

LIONBRIDGE



ライオンブリッジ 2023 年 機械翻訳レポート

ライオンブリッジの専門家が、機械翻訳のトレンドを分析して今後の見通しについて考察し、機械翻訳テクノロジーの進歩が企業の発展にどのように貢献するかを予測します。

◎ エグゼクティブ サマリー

この数年、企業による機械翻訳 (MT) テクノロジーの導入が進んでいますが、2022 年には MT の品質に目立った動きはありませんでした。業界で最も長期にわたって主要な MT エンジン进行评估してきたライオンブリッジの機械翻訳トラッカーにより、2022 年には 5 大 MT エンジンに大きな品質向上が見られていないことが実証されています。

この評価結果は、2023 年に出現しつつある新しい機械翻訳テクノロジーの前触れであり、この新しいテクノロジーが

定着すれば MT の質は大幅に向上するでしょう。多くの企業がカスタマー エクスペリエンスの改善に向けて機械翻訳テクノロジーに注目しているのは、人による翻訳よりもはるかに高速で効果的な MT 翻訳手法を通じて、世界中の顧客とリアルタイムで対話することが可能になるためです。

機械翻訳はまた、言語や地域の壁を越えて従業員の生産性を向上させるためにも役立ち、新型コロナウイルス感染症のパンデミックの影響下においても、企業による従業員のエクスペリエンス向上にも貢献してきました。

機械翻訳のビジネス事例はここ数年で増加しています。

さまざまな分野の企業がビジネス上の利益を求めてどのように機械翻訳を活用しているかをご紹介します。

◎ e コマース / リテール

リテール業者は MT を活用して商品説明やマーケティング資料を短時間で効率的に翻訳し、複数の市場に提供しています。MT を利用することで、国境を越えた取引が飛躍的なスピードで増加しました。

◎ 旅行 / ホスピタリティ

旅行関連企業では、MT を使用してホテル、賃貸物件、レストラン、観光スポットなどの目的地の説明文や顧客レビューの翻訳を行っています。MT を使用することで、カスタマー エクスペリエンスを向上させ、見込み顧客とより迅速につながるできるようになっています。

◎ 医療

医療関連機関では、医学研究や臨床試験データの翻訳に MT が使用されています。MT の導入によって、情報をよりスムーズに一般公開することが可能となり、治療や患者の転帰の改善につながります。

◎ 法律 / 金融サービス

法律 / 金融分野のサービス プロバイダーは MT をいち早く採用し、e ディスカバリのプロセスや市場調査活動に活用してきました。MT によって、多言語コンテンツを大量に処理することが可能になりました。

◎ 公的機関

政府機関は、MT を通じて地域コミュニティで話されている言語で自治体向けサービスを提供することで、費用対効果の高い方法で言語の障壁を克服しています。

◎ あらゆる分野のグローバル企業

世界中に従業員を抱えるグローバル企業では、チャットや社内文書など、社内のコミュニケーションを円滑にするために MT が取り入れられています。

MT は、サポートチームが対応可能な範囲の拡大や、問題解決までの時間の短縮にも役立っています。グローバル企業は、MT を使用することでオフィスの生産性向上を実現しています。



今や、相互につながりあった世界規模の経済の中で成功を収めるために、グローバル企業は機械翻訳を導入して最大限に活用する必要があります。

MTは日々進歩しており、そのユースケースも増加傾向にあります。いまだに制限事項や限界があります。MTの品質に関する以前からの問題として、適切なフォーマルさの度合いやトーン（語調）、否定の表現が一貫した方法で表現されないことが挙げられ、このような制限がMTの成長の妨げとなっています。大規模言語モデル(LLM)の研究と利用によってこれらの問題が解決され、機械翻訳テクノロジーが新たな飛躍を遂げることが大いに期待されています。

ChatGPT、GPT-3、GPT-4モデルを開発したOpenAIにマイクロソフトが100億ドルを投資するなど、大手IT企業によるLLMテクノロジーへの投資は、LLMテクノロジーの開発を加速させ、自然言語処理(NLP)分野の発展を促進しています。こうした進歩は、翻訳・ローカライゼーション業界に大きな変化をもたらし、今後企業がコンテンツを作成・翻訳する方法が変化していくことは避けられないでしょう。

NLPの中でも特にLLMが急激に進歩することによって、コンテンツの制作やローカライズ方法が一変し、最終的には翻訳者が処理できるコンテンツの量が大幅に増えることで、生産性とスピードが飛躍的に向上することが予想されます。

コンテンツエンジンでAIを巧みに使いこなして活用できる企業が、デジタル化が進む経済社会において大きな競争優位性を獲得するでしょう。

機械翻訳の未来はどうなるのか

ライオンブリッジでは長年にわたって主要なMTエンジンを追跡し、新しいテクノロジーを巧みに活用してきました。当社はこうした知識に基づき、2023年以降の動向として、MTの既存のニューラル機械翻訳(NMT)パラダイムは終わりを迎え、おそらくChatGPTのような大規模言語モデル(LLM)に基づいた新しいパラダイムがNMTに取って代わることになると予測しています。実際に、すでにGPT-4のリリースやLLMの発展はビジネスに大きな影響を及ぼしています。

予想される変化には次のようなものがあります。

- ワークフローの自動化を含むMT品質の大幅な向上
- 制作されるコンテンツ量の増加
- 優秀な翻訳者の供給不足
- 機械翻訳の導入の増加
- カスタマーエクスペリエンス(CX)を向上させる手段としての機械翻訳

相互につながりあった世界規模の経済の中でグローバル企業が成功を収めるには、機械翻訳を導入して最大限に活用する必要があります。以下では、2022年における機械翻訳テクノロジーの進歩の状況、2023年時点の現状、そして今後数年間で予想される展望について考察していきます。

2022 年の機械翻訳の品質からわかる テクノロジーの進化

機械翻訳 (MT) の勢いはとどまるところを知らず、企業は MT テクノロジーを利用することのメリットをますます強く認識するようになってきました。この現象は、これまで MT テクノロジーの採用に消極的であったライフサイエンスのような分野でも同様に見られます。

しかし 2022 年は、MT が提供する翻訳品質については大きな変化が見られず、新しい刺激的な変化のための準備期間となりました。

変化としては、2022 年末にマイクロソフトが小規模の改善を行い、Amazon が若干の改良を加えたことくらいですが、これらの小規模の改善は過去の飛躍的な品質改良に比べればそれほど大きなものではなく、主要な MT エンジンの品質は同程度の水準となっています。

統計的 MT パラダイムの終了時にも、品質向上の速度が横ばいとなり、すべての上位エンジン間で MT 翻訳品質が均一化されるという同様の事象が見られたことから、当社は、現在のニューラル機械翻訳 (NMT) パラダイムが終わりに近づいているという一つの結論に達しました。

当社では、マルチモダリティと多言語を含む膨大なコンテンツを活用する大規模言語モデル (LLM) が、これから先の展開を大きく左右するだろうと予想しています。

コンテンツエンジンで AI を巧みに使いこなして活用できる企業は、デジタル化が進む経済社会において大きな競争優位性を獲得するでしょう。



機械翻訳入門

企業が機械翻訳を最大限に活用してその大きなメリットを享受し、ライオンブリッジが掲げる「Localize everything™」(あらゆるものをローカライズ)を初めて実現するためには、MTテクノロジーの進化について基本的な理解を深める必要があります。

MTの定義、MTが世界的に普及したきっかけ、MTの長所とMT使用時に避けるべき落とし穴、そしてMTが進化してきた背景について見ていきましょう。

人工知能


最も基本的な定義として、機械翻訳では人工知能(AI)、つまり機械が持つ「知能」を利用して、学習や問題解決など、人間のような思考が必要と思われるタスクを行います。機械翻訳の場合はAIを利用して翻訳を行います。ここ数年、AIはますます高まるコンピューティング能力を原動力として、飛躍的に性能が向上しています。強力なコンピューターによって目下のタスクを集中的に処理できるだけでなく、より先進的なマシンラーニングが可能になっていま

す。マシンラーニングとは、AIを活用するにあたって必要となる知識をコンピューターが得るための方法です。

マシンラーニング

マシンラーニング(ML)はコンピューターサイエンス分野の一つであり、大量のデータを活用してタスクを行う方法をコンピューターに教えることを指します。マシンラーニングでは特定のタスクに関連するデータを調査し、データからパターンを見出してそれらに関連付けます。そして学習した内容を利用して、コンピューターがタスクを行う方法を決めていきます。

この分析後、コンピューターによるタスクの実行が上達した場合は、マシンラーニングが発生したということになります。膨大な言語データとローカリゼーションデータが存在していることから、私たちはすでにマシンラーニングを利用して、天気予報、自動銘柄選択、機械翻訳といった、あらゆる場面でのコンピューターのパフォーマンスを強化しています。



マシンラーニングでは特定のタスクに関連するデータを調査し、データからパターンを見出してそれらに関連付けます。そして学習した内容を利用して、コンピューターがタスクを行う方法を決めていきます。



機械翻訳テクノロジーとは

機械翻訳とは自動翻訳を意味します。ある言語の文章をコンピューターに入力すると、その文章が他の言語で出力されます。MT テクノロジーは完璧ではありませんが、使用可能なレベルの翻訳をより効率的に生成するための最も強力なツールの一つです。

過去数十年間、出力物の品質と対応言語の幅広さという点で MT は進化を続けてきました。単純な単語置換システムであった黎明期の MT から、明示的にコーディングされた文法と辞書を持つルールベースの MT、大量の演算を行う枠組みを持つ統計的 MT、ディープラーニング（深層学習）とニューラルネットワークに基づくニューラル MT へと、MT の進化は、コンピューターの使用法がますます高度化している状況を反映しています。

機械翻訳の進化


ルールベース機械翻訳

当時の機械翻訳エンジンではルールベースの手法が使用されていました。その名が示すとおり、ルールベース機械翻訳 (RBMT) では、機械翻訳を実行するために開発されたルールや辞書に基づいて翻訳を処理し、単語ベースの処理を特徴としています。

ルールベース機械翻訳システムでも正確な翻訳を生成することはできますが、この方法は、簡略化された言語で、簡略化された文法と小規模な辞書を使用した場合に最も効果的です。

複雑さを軽減することで、機械翻訳エンジンの性能を向上させることができます。ルールベース機械翻訳では、特定の分野に関する翻訳が必要な場合や、慣用表現が使用されている場合、またはテキストに曖昧さがある場合は、優れた翻訳が生成される確率が低下します。とはいえ、RBMT が導入されて以来、言語テクノロジーは大きく進化してきました。

○ 機械翻訳は完璧ではありませんが、使用可能なレベルの翻訳を効率的に生成するための最も強力なツールの一つです。



次に、その学習アルゴリズムにより、翻訳対象の言語であるターゲット言語で一定の単語とフレーズが隣り合って出現する可能性を計算する言語モデルが構築されます。

統計モデル

統計的機械翻訳 (SMT) は、1つの原文の翻訳候補を大量に生成し、ターゲット言語で連動して出現する可能性が高い単語やフレーズに基づき、最適なものを1つ選択します。

SMTは「nグラム」というレンズを通して翻訳を学習します。nグラムとは、ソース言語(翻訳元の言語)とターゲット言語で連動して出現する単語の小さなグループです。SMTシステムには、トレーニング資料、つまり、ソース言語の原文と、それをターゲット言語に訳した訳文のサンプルが大量に与えられます。

学習アルゴリズムでは原文と訳文をnグラムに分割し、特定のソース言語nグラムが文章に出現したときに、どのターゲット言語nグラムが訳文で出現する傾向が高いかを判断します。

次に、学習アルゴリズムによってターゲット言語で一定の単語とフレーズが隣り合って出現する可能性を計算する言語モデルが構築されます。新しい文章の翻訳を行う際に、

SMTシステムは新しい原文をnグラムに分割し、関連性の高いターゲット言語nグラムを見つけて翻訳候補を生成します。

最終的な訳文は、ソース言語nグラムと最も関連性の高いターゲット言語nグラムを持ち、連動して出現する可能性が最も高いターゲット言語の単語を含む文章となります。SMTシステムをまったく言語的ではないと捉える見方もありますが、SMTは驚くほど有効に機能します。実際にSMTシステムはnグラムのみを考慮しており、文章全体は考慮していません。

