他社のために、著作物を収集・改変して学習用データセットを生成し提供する行為」などが可能です。



## 改訂履歴

Ver. 1.0 発行 (2022年9月1日)

一般社団法人アジア太平洋機械翻訳協会 (AAMT)

ウェブサイト: <https://aamt.info/>

## 制作者について

AAMTの「MTユーザーガイド委員会」が制作しました。

委員は以下の通りです (順序不同)。

## MTユーザーガイド委員会

### 委員長

山田優 (立教大学、AAMT理事)

### 委員会メンバー

安達久博 (株式会社サン・フレア、AAMT理事)

内山将夫 (情報通信研究機構、AAMT理事)

柴田知恵 (株式会社知財コーポレーション)

中澤敏明 (東京大学、AAMT理事)

永田哲 (個人翻訳者)

永田昌明 (日本電信電話株式会社、AAMT理事)

伊藤 (鍋田) 尚江 (株式会社フランシール)

波多野一真 (創価大学)

藤田篤 (情報通信研究機構)

ローズ三浦 (エッジ・トランスレーション)

宮本伸也 (日本ビジネス翻訳株式会社/NIPTA)

森口功造 (株式会社川村インターナショナル、AAMT理事)

## 執筆担当

1章 藤田篤・中澤敏明

2章 ローズ三浦・永田哲

3章 伊藤 (鍋田) 尚江・宮本伸也

4章 AAMT

クイックガイド 永田昌明

## 著作権

本ユーザーガイドの著作権は、一般社団法人アジア太平洋機械翻訳協会 (AAMT) が保有しています。

Copyright ©2022 Asia-Pacific Association for Machine Translation (AAMT)

## 利用許諾

MTユーザーガイドは、クリエイティブ・コモンズ・ライセンス「表示 4.0 国際 (CC BY-ND 4.0)」の下で利用が可能です。

同ライセンスの要約: <https://creativecommons.org/licenses/by-nd/4.0/deed.ja>

## 免責

本ユーザーガイドおよび関連文書は現状有姿で提供され、アジア太平洋機械翻訳協会 (AAMT) は、その正確性や使用について一切の保証をいたしません。これらの文書には、誤りや不正確な情報が含まれる可能性があります。AAMTは、事前の通知なくこれらの文書の記載内容を変更する権利を有します。