ハイブリッド機械翻訳

次に、企業各社は統計的機械翻訳システムとルールベース機械翻訳システムの出力を組み合わせた「ハイブリッド機械翻訳 (HMT)」の実験を開始しました。これらの技術革新は、機械翻訳テクノロジーを世に知らしめ、世界規模の採用へとつながりました。そしてもう一つの技術的飛躍は、MTへの新しいアプローチであるニューラル機械翻訳によってもたらされます。

ニューラル機械翻訳

ニューラル機械翻訳 (NMT) は、SMT の最も大きな欠点である n グラム解析への依存を克服しました。NMT は、トレーニング資料を受け取るという点では SMT と同じですが、資料を受け取った後はそのデータの学習方法を自らが決定します。

NMT システムは原文の情報をベクトルに変換し、単語とその周囲の単語の情報を関連付けます。一部のシステムでは 1 つの単語につき数百もの情報ピースがあり、その分精度が高まっています。

NMT システムは、ディープラーニングにより単語や原文に関する大量の情報を捉えて、アテンションモデルと呼ばれる仕組みを使用し、大量のデータストリームの解析を通じて、翻訳プロセスにとって重要だと学習した特徴に焦点を合わせます。

その結果として生み出される翻訳は流暢さが劇的に改善されています。これは、コンピューターによる翻訳が今後ますます自然な文章になっていくことを意味します。また、ニューラル機械翻訳では、自動翻訳されたテキストの品質 (可読性) や、特定言語との非互換性の問題など、長らく存在した機械翻訳の欠点が解消されました。

ニューラル機械翻訳の開発によって機械翻訳は実用的な技術となり、膨大な量の文書を理解したり要点を把握したりするだけでなく、重要度の低い一般的なビジネス文書にも対応できるようになりました。3 つのアプローチの中では NMT が明らかな勝者と言えるでしょう。NMT は MT の使用を促進して制作プロセスを高速化することで、業界に大きな変革をもたらしました。

ニューラル機械翻訳 (NMT) の仕組み

NMT では統計的機械翻訳と同じくトレーニングに翻訳メモリ (TM) を使用する一方で、ディープラーニングと膨大な量のトレーニング用データを使用し、人工的なニューラルネットワークを構築します。

ニューラル機械翻訳はニューラルシステムを「育て上げる」ものと表現されることがあります。まるでピアノを弾くように、弾き間違えたら、立ち戻り、もう一度弾き直して、正確に弾けるようになるまで繰り返すのです。ニューラル MT システムがニューラルネットワークを経由して答えを見つけようとするのも、これと同様です。

トレーニング資料をディープラーニングのアルゴリズムに追加する際は、処理方法を必ずしも指定する必要はありません。ソース言語の原文周辺から得られるコンテキスト情報など、システムが自発的にパターンを見つけるようにします。しかし、そのプロセスの詳細は依然として、多くの点で謎に包まれています。

プロセスの大部分は複雑なデータの「隠れたレイヤー」の内部で処理されるため、ニューラルネットワークが実際どのような処理を行っているかを視覚化し、NMT のプロセスを具体的にイメージすることは困難です。そのため、必然的に人間はトレーニング資料のみを用意し、処理はアルゴリズムに任せ、翻訳が不正確だった場合にトレーニング資料を微調整するという流れになります。

○ トレーニング資料を受け取ると、システムはそのデータの学習方法を自らが決定します。

機械翻訳の歴史

機械翻訳は 1950 年代に開発が始まって以来、長い道のりを歩んできました。

以下は、機械翻訳の歴史における主なマイルストーンです。

- 1954 年 ○ 米国ジョージタウン大学の研究者が初期の MT システムの公開デモを初めて実施。
- 1962 年 ○ 米国で機械翻訳およびコンピューター利用言語学協会 (Association for Machine Translation and Computational Linguistics) が設立。
- 1964 年 ○ 米国科学アカデミー (National Academy of Sciences) が MT 研究のための委員会 (ALPAC) を設立。
- 1970 年 ○ フランス繊維協会 (IFT) が MT システムを使用して抄録の翻訳を開始。
- 1978 年 ○ Systran が技術マニュアルの翻訳を開始。
- 1989 年 ○ Trados が初の翻訳メモリ テクノロジーを開発し、市場に投入。
- 1991 年 ○ ウクライナ、ハルキウ国立大学でロシア語、英語、ドイツ語とウクライナ語の初の商用 MT システムを開発。
- 1996 年 ○ Systran と Babelfish が Web 上で少量のテキストの無料翻訳を提供。
- 2002 年 ○ ライオンブリッジがルールベースの MT エンジンを使用した初の商用 MT プロジェクトを実施。
- 2000 年代
中頃 ○ 統計的 MT システムが一般に公開。2006 年に Google 翻訳が、2007 年に Microsoft Live Translator が公開。
- 2012 年 ○ Google 翻訳では毎日書籍 100 万冊分のテキストが翻訳されていることを Google が発表。
- 2016 年 ○ Google とマイクロソフトがニューラル機械翻訳 (NMT) を実現したことにより、語順の誤りが大幅に減少し、語彙と文法が大幅に改善。
- 2020 年 ○ 10 月時点で Google Neural Machine Translation (GNMT) が 109 言語に対応。
- 2022 年 ○ 11 月にコンテキストに基づいて人間が書いたかのようなテキストを生成できる大規模言語モデル (LLM) である ChatGPT が注目を集め、機械翻訳に大きな影響が波及。
- 2023 年 ○ LLM が進化し、MT に破壊的イノベーションをもたらす中、MT の大きなパラダイム シフトが起こると予測。

機械翻訳のメリットとは

機械翻訳が非常に魅力的な理由は主に2つあります。

- 驚くほど高速で、数秒の間に大量の翻訳を提供
- 費用対効果が高く、企業が可能だと考えていた以上の翻訳量を実現

スピード + コスト削減 = ローカライズ版コンテンツの増加とカスタマー エクスペリエンスの向上

自動翻訳は「Localize Everything」を実現し、より多くのコンテンツをより短時間で市場に送り出したいという増え続ける需要に、これまでと同じ、あるいはそれ以下の予算で対応できるようになります。

企業が機械翻訳を導入すべきその他の理由

時間の節約とローカリゼーションコストの削減に加え、重要性の低いコンテンツに MT を使用することで、企業は以下のことを実現できます。

- より多くの資料をローカライズすることで、翻訳する言語の数と参入する市場の数に基づいてリーチを拡大
- 用語とブランド ボイスの一貫性の向上
- 顧客とのリアルタイムのやり取りをローカライズすることで顧客満足度を向上
- コミュニケーションの障壁を取り除くことで、世界中に広がるチームの生産性が向上

機械翻訳を導入することで翻訳に必要な予算を抑えることができるため、企業はコンテンツ制作により多くのリソースや予算を割り当てることができます。

予算の再配分を行うことにより、企業はコンテンツ制作のスピードを向上させてより多くの読者を獲得し、成長を促進することができます。これはデジタルファーストの時代に競い合う企業各社にとって、特に大きなメリットとなります。

ワークフローの一環として機械翻訳を利用されている当社のお客様の多くは、以下のようなメリットを得ています。



翻訳効率の大幅な向上

最大 40% のコスト削減。



納期の短縮

納品までの時間を最大 60% 短縮。



カスタマー エクスペリエンスの向上

場所を問わずに、顧客の母語でのコミュニケーションが可能。

機械翻訳の主な欠点

MT エンジン はますます洗練され、翻訳の品質も向上していますが、ニュアンスやトーン（語調）、皮肉、ユーモア、意味合いなどに対する理解については、まだ人間の能力に到底及んでいません。

機械翻訳を単体で実施した場合はトーンが人工的になり、メッセージが正しく伝わらない可能性があるため、顧客にサービスの品質が低い、または信頼できないと思われるリスクが生まれます。

公開コンテンツの翻訳は、読みやすく、対象となる市場の文化的な文脈に自然になじむものでなければなりません。質の高いローカライズ版コンテンツを制作するためには、多少コストがかかってもネイティブスピーカーによる MT 出力のレビューを行う必要があります。このレビュー作業は「機械翻訳ポストエディット (MTPE)」と呼ばれます。

機械翻訳ポストエディットによって 翻訳者のスキルを通じて MT 翻訳を最適化

機械翻訳ポストエディット (MTPE) とは MT と人間の手による翻訳を組み合わせたプロセスで、その名が示すとおり機械翻訳処理後に翻訳者が編集・校正を行います。最初にソフトウェアによって対象となる文章から初期翻訳が生成され、次に翻訳者がその文章をチェックし、正確性、明瞭性、流暢性、対象地域への適合性などの観点に基づいて訳文を編集します。このアプローチにより、機械翻訳のスピードと、翻訳者による注意力や繊細さのメリットの両方を取り入れることができます。

機械翻訳エンジンによって短時間で処理された一括翻訳を翻訳者が自然な翻訳に直していきます。機械翻訳とポストエディットの組み合わせは、顧客向けコンテンツの翻訳品質を最大限に高めながらも納期をできるだけ短縮したいというお客様のニーズに応えるもので、高品質な翻訳を短期間で得るための費用対効果の高い方法の一つと言えます。



機械翻訳に関連するその他のリスク

MT は大幅に進歩しましたが、そのテクノロジーはまだ完璧ではありません。MT テクノロジーを導入する際、特に MT の欠点を軽減するための対策が講じられていない場合は、出力された翻訳のトーンが人工的になること以外のリスクも存在します。

具体的な問題

機械翻訳の使用時には以下のような問題が起こる可能性があります。

- 機械翻訳エンジンが中立的な言葉を使うタイミングや対象のオーディエンスに適したフォーマルさの度合いを識別できないことが原因で、表現が包括性に欠ける場合や意図するフォーマルさの度合いにならない場合があります。
- 翻訳出力に大小さまざまな誤りが含まれる可能性があります。

こうした問題はそれほどマイナスに作用しないこともありますが、誤りの場合は企業全体に特に悪影響を及ぼすおそれがあります。



○ **いまだコンピューターでは人間と同じような判断はできないのです。**

MT の利用に伴うリスク - 誤り

MT で発生する誤りの種類

MT エンジンでは、大きく分けて 2 種類の誤りが発生します。

○ 標準的な誤り

ターゲット言語に含まれるこうした誤りは、コンテンツの言語的特徴と関連しています。文法、スペル、句読点などの誤りが標準的な誤りに分類されます。こうした間違いはネイティブスピーカーであれば容易に気付けるものですが、これらの誤りが企業に壊滅的な損害を与えることはまれです。

○ 重大な誤り

MT による重大な誤りは言語的特徴にとどまらず、翻訳が原文のメッセージの意図から逸脱した場合に発生します。こうした誤りから生じた誤情報や誤解は、企業の評判への悪影響、財務的損失、法律上のトラブルにつながるほか、公共安全や公衆衛生に有害な結果を招く可能性があります。

MT エンジンによる重大な誤りが発生する理由

重大な誤りは、MT エンジンの動作エラーと捉えることができます。こうしたエラーは、たとえばある 1 つの語句に意味が 2 つある場合や、原文にスペルミスがある場合など、MT エンジンが原文の文脈を理解できない状況で発生します。また、MT エンジンのトレーニング不足や使用している用語集の不備によっても、こうした誤りが繰り返し発生する可能性があります。こうした重大な誤りが起こるのは、今日の MT エンジンがますます進化しているとはいえ、まだまだ完璧ではないためです。いまだコンピューターでは人間と同じような判断はできないのです。

重大な誤りは、主に 3 つの状況で発生します。

○ 重要な名詞の誤訳

これは、固有名詞 (人物や組織)、重要な数字、測定単位に関する翻訳上の誤りを指します。

○ 否定と逆の意味

この種の誤りは、ターゲット言語に、原文の意図とは逆の意味になる間違いが含まれる状況のことです。

○ ハルシネーション

この種の誤りは、原文にはない内容が MT によって訳文に追加された場合に起こります。場合によっては、不快な語句、下品な表現、攻撃的な表現、扱いに注意を要する表現が出力されることもあります。この種の重大な誤りが起こる場合、通常は MT エンジン ソフトウェア自体に問題があります。

そのような人名に関する重大な誤りが、スペイン政府の Web サイトで実際にありました。ある省の公式サイトにおいて、一部署の長であった Dolores del Campo 氏の名前の代わりに、その逐語訳である「It is pain of field」(「野原の痛み」の意)と記載されていたのです。

重大な誤りを最小限に抑える方法

コンピューター科学者が既存の MT テクノロジーを改善し、重大な誤りが最初から発生しないようにすることで初めて、自動化テクノロジーを活用して、翻訳プロセスで起こり得る問題を予め特定し、問題のある文を修正して翻訳の精度を高めることができます。たとえば、当社ではプロの翻訳者が介入する「ポストエディット」プロセスを最小限に抑えつつ、MT 処理のスピードを損なわずに重大な誤りを検出できる Smart MT™ サービスと、最先端の Smart Content™ 言語 AI とを組み合わせ、翻訳済みテキストに自動品質チェックを適用することで、上述の重大な誤りを早期段階で検出します。

この自動処理では以下のような誤りが検出されます。

- **個人名や組織名などの固有名詞の誤訳**
原文に含まれる語句から、固有名詞にも一般名詞にもなり得る語句を特定
- **不快な語句、下品な表現、扱いに注意を要する表現**
教師ありマシン ラーニング (ML) アルゴリズムと問題のある語句のリストを組み合わせ使用
- **原文と訳文とで意味が逆になる誤訳**
否定助詞 (たとえば英単語の not やその短縮形 n't) を含む文が、原文 (英語) と訳文の両方ではなく一方だけに含まれているケースを特定
- **訳文中のハルシネーション (訳文には含まれるが原文には含まれない語句) を辞書を使用して検出**
侮辱的な語句が混入するケースは問題のある語句のリストを使用して検出

ただし、自動品質チェックで重大な誤りを 100% 防ぐことはできません。自動チェックが誤りを検出できず、本来検出されるべき問題がそのまま訳に含まれてしまう結果となることもあります。それでもなお、問題箇所を特定するうえで自動品質チェックは非常に効果的です。

このアプローチを採用することで、翻訳者は要注意として検出された文に集中でき、文書全体の再チェックが不要になります。翻訳者に問題の起こりそうな箇所への注意を促すことで、ローカリゼーション プロセスの効率が向上します。

こうした重大な誤りが起こるのは、今日の MT エンジンがますます進化しているとはいえ、まだまだ完璧ではないためです。いまだコンピューターでは人間と同じような判断はできないのです。





MT の利用に伴うリスク - 不適切なフォーマルさの度合い

言葉遣いにおける「フォーマルさの度合い」とその使い分けが重要な理由

言葉遣いのフォーマルさの度合いとは、自分自身を表現するために使用するフォーマルさの程度のことです。私たちは通常、親しい友人と上司に対してはそれぞれ異なる語彙や文法を選んで使用し、同じ話をする場合でも同じように話すことはありません。

多くの人は、ビジネスシーンやフォーマルな場面ではフォーマルな言葉遣いを用いて丁寧さと敬意を表し、リラックスした場面や親しい人と話すときはカジュアルな言葉遣いを用います。

文化によっては、適切な度合いのフォーマル表現を用いていないコンテンツは失礼にあたる可能性があります。たとえば、韓国語では子どもに話しかけるように年長者に話しかけることはありません。韓国語では7つのスピーチレベルごとに異なる動詞の語尾が使用され、そのうち5つはフォーマル、残りの2つはカジュアルに分類されます。フォーマルさの度合いを間違えると恥ずかしい思いをすることもあります。

逆に、フォーマルな言葉遣いがふさわしくない事例も多数あります。多くの場合は、それが不快だからではなく、その場にまったくそぐわないことが理由です。こういった誤認識や表現のギャップにより、オーディエンスが離れてしまう可能性もあります。たとえば、コンピューターゲームや学生向けのプログラムの翻訳では、フォーマルな表現よりもカジュアルなトーンが適していると言えます。

また、コンテンツの翻訳にあたり、企業が費用や手間を惜しまずに対象オーディエンスに適したスタイルを取り入れる努力をすることで、顧客を重視しているという暗示的なメッセージを対象市場に届けることにもつながります。こういった努力は潜在的顧客の注意も引き、効果的なアピールとなり得ます。多くのオーディエンスと効果的に関係を構築・強化できれば、成果にもつながります。

文化によっては、適切な度合いのフォーマル表現を用いていないコンテンツは失礼にあたる可能性があります。

MT エンジンがフォーマルさの度合いを適切に判断するのが苦手な理由

機械翻訳エンジンでは、フォーマルさの度合いが不適切な訳文や、フォーマルさに一貫性のない訳文が生成されることがあるため、機械翻訳を利用する際のフォーマル/カジュアルな言葉遣いの使い分けには注意が必要です。これは、MT モデルでは通常、入力セグメントごとに1つの翻訳が返されることが原因です。入力セグメントに曖昧さがある場合は、妥当な選択肢の中から1つの翻訳が選ばれますが、このとき対象のオーディエンスについては考慮されません。

妥当性のある複数の選択肢の中から1つをモデルに選択させると、翻訳の一貫性が損なわれる場合や、フォーマルさの度合いが適切でない訳文が生成されることがあります。

ターゲット言語に、ソース言語よりも多様なフォーマルさの段階がある場合は特に、適切な出力を得るのが難しくなります。たとえば、フランス語では、親密さの度合いによって「tu」(カジュアルな二人称)と「vous」(フォーマルな二人称)を使い分けますが、英語にはこのような区別はありません。

当社のエンタープライズ向け機械翻訳ソリューション Smairt MT™ では、ターゲット言語のテキストに対して言語ルールを適用でき、スタイルやフォーマルさの度合いを指定して機械翻訳を生成することができます。

MT 利用時の言葉遣いにおけるフォーマルさの度合い
企業は、ルールを使用して望ましくないスタイルを正しい翻訳に置き換えるルールベースの手法と、カスタム MT モデルを使用するルールに基づかない手法によって、言葉遣いのフォーマルさの度合いを制御することができます。

ほとんどの商用 MT システムは、フォーマルさの度合いや性別の違いに対応していませんが、進歩は見られています。DeepL (API) と Amazon (コンソールおよび SDK) には、フォーマルさを制御する機能が備わっています。

フォーマルさの度合いは、3つのオプション(デフォルト、フォーマル、カジュアル)から選ぶことができます。デフォルトでは、ニューラル機械翻訳(NMT)出力のフォーマルさの度合いは変更されません。フォーマル/カジュアルのオプションでは、ユーザーがフォーマルまたはカジュアルなトーンを選択できます。具体的には、この機能では、翻訳で使用される代名詞と関連する単語が特定されます。

ライオンブリッジのエンタープライズ向け機械翻訳ソリューションである Smairt MT では、ターゲット言語のテキストに対して言語ルールを適用でき、スタイルやフォーマルさの度合いを指定して機械翻訳を生成することができます。中立な表現を使用し、たとえば「salesman」と「salesperson」の違いに見られるようなローカリゼーション中に生じる偏見を排除することもできます。

当社の専門家によって最新の状態に維持されたルールデータベースを MT 出力の解析に適用し、MT 出力を制御します。十分なコンテンツがある場合は、ルールベースの手法とカスタム MT モデルを組み合わせることで最適な結果を得られます。

MT を利用して適切なフォーマルさの度合いで翻訳を生成することは、容易ではありませんが実現可能です。本ホワイトペーパーの後半では、言葉遣いのフォーマルさに関連する問題を解決する可能性がある画期的なテクノロジーの進歩についてご紹介します。

◎ 機械翻訳の1年間のまとめ(2022年)

2022年の主要なMTトレンド

ライオンブリッジのMT専門家は、2022年に起こったことと起こらなかったことの両方に注目すべきだと考えています。過去数年間、MT関連のテクノロジーの数々の進歩を見てきた当社では、同様のことが起こると予想していましたが、当社の機械翻訳トラッカーによる評価結果からもわかるようにMTに大きな進歩は見られませんでした。

まれな例外を除いて、主要MTエンジンにはこの1年間ほとんど改良が加えられていません。この傾向は将来にも影響を及ぼすことが予想されますが、まずは2022年の状況を詳しく見ていきましょう。

2022年における上位MTエンジンの性能

企業がMTの利用を開始する際や現行のMTの利用方法を改善する際には、特定のニーズを満たすために最も効果的なMTエンジンを特定することが非常に重要です。また、2022年における主要なMTエンジンの性能を掘り下げていくとはっきりするのは、1つのエンジンですべてをこなすことはできないということです。

言語ごとのMTエンジンの性能の比較

スペイン語のコンテンツを扱うある企業では、自動翻訳にDeepLを採用してその恩恵を受けていますが、日本語の翻訳にはより日本語翻訳に適したエンジンを利用しています。これは、各エンジンの性能が、取り扱う言語によって異なるからです。

当社は、3つの主要エンジンが英語から多言語への自動翻訳をどの程度の品質で処理するのかを計算しました。翻訳

の品質は平均編集距離を算出することで測定します。編集距離は、MTの出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量を算出するものです。

編集距離が小さければ小さいほど自動翻訳が効果的であると言えます。図1に分析結果が示されているように、こうした情報に注意を払うことは企業にとって有意義です。

当社の分析では、特定の状況において以下のような結果が得られました。

- DeepLのスペイン語翻訳はGoogleやマイクロソフトよりも優れた性能を示しました
- Googleの日本語翻訳はDeepLよりも優れた性能を示しました
- マイクロソフトのポーランド語翻訳はDeepLよりも優れた性能を示しました
- 3つのエンジンは、イタリア語、トルコ語、ヘブライ語で同程度の性能を示しました

これらの分析結果は、異なる言語、文化、分野のニュアンスや複雑性を扱う機械翻訳に特有の複雑さと課題を示しています。

単一のアルゴリズムや手法によってすべての言語とすべての種類のコンテンツに完璧に対応することは不可能であるため、MTエンジン間で性能にばらつきがあるのも無理はありません。



MT の性能が言語によって異なる理由

このような性能の違いが生じる理由として考えられるのは、各言語で利用可能なトレーニングデータの質と量に違いがあることです。機械翻訳モデルは、精度の高い翻訳を生成する方法を学習するために高品質のバイリンガル言語データから構成される大規模なコーパスを必要としますが、特定の言語に対して十分な言語データが提供されていない場合、MT エンジンは精度の高い翻訳を生成することが困難になります。

さらに、MT エンジンにはそれぞれ異なる技術的・言語的特徴があり、各エンジンごとに独自のアルゴリズム、アーキテクチャ、機械翻訳へのアプローチが採用されています。特定のエンジンが、そのモデルの特徴や能力に応じて、特定の言語ペアや分野で高い性能を発揮することもあります。さらには MT エンジンの最適化とカスタマイズのレベ

ルに応じて性能が異なる場合もあります。一部の MT エンジンは特定の言語や分野向けに最適化・カスタマイズされており、こうしたユースケースでは、より高精度で効果的な翻訳を生成することができます。

全体的にはこれらの要因が機械翻訳の複雑さとばらつきに寄与しており、さまざまな言語やユースケースに対する各種 MT エンジンの長所と限界を理解することの重要性を浮き彫りにしています。

特定のユースケースに対して最適な MT エンジンを選択することは簡単な作業ではなく、機械翻訳に関する専門知識が必要となります。最善の結果を得るために非常に重要となるのは、慎重に評価を行うことと最適化を継続することです。

言語ごとの MT エンジンの平均性能

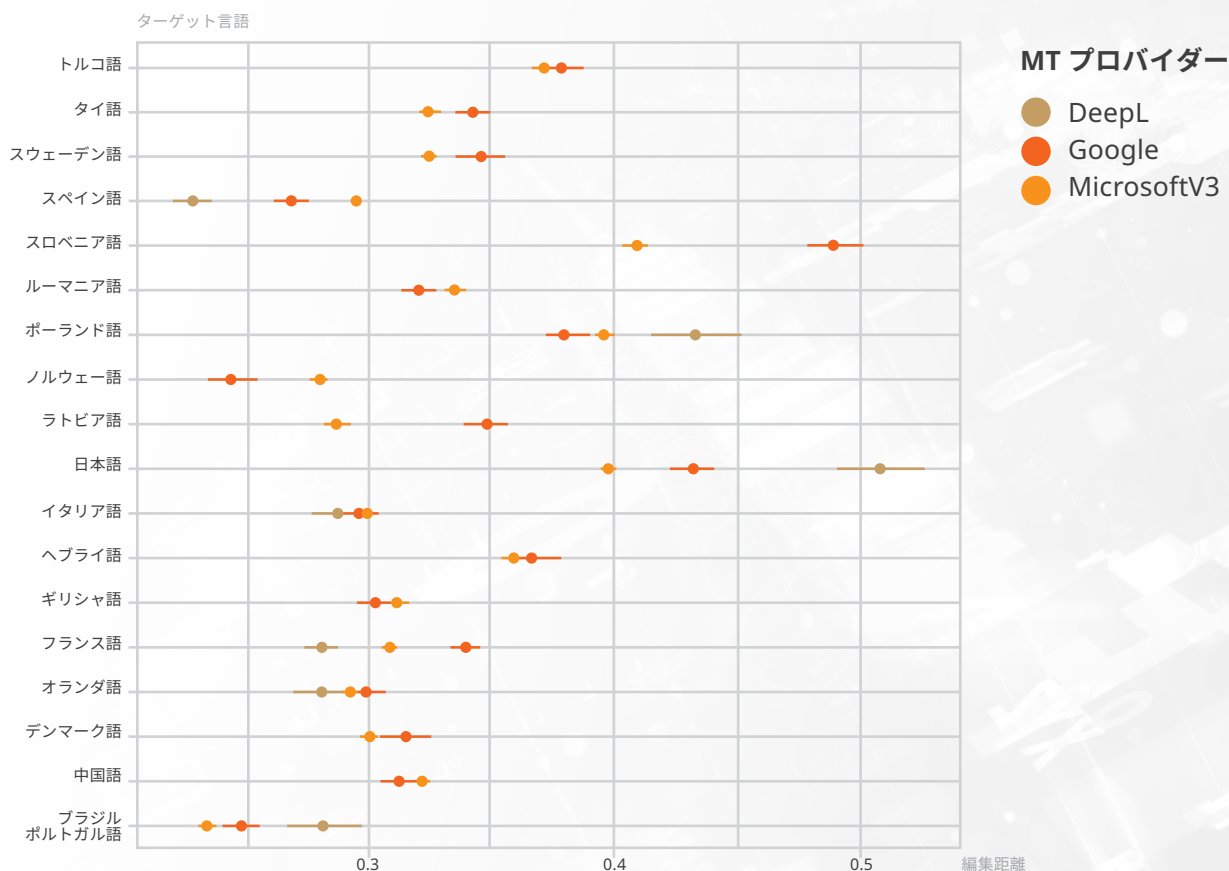


図 1. 言語ごとの平均編集距離と MT プロバイダー

分野ごとの自動翻訳の性能

コンテンツがクリエイティブであればあるほど、MT エンジンで効果的に翻訳することは難しくなります。

分野ごとの翻訳品質の違いを見るとその点は明らかです。当社は、主要な分野の平均編集距離 (MT の出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量) を追跡し、MT エンジンがさまざまな種類のコンテンツをどの程度の品質で処理するのかを調査しました。

編集距離が小さければ小さいほど翻訳品質は優れています。

図 2 に示すとおり、当社の分析では以下の結果が得られました。

- MT エンジンによる翻訳が最も困難なコンテンツは、独創性が特徴であるメディア・マーケティング分野のコンテンツであるという結果が得られました。
- MT エンジンによる処理が 2 番目に困難なコンテンツは、想像力が豊かでニュアンスに富んだ表現が多用される繊維・ファッション分野のコンテンツであるという結果が得られました。
- 機械翻訳の品質が最も優れていたのは、表現が明確な自動車・機械分野のコンテンツでした。

当社ではこうした結果が出ることを予想していました。汎用型 MT エンジンは、比較的シンプルでわかりやすく、構造と語彙が明確で、専門性や技術性が低い明瞭なコンテンツを処理する際にその威力を発揮します。

分野ごとの MT 出力の品質

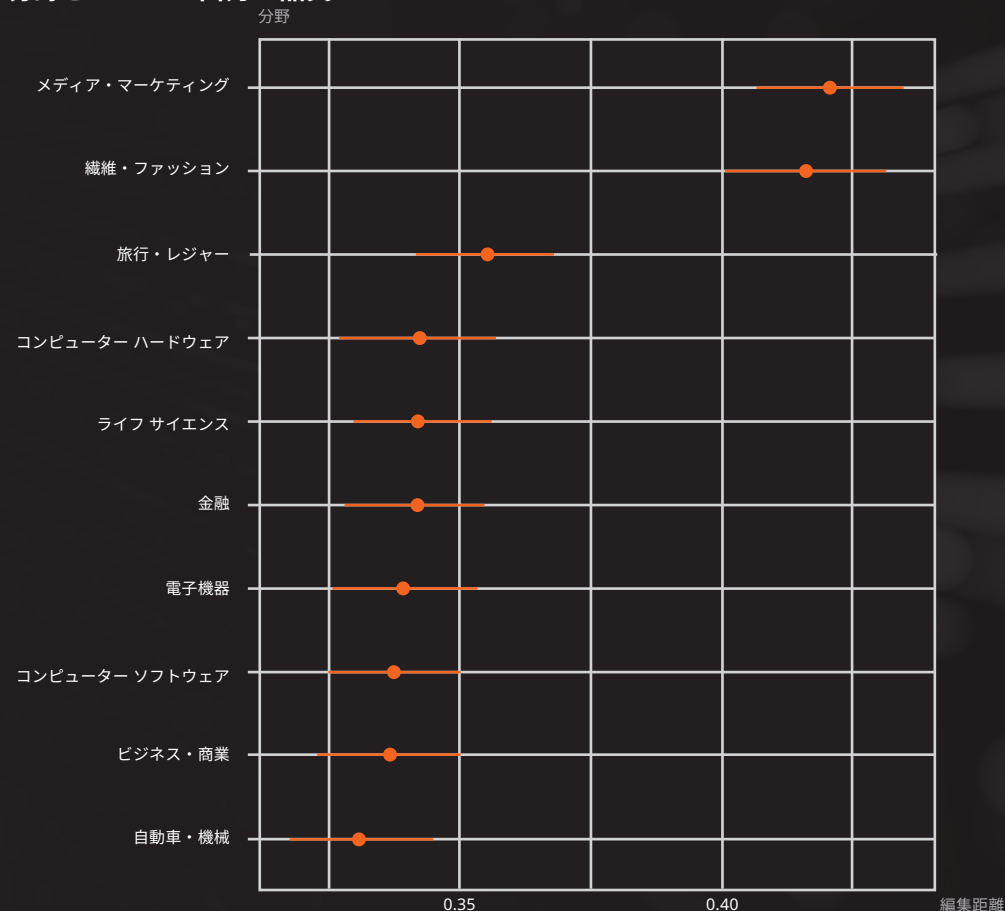


図 2. 分野ごとの平均編集距離

分野と MT エンジンに基づいた MT エンジン性能の比較

メディア・マーケティング分野のコンテンツを翻訳している企業では、DeepL を MT プロバイダーとして選択したことがプラスに働いた一方で、繊維・ファッション分野のコンテンツを翻訳している企業にはもっと適した選択肢がありました。利用可能なデータに基づき、当社は 4 つの主要翻訳エンジンがさまざまな分野のコンテンツをどの程度の品質で処理するのかを考察しました。翻訳の品質は平均編集距離 (MT の出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量) を算出することで判断しました。編集距離が小さければ小さいほど自動翻訳が効果的であると言えます。図 3 に示されるように、翻訳エンジンは特定の状況において、分野やコンテンツの種類を問わず同様の性能を発揮するわけではないことが明らかになりました。

当社の分析では、特定の分野について以下のような結果が得られました。

- **メディア・マーケティング** - DeepL は Google やマイクロソフトよりも優れた翻訳性能を示しました
- **ライフサイエンス** - DeepL はマイクロソフトや Google よりも優れた翻訳性能を示しました

- **金融** - マイクロソフトと DeepL は Google よりも優れた翻訳性能を示しました
- **自動車・機械** - Google はマイクロソフトや DeepL よりもわずかに優れた翻訳性能を示しました

ただし、当社がお客様に適切なシステムをお勧めするためには、お客様のコンテンツを分析してお客様のニーズを把握し、求められる要件に最も適したエンジンはどれであることを判断する必要があります。

MT エンジンの性能は、先ほど述べたように言語ペアに影響されるだけではなく、特定の分野や業界で使用される用語や言い回しにも左右される可能性があることが上記の分析結果で示されています。

このため、MT エンジンを選択する際には、翻訳対象の言語ペアと、特定の分野および業界のコンテキストを考慮することが非常に重要となります。これは言語と特定の分野に関する専門知識と見識が必要なアプローチですが、うまく活用することで特定のコンテンツの種類や業界に関連する翻訳の品質と精度を向上させることができます。

分野ごとの MT エンジンの平均性能

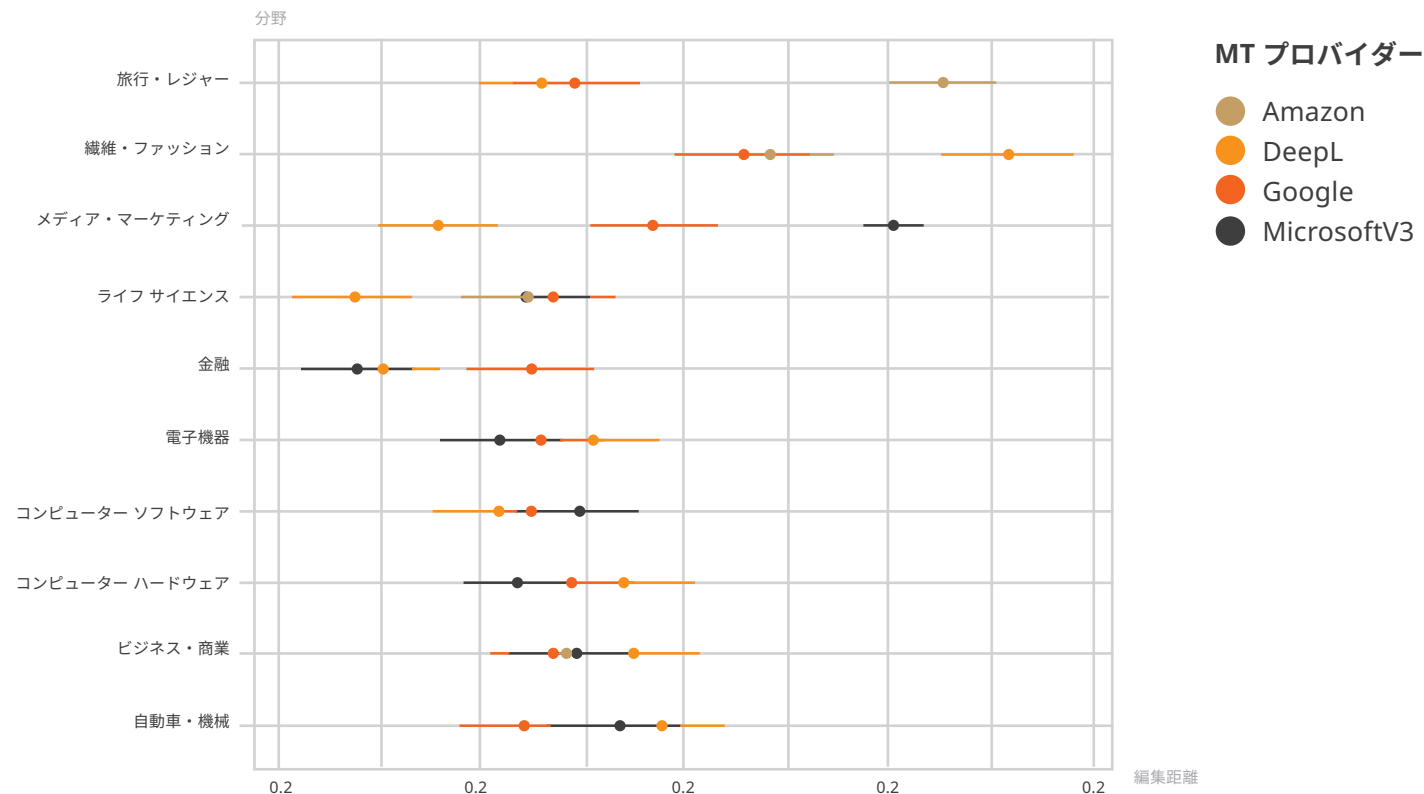


図 3. 分野ごとの平均編集距離と MT プロバイダー

機械翻訳エンジンの品質の比較

主要な MT エンジンに総合的な性能の違いは見られるのでしょうか。図 4 は、2018 年 5 月から 2022 年 12 月までの期間に逆編集距離を使って測定された、上位 5 つの MT エンジンによる、ドイツ語、スペイン語、ロシア語、中国語の出力品質を当社が比較したものです。

編集距離は、MT の出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量を算出するものです。逆編集距離は、結果の数値が大きいほど品質が優れていることを意味します。

図 5 は、各エンジンのトレンドラインを示すものです。図 4 と図 5 を比較することで、上位の MT プロバイダーのエンジンの品質の差が縮まり収束しているという興味深いインサイトが得られます。

当社の分析結果は以下のことを示しています。

- Microsoft Bing がハイペースで上位陣を追い上げています
- Amazon と Google のトレンドラインはほぼ平行です
- テクノロジー全般として、2022 年に機械翻訳に大きな進歩は見られません

2023 年初頭には、主要エンジン間の差はほとんどなくなっています。ニューラル機械翻訳 (NMT) のパラダイムが優勢であり続け、MT プロバイダーもこれまでと同様のペースで投資を継続すると想定した場合、当社は、言語ペアや分野によって多少の差はあるとしても、すべての主要 MT エンジンの性能は 1 年以内にほぼ同じ水準に収束すると予想しています。



○ 2023 年初頭には、主要エンジン間の差はほとんどなくなっています。

MT エンジンの性能

逆編集距離

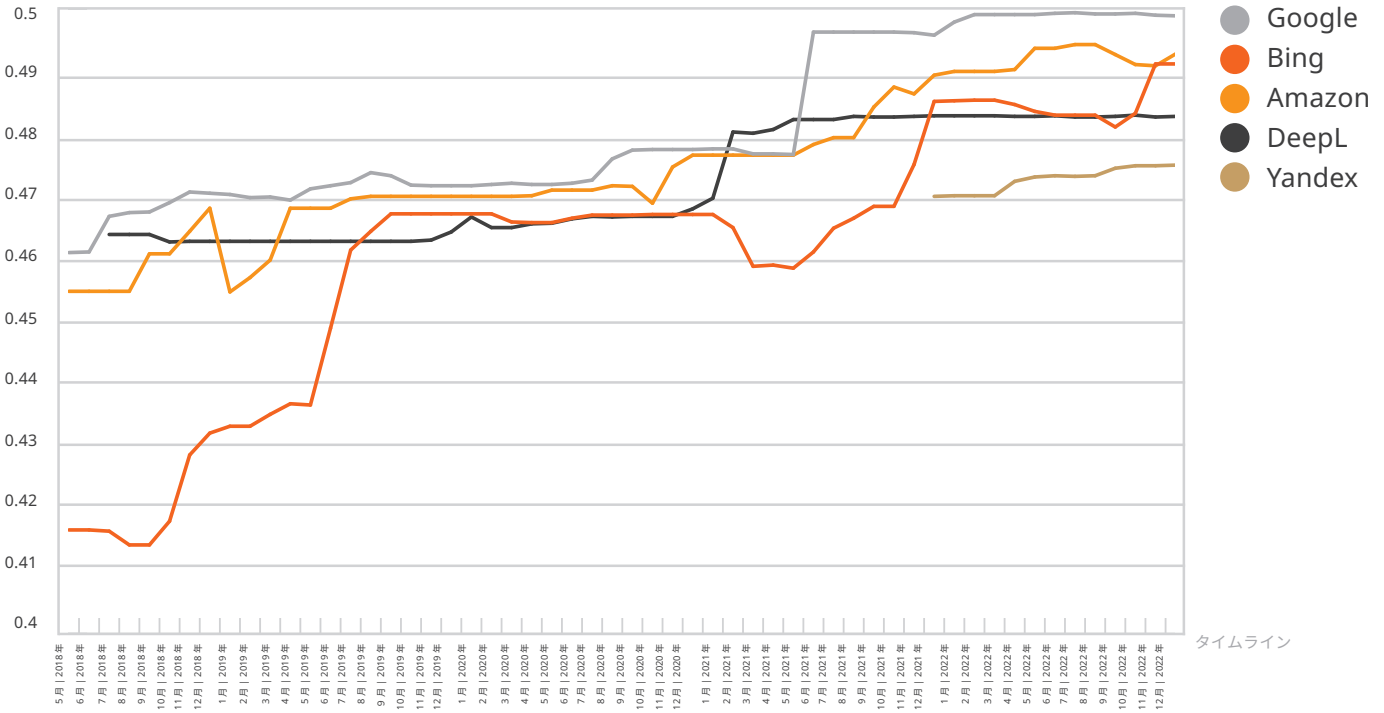


図 4. 逆編集距離に基づく総合的な MT 品質の比較

MT エンジンの性能とトレンドライン

逆編集距離

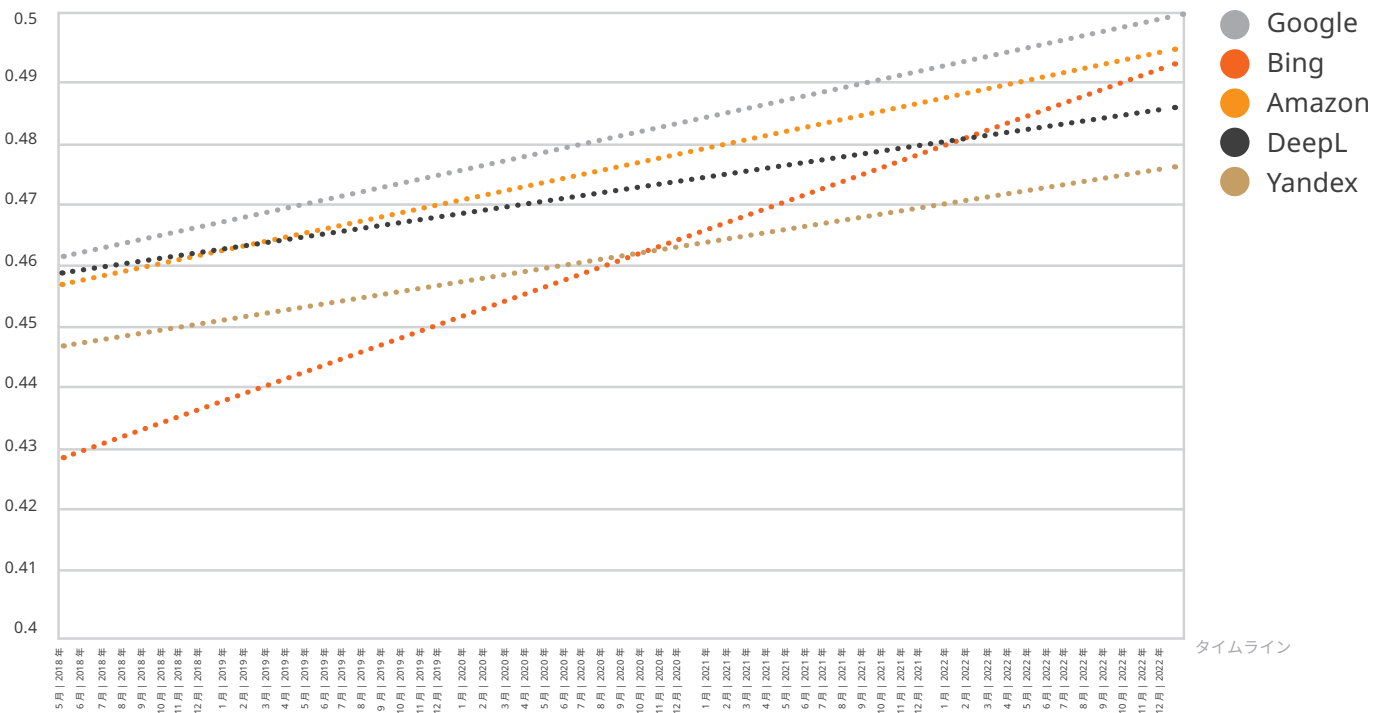


図 5. つの主要な MT エンジンの性能のトレンドライン

言語ごとの機械翻訳エンジンの品質の比較

当社は、2022年時点での主要な MT エンジンによるドイツ語、スペイン語、ロシア語、中国語の翻訳の品質を比較しました。翻訳品質は逆編集距離を使って測定したものです。

編集距離は、MT の出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量を算出するものです。

逆編集距離では、結果の数値が大きいほど品質が優れていることを意味します。

図 6 に示すように、以下の結果が得られました。

- 逆編集距離の測定に使用された尺度にも反映されているように、全体的に MT の進歩はほとんど見られなかった
- Microsoft Bing では、10月から11月にかけてドイツ語、スペイン語、中国語で細かい改善が見られた
- 2022年は変化に乏しい年であった

ニューラル機械翻訳の進歩は停滞期を迎えたと結論付けることができます。MT の品質を大幅に向上させるには、新たなイテレーションが必要となるでしょう。

特定の言語ごとの逆編集距離に基づく機械翻訳エンジンの性能

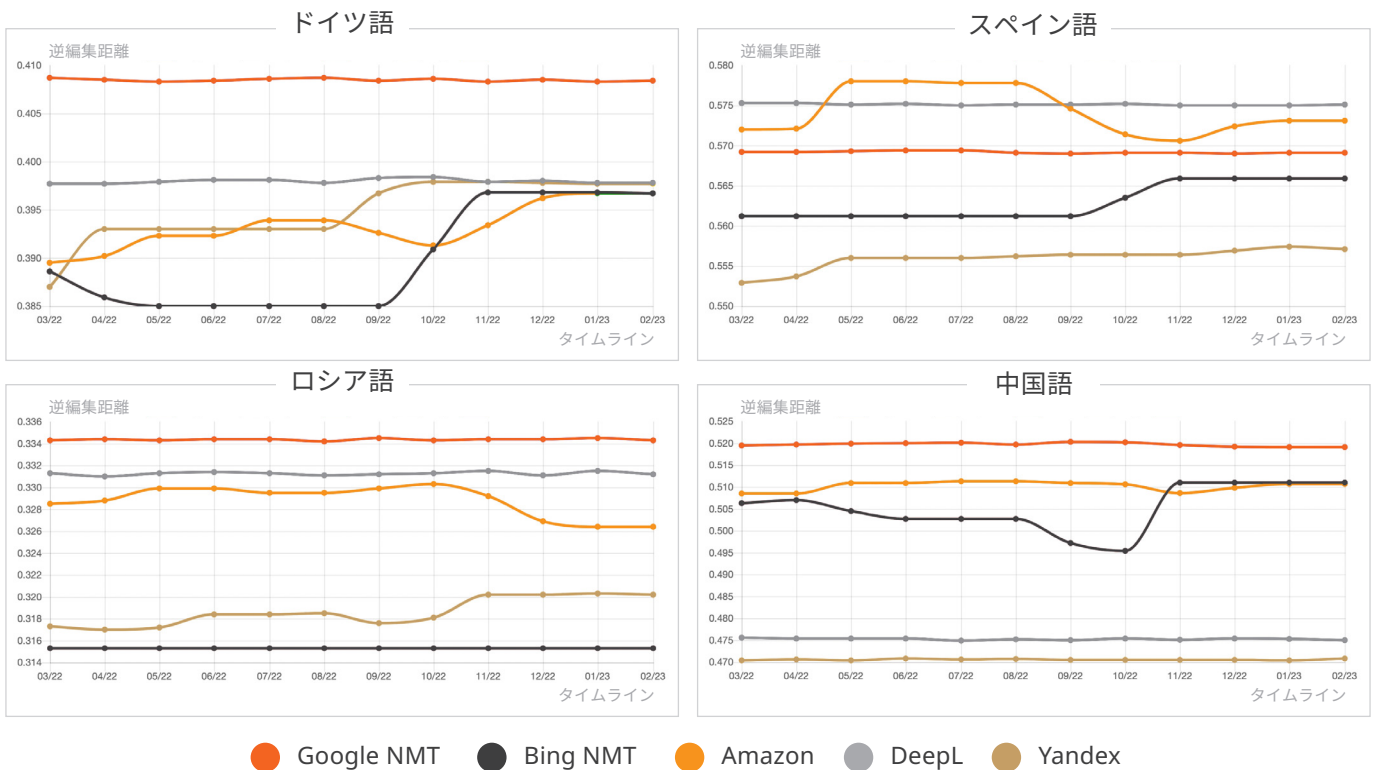


図 6. 逆編集距離に基づく言語ごとの MT 品質の比較

○ ニューラル機械翻訳の進歩は停滞期を迎えたと結論付けることができます。

分野ごとの機械翻訳エンジンの品質の比較

図7は、2022年時点での特定の分野のMTエンジンの翻訳出力品質を、当社が逆編集距離を使って測定・比較したものです。

編集距離は、MTの出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量を算出するものです。逆編集距離は、結果の数値が大きいほど品質が優れていることを意味します。

当社の分析結果は以下のことを示しています。

- **機械翻訳**は、クリエイティブなコンテンツよりも手順などを説明したコンテンツを翻訳する際により優れた性能を発揮します。これは、コンピューターにとって手順などを説明したコンテンツの方がより明瞭で、処理しやすいからです。そのため、リテール、マーケティング、旅行・観光のようにコンテンツがより複雑な業界よりも、自動車やコンピューターソフトウェア業界でMTの翻訳品質が優れていたという結果は当然のものです。このデータは、分野と平均編集距離に基づく自動翻訳の性能を測定した際に得られた結果と類似しており、その結果をさらに裏付けるものです(図2を参照)。
- **Google MT**と**DeepL**は、**Amazon**と**Bing**と比較して、年間を通して安定した性能を示しました。
- 分析期間中、**Amazon**に大きな進歩は見られませんでしたが、上下の動きは数回見られますが、2つの分野(法務・法律、メディア・広告・マーケティング)以外では、測定期間の終わりまでに値が開始時と同じ位置に戻っています。この2つの分野については進化が見られ、法務・法律の分野では競合他社をわずかに上回り、メディア・広告・マーケティングの分野では**Google**とともにランキングをリードしています。

- **Bing**は、親会社であるマイクロソフトに本質的に関連するコンピューターソフトウェアと金融という2つの領域を明らかに牽引しています。Bingはこの1年間で大幅な改善を示し、ランキング3位からトップに躍り出ました。どの分野においても、ランキングでこの水準の上昇を示したMTエンジンは他にありません。
- **Yandex**は、分析対象となったMTプロバイダーの中で唯一、どの分野でも首位を獲得していません。最終的に、2022年末までに、自動車とコンピューターソフトウェアの2つの分野のみで**DeepL**を上回る性能を示しています。



特定の分野ごとの機械翻訳エンジンの性能

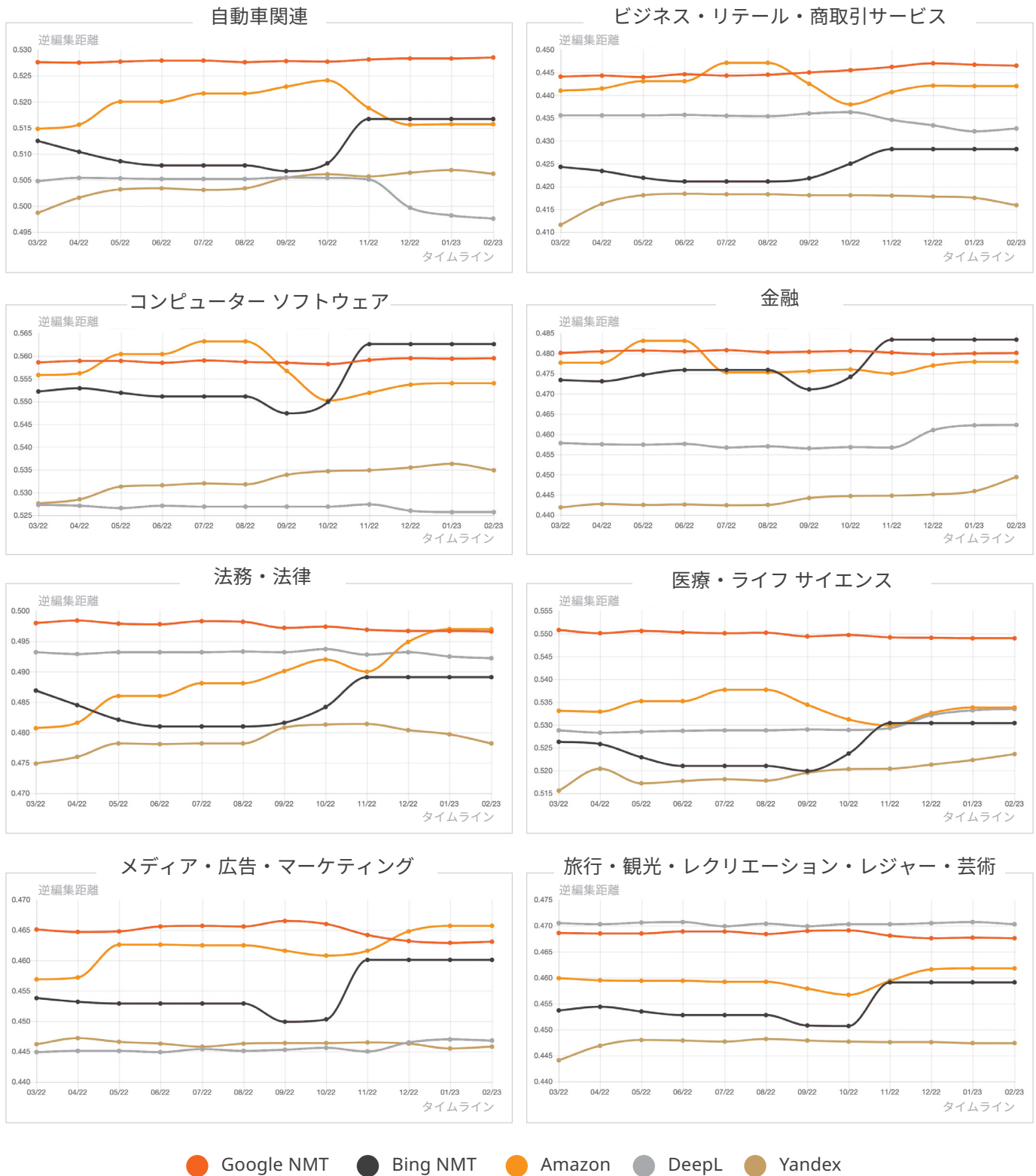


図 7. 逆編集距離に基づく分野ごとの MT 品質の比較

機械翻訳を最大限に活用する方法

MTテクノロジーを最大限に活用するために、以下のステップの実施を検討してみてください。

- MT戦略を策定して開拓すべき市場を決定するために、当社が収集したデータを参照し、MTエンジンによる特定の言語ペアの翻訳しやすさ(言語の機械翻訳可能性)を確認する。
- MTの利用時に問題がよく発生する分野であっても、出力品質の改善に向けて用語集を効果的に使用する。
- MTのカスタマイズまたはトレーニングを実施するタイミングを検討する。どちらの方法でもMT出力は改善できますが、互いにそれぞれを置き換えることはできません。

機械翻訳可能性

特定の言語ペアをMTエンジンで処理するうえでどれだけの課題があるかを明らかにすることは、複数の言語を対象とした翻訳のコストについて予算を検討する際に大いに役立ちます。どの言語ペアの翻訳がより多くの労力を必要とするかを、より深く理解することができます。言語の複雑さについて十分に理解しておくことは、ビジネス上の適切な判断につながるだけでなく、以下のような疑問を解消できます。

- より複雑な言語ペアのポストエディット作業に予算を多めに配分する必要があるか。
- 予算が限られている場合に、一部の言語で、簡易的な「ライト」ポストエディット、または重要な箇所のみを対象とした「フォーカス」ポストエディットで対応できるかどうか。これらのポストエディットをどの言語に対して実施するか。
- 特に低予算のプロジェクトで予算の最適な配分を検討する際、ビジネス要素や文化的な要素に加えて言語ランキングを考慮すべきかどうか。比較的低品質でも許容される場合に、機械翻訳可能性ランキングが低い言語を翻訳すべきかどうか。

翻訳しやすさによって言語にランク付けをするのは簡単なことではありませんが、さまざまな指標を用いて評価を行うことはできます。編集距離(機械翻訳によるテキストを人による翻訳の品質まで高めるために、ポストエディット作業を通じて加えられる変更の量)は、それぞれの言語ペアに関するMTの複雑さと機械翻訳可能性を把握するうえで役立ちます。言語を比較しやすいように、当社では英語から上位28のターゲット言語への機械翻訳可能性をランク付けしました。

表1に示されるように、ポルトガル語、スペイン語、フランス語、イタリア語など、ロマンス諸語の多くを英語から翻訳する際は、少ない変更で高い品質レベルを実現することができます。これらのターゲット言語は機械で処理することが非常に容易であることが明らかになっており、機械翻訳可能性のランキングで上位4位を占めています。

ハンガリー語とフィンランド語はどちらもウラル語族の言語ですが、これらの複雑度はより高く、このランキングでは27位と28位で最下位となっています。同じ語族に属するもう一つの言語であるエストニア語も、複雑度の高い言語の1つです。当社が今まで処理してきた数百万に及ぶ文章に基づくこのような結果は、語族の分類がMTの結果に及ぼす重要性を明確に示しています。

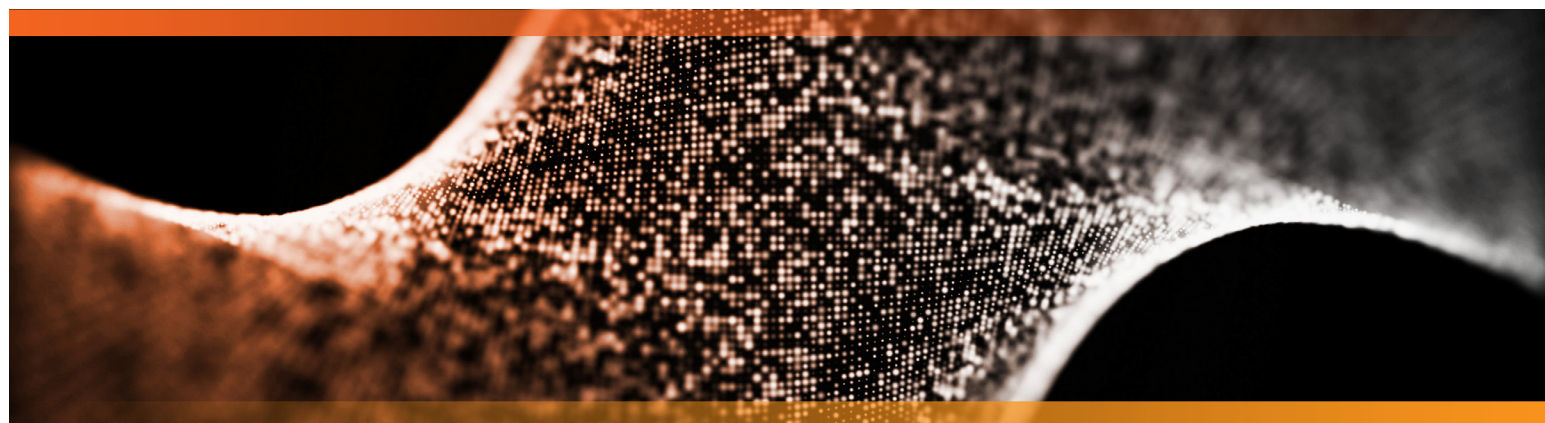
言語の比較には限界がありますが、このランキングによって多言語プロジェクトを管理するうえで興味深いインサイトが得られます。




表 1. 言語の機械翻訳可能性ランキング

ランク	言語 (英語から)
1	ポルトガル語
2	スペイン語
3	フランス語
4	イタリア語
5	中国語 (簡体字)
6	オランダ語
7	デンマーク語
8	日本語
9	ギリシャ語
10	ルーマニア語
11	タイ語
12	ノルウェー語
13	ドイツ語
14	スウェーデン語

ランク	言語 (英語から)
15	トルコ語
16	スロバキア語
17	ヘブライ語
18	ラトビア語
19	ポーランド語
20	中国語 (繁体字)
21	リトアニア語
22	チェコ語
23	アラビア語
24	エストニア語
25	韓国語
26	ロシア語
27	ハンガリー語
28	フィンランド語





○ MT エンジンの十分なトレーニング、用語集のカスタマイズ、前処理と後処理のルールの明確化を組み合わせることで、MT による翻訳において適切な用語が使用され、お客様の文書と同様のスタイルに従った訳文が生成されるようになります。

機械翻訳の品質を向上させるための用語集

先ほど述べたように、用語の観点から見た場合に、汎用型 MT エンジンでは誤った翻訳を出力することがあり、特に、特定の分野で望ましくない結果につながる可能性があります。とりわけ医療や法律の分野では、この影響がマイナスに作用するおそれがあります。用語集を効果的に使用することで、主題がどのようなものであっても MT の品質を向上させ、正確で一貫性のある翻訳を実現できます。

専門用語を含む、特定の分野に特化したバイリンガルテキストを使用して、カスタマイズされた MT システムをトレーニングすることが不可欠です。しかし、特定の分野に特化したテキストでエンジンのトレーニングを行ったとしても、用語が一貫性のある形で使用されていない場合、翻訳の正確性は必ずしも保証されません。この分野の研究では、ニューラル機械翻訳 (NMT) システムに言語情報を組み込むことが提案されています。手動または半自動のアノテーションの導入は、用語集などのリソースの有無や、時間、コスト、アノテーションを担当する人員といった制約に左右されます。

ライオンブリッジの Smart MT では、ソース言語 (翻訳元の言語) とターゲット言語 (翻訳先の言語) のテキストへの言語ルールの適用を可能にし、機械翻訳の用語に対応するための特定のプロファイルに加えられた「翻訳しない

語句 (DNT)」と用語集のリストに基づいて用語を適用できます。当社は、新しい用語の追加や古くなった用語の除外などの用語集の定期的なメンテナンスや、新しい用語集の作成など、用語集に関連するあらゆるニーズに対応しています。また、Smart MT では用語集を一度作成するだけですべての MT エンジンで使用できるようになるため、時間とコストの節約につながります。

MT プロジェクトで用語集を使用するのは、想像するほど簡単ではありません。用語集が不適切に使用された場合、機械翻訳全体の品質にマイナスの影響を及ぼす可能性があります。MT における用語の順守を徹底するための最適な方法は MT のトレーニングです。MT エンジンの十分なトレーニング、用語集のカスタマイズ、前処理と後処理のルールの明確化を組み合わせることで、MT による翻訳において適切な用語が使用され、お客様の文書と同様のスタイルに従った訳文が生成されるようになります。

MT のカスタマイズとトレーニング

MT のカスタマイズと MT のトレーニングは、MT の出力を最大限に活用するために役立ちますが、これらの方法を適用するタイミングを計画的に見極める必要があります。表 2 は、機械翻訳のカスタマイズとトレーニングに関する概要と、それぞれの方法を評価する際に検討すべき事項を示しています。

表 2. 機械翻訳のカスタマイズと機械翻訳のトレーニングの比較

	MT のカスタマイズ	MT のトレーニング
定義と仕組み	機械翻訳により生成される翻訳の精度を高めるために、既存の MT エンジンに用語集や翻訳しない語句 (DNT) リストを取り込んで使用する	機械翻訳により生成される翻訳の精度を高めるために、コーパスと翻訳メモリ (TM) からの膨大なバイリンガル言語データを利用して MT エンジンを構築・トレーニングする
達成される内容	MT が出力する翻訳の精度を高めて、ポストエディットの必要性を低減する	MT が出力する翻訳の精度を高めて、ポストエディットの必要性を低減する
具体的なメリット	適切なブランド名や用語を使用して地域別のバリエーションを用意できる	企業独自のブランド ボイス、トーン、スタイルを維持しながら、地域別のバリエーションを用意できる
使用上のリスク	適切に実行しないと MT 翻訳の品質が低下し、全体的な品質を損なう可能性がある	エンジンをトレーニングするために必要な高品質のデータが十分でない場合は、トレーニングによる十分な成果を得ることができない可能性がある。十分な経験のない作成者がコンテンツで用語を多用しすぎると、MT によって低品質の翻訳が出力され、全体の品質を損ねる可能性がある。
使用すべき状況	<p>技術的なコンテンツや細部を重視するコンテンツ、さらに以下が該当するコンテンツに理想的な選択肢:</p> <ul style="list-style-type: none"> 用語を正確に翻訳する必要がある 地域別のバリエーションを用意する必要があるものの、MT のトレーニングを実施するには十分なデータがない 	<p>専門性の高いコンテンツ、マーケティング資料やクリエイティブなコンテンツ、さらに以下が該当するコンテンツに理想的な選択肢:</p> <ul style="list-style-type: none"> 独自のブランド ボイスやトーン、スタイルを維持する必要がある 地域別のバリエーションを用意する必要があり、MT のトレーニングを実施するための十分なデータがある
成功要因	インプット / アウトプットの標準化ルール、用語集、および DNT を適切に管理できる経験豊富な MT エキスパート	エンジンを十分にトレーニングするために必要な少なくとも 15,000 個の固有のセグメント
コスト上の検討事項	MT エンジンに設定するプロファイルを更新するための一度限りの費用と、用語集を維持・管理するための費用が継続的にかかる。潜在的な利点を考慮すると比較的安価。また、一般的には MT のトレーニングにかかる費用よりも安価で済む。	初回のトレーニングにかかる費用に加え、MT の性能をモニタリングする中で、改善の余地があると判断された場合に追加のトレーニング費用がかかる可能性がある。潜在的な利点を考慮した場合に、特定のケースでトレーニングを実施する投資価値がある。

結論と機械翻訳の展望

機械翻訳を取り巻く状況について、2022年のデータと品質性能の進歩の停滞が示された驚くべき分析結果からはどのような結論が得られるのでしょうか。機械翻訳テクノロジーは成熟しており、ビジネスレベルのテクノロジーとしての価値が明確に証明されているため、今後も普及は進んでいくでしょう。

人が介入するかどうかや、ハイブリッドのアプローチを採用するかどうかにかかわらず、ほぼすべての翻訳事例で機械翻訳が有用であることが認識されています。Global Market Insights のレポートによると、翻訳市場は、2022年から2030年にかけて年平均成長率(CAGR)30%で成長すると見込まれています。ゲーム業界やライフサイエンス業界など、これまではMTの導入に抵抗があった業界でも、企業によるMTの利用は増えていくでしょう。ワークフローと翻訳者の選択を自動化するAI主導型テクノロジーと組み合わせてこのテクノロジーを十二分に活用することができれば、企業はコンテンツ制作のスピードを向上させ、常にブランドのイメージに合った魅力的な多言語コンテンツを制作して市場を拡大し、非常に競争の激しいデジタル市場で成功を収めることができます。

機械翻訳の未来

2022年の機械翻訳を取り巻く状況から、現在のニューラル機械翻訳について次のような疑問が生まれました。

- NMTの進歩は停滞期に入っているのか。
- 従来のエンジンで大きな前進ができないと考えると、新たなパラダイムシフトが必要なのではないか。
- その後何が起こりうるのか。

当社では、マルチモダリティと多言語を含む膨大なコンテンツを活用する大規模言語モデル(LLM)が、将来のパラダイムに関わりを持つと予想しています。

そう考えるのは、ChatGPTとMTエンジンの翻訳品質を比較した画期的な分析結果があるからです。

OpenAIのChatGPTは、指定したMTエンジンよりも結果は劣っていましたが、大差はありませんでした。GPT-4は、1つの分野と1つの言語ペアにおいて、ある主要なニューラル機械翻訳エンジンを上回る性能を示しました。これらの結果が機械翻訳の未来に影響を与えるのは疑いのないところです。

新しい機械翻訳パラダイムが進行していると思われる理由

現在のMTエンジンの傾向を見るにつけ、既視感に襲われます。NMTに取って代わられた統計的機械翻訳時代の終末期、MTの品質は変化がないも同然で、各種のMTエンジンの品質が同等の水準に収束していました。現在がまさにこの状況です。

NMTがすぐにも取って代わられることはないかもしれませんが、しかし、指数関数的成長と収穫加速の理論が正しいとして、ルールベースMTの30年の運用と統計的MTの10年にわたる隆盛を考慮し、さらにNMTが登場から6年目であることを踏まえると、新たなパラダイムシフトは近づいています。

次の機械翻訳パラダイムとなりえるのは何か

2022 年中の LLM の重要な進歩によって、このテクノロジーが 2023 年に参入する下地は整いました。LLM は、さまざまなことを実行できるようトレーニングされた汎用モデルです。しかし、2022 年末までに、何らかに特化した (微調整された) LLM が、いくつかの特定の分野できわめて重要な進歩を遂げたことがわかりました。つまり、何らかの追加トレーニングを行えば翻訳を実行できるポジションに立ったのです。

たとえば、ChatGPT の場合、OpenAI はこのモデルを調整して、質問応答の対話ができるようにしましたが、このモデルは汎用 LLM で実行可能なことは引き続き何でも実行できます。OpenAI は GPT-4 モデルでさらに改良を加えており、時間の経過とともにより多くの LLM の調整が行われることが予想されます。

大規模言語モデルでの翻訳処理に必要な調整方法

よりバランスの取れた言語コーパスで機械がトレーニングされれば、翻訳に LLM を使用できる可能性はさらに高まります。GPT-3 のトレーニング コーパスは 93% が英語で、それ以外のすべての言語のコーパスはわずか 7% です。GPT-4 にはより多くの言語のトレーニング データがあると推測すると、出力品質が向上した理由を説明できます。言語バランスの取れたコーパスは、翻訳に特化している LLM に加えて、調整モデルを構築するベースになる可能性があります。

この LLM をベースとした仮説としての新たな MT パラダイムには、興味深い側面がもう一つあります。マルチモダリティの傾向です。当社は、LLM のトレーニングに、言語データに加えて画像や動画といった他のトレーニング データを使用することがあります。このタイプのトレーニングによって、より質の高い翻訳につながるような、世界についての知識がさらに提供される可能性があります。

大規模言語モデルがニューラル機械翻訳に取って代わるパラダイム シフトが起こるのか

LLM が NMT に取って代わるものとして有望なのかを評価するために、MT 品質トラッキングで当社が使用している 5 つの主要 MT エンジンと、ChatGPT と GPT-4 の翻訳品質を比較しました。図 8 に示すように、ChatGPT は特化型エンジンに匹敵する結果を残しました。図 9 で示されているとおり、GPT-4 は、英語から中国語への言語ペアについて Yandex よりも優れた翻訳品質を示しています。

ChatGPT と汎用型 MT エンジンの品質比較の評価方法

エンジンの品質レベルは、複数の参照訳をサンプルとして、逆編集距離に基づいて算出しました。編集距離は、MT の出力に対して、翻訳者による翻訳品質と同等にするために人間が行わなければならない編集の量を算出するものです。今回は翻訳者による翻訳 1 種類のみではなく、10 種類、つまり複数の参照訳と、MT のみの出力を比較しました。逆編集距離では、結果の数値が大きいほど品質が優れていることを意味します。

言語バランスの取れたコーパスは、翻訳に特化している LLM に加えて、調整モデルを構築するベースになる可能性があります。



逆編集距離に基づく複数参照評価

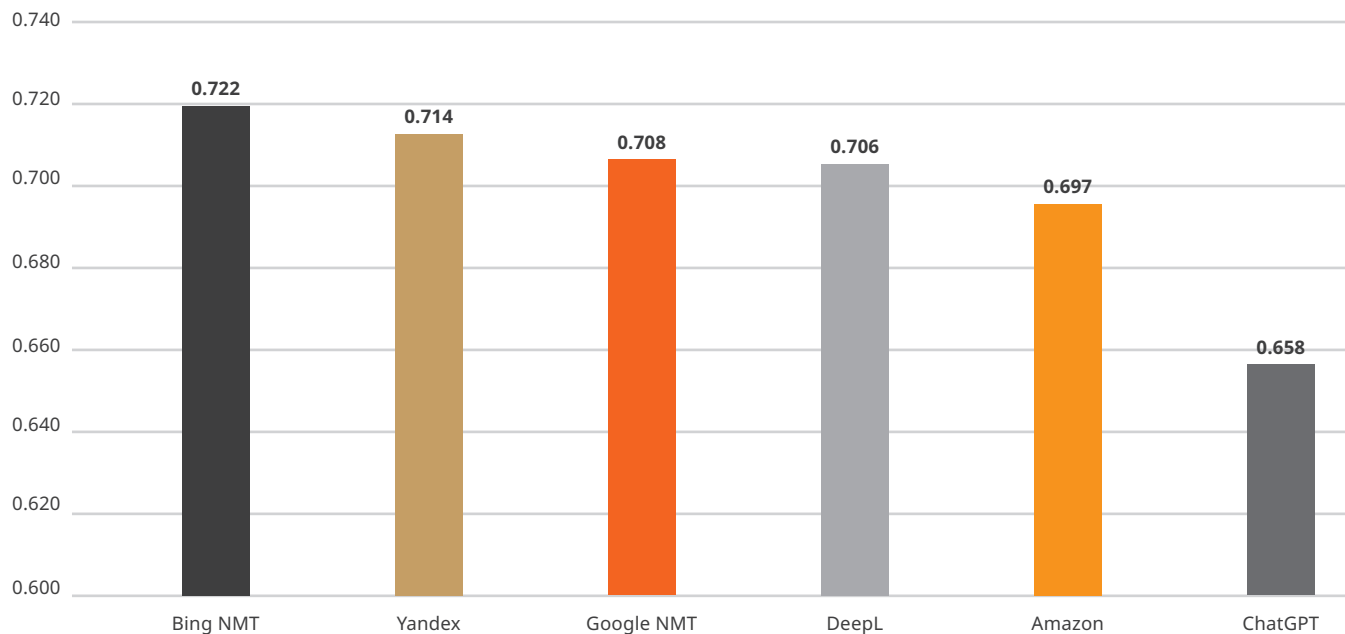


図 8. 英語からスペイン語に翻訳する言語ペアの複数の参照訳を使用した逆編集距離に基づく、ChatGPT と主要機械翻訳エンジンによる自動翻訳品質の比較

英語から中国語の翻訳品質

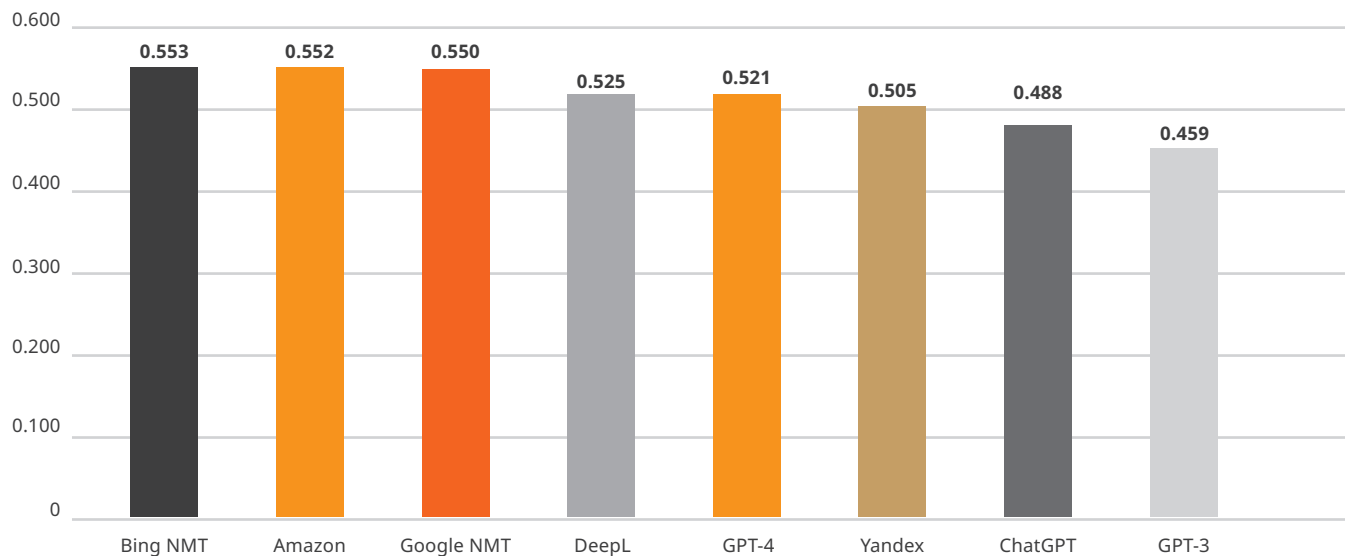


図 9. 英語から中国語に翻訳する言語ペアの複数の参照訳を使用した逆編集距離に基づく、GPT モデルと 5 つの主要ニューラル MT エンジンによる自動翻訳品質の比較



これらの大規模言語「汎用」モデルの優れた点は、さまざまなことを実行でき、それらのタスクのほとんどで優れた品質を示せることです。

LLM 翻訳結果が注目に値する理由

比較分析の結果は目を見張るものです。というのも、MT エンジンが翻訳という単一の自然言語処理 (NLP) タスクに特化してトレーニングされているのに対し、汎用モデルは多種多様な NLP タスクを実行するようトレーニングされているからです。GPT は翻訳の実行用にトレーニングされているわけではないのに、その翻訳品質は非常に優れています。

大規模言語モデルの影響を受けた機械翻訳の進化

世間の関心とこの技術へのテクノロジー企業の大型投資を踏まえて LLM の成長を考えると、MT が新しい LLM のパラダイムを採用するようになるかがわかるまでに、それほど時間はかからないかもしれません。

LLM が MT のベースとして使用される可能性はありますが、その際は機械翻訳向けにテクノロジーのチューニングが行われる可能性があります。これは、機械と人間の会話によるコミュニケーションを可能にするといった特定のユースケース向けに汎用モデルを向上させるべく、OpenAI や他の LLM 企業が行っていることと同様のことです。特化することで、実行するタスクの精度が高まるのです。

大規模言語モデルの総体的な展望

これらの大規模言語「汎用」モデルの優れた点は、さまざまなことを実行でき、それらのタスクのほとんどで優れた品質を示せることです。たとえば、DeepMind の GATO は別の汎用 AI モデルですが、600 以上のタスクでテストされ、うち 400 で「現時点での最先端レベル (SOTA: State-of-the-Art)」という結果を示しました。

GPT や GATO のような汎用モデルと、これらの汎用モデルをベースに特定の目的のために特化したモデルという 2 つの開発ラインは今後も存続するでしょう。汎用モデルは、汎用人工知能 (AGI) を進化させるうえで、そしておそらくは長期的な、より目覚ましい発展を進めるうえで重要です。そして特化型モデルは、特定の分野で短期的に実用化されるでしょう。

LLM で注目すべきことのひとつに、両方のラインが並行して同時に進化・機能できる点があります。

機械翻訳におけるパラダイムシフトの意味

現在のニューラル機械翻訳テクノロジーは限界に達し、有力な機械翻訳テクノロジー（おそらくは LLM をベースとするもの）が新たに登場しているため、MT 分野で何らかの変化が起こることが予測されます。それによる影響の多くは企業にとってメリットとなりますが、人による翻訳を求める企業にとってはいくつか課題が生じると考えられます。

以下でその課題を説明します。

品質向上

技術的な進歩によって、言葉遣いのフォーマルさの度合いやトーン（語調）に関する品質の問題といった長年の問題を解決できるので、機械翻訳の品質は飛躍的に向上します。LLM は、MT エンジン最大の問題である「世界についての知識」の不足すら解決できるかもしれませんが、これは、マルチモダリティトレーニングを通じて達成できる可能性があります。

テクノロジストは、最新の LLM のトレーニングに大量のテキストだけでなく、画像や動画も使用します。この種のトレーニングによって、LLM は、機械がテキストの意味を解釈するうえで役立つ、より関連性のある知識を持てるようになります。

コンテンツの増加と優秀な翻訳者の供給不足

企業はより多くのコンテンツを迅速に制作できるようになりましたが、現状では、そのペースがそれを翻訳する能力がある翻訳者数の増加を上回っています。MT と翻訳者の生産性の向上をもってしても、現存の翻訳コミュニティが翻訳需要を満たすのは難しい状況が続くでしょう。

機械翻訳の導入の増加

新しいテクノロジーが利用できるようになり、機械翻訳の品質も向上しているため、翻訳サービスの需要は成長を続け、より多くの状況やユースケースで導入が進むと思われる。

機械翻訳の使用によるカスタマー エクスペリエンスの向上
MT 品質の向上と、よりパーソナライズ / カスタマイズされたカスタマー エクスペリエンスのニーズを背景として、企業は MT の使用を増やして、世界中の顧客のデジタルエクスペリエンスを向上させ、強い信頼関係を築くようになるでしょう。



LLM は、MT エンジン最大の問題である「世界についての知識」の不足すら解決できるかもしれません。

結論

テクノロジー企業は、LLM テクノロジーに多大な関心を示しています。OpenAI に対するマイクロソフトの投資額は100億ドルです。NVIDIA、Google、その他の企業も、LLM や AI テクノロジーに多額の投資を行っています。GPT とローカリゼーション、そして言語サービス業界に価値をもたらす大規模言語モデルの真の可能性について理解を深めるには、以下のような作業や準備が必要になります。

- 実際の状況に即したテストを大規模に実施し、ローカリゼーション・編集の各タスクについてエラー発生率の評価を行う。
- ローカリゼーションのバリューチェーン内におけるユーザージャーニーをマクロとミクロの両面から詳細に分析し、このようなタイプのテキスト自動生成で混乱が生じる可能性が高いポイントを特定する。
- 大規模なプロジェクトでのプロンプトの使い方、関連するコンテキストを GPT や LLM に渡す方法を理解し、注意すべき点やベストプラクティスを文書化する。
- 自動化と人間による編集作業が混在するワークフローを新たに確立し、このような AI を取り入れることでポストエディットや QA が今後どのような役割を果たしていくのかを再考する。
- 改善の可能性があればその都度、ローカリゼーションエージェント向けと顧客向けの両方について、新しい自動化機能とユーザーエクスペリエンス (UX) のインタラクションコンテキストを設計する。
- ライセンス、導入、保守のコストについて、自社のビジネスにおいて経済的合理性があるかどうかを確認する。

当社では今後のトレンドや動向についても興味深く見守ってまいります。引き続き LLM を評価して皆さまと有益な情報を共有できるよう、この心躍る進化の最新情報をお届けいたします。

テクノロジー企業は、LLM テクノロジーに多大な関心を示しています。



ライオンブリッジ機械翻訳トラッカー専門家



ラファ モラル

イノベーション担当バイス プレジデント

機械翻訳、コンテンツのプロファイリングと分析、用語の抽出、言語品質の保証と管理などの言語・翻訳に関わる研究開発を指揮しています。



ヨランダ マーティン

MT 専門家

カスタマイズされた翻訳モデルの構築、品質分析、チューニング戦略の開発を担当しています。また、研究開発部門と協力し、新しい言語ツールやリソースの開発も行っています。



トーマス マッカーシー

MT ビジネス アナリスト

ライオンブリッジのお客様とその関係者に MT 関連のテクノロジー、サービス、コンサルティングの利益を最大限にお届けできるようサポートします。



ライオンブリッジの専門家は、お客様が機械翻訳を効果的に導入してビジネス上の利益を獲得できるよう支援をご提供しています。ライオンブリッジでは、機械翻訳業界が急速に変化していく中で、お客様にさらなる利益を提供すべく新しいテクノロジーを綿密に評価しています。

自動翻訳を最大限に活用するための当社の各種サービスについては、ぜひ当社までお問い合わせください。

ライオンブリッジについて

ライオンブリッジはさまざまなお客様のために架け橋を築き、障壁を打ち破ってあらゆる課題を解決に導きます。25年以上、350を超える言語で翻訳とローカリゼーションのソリューションを提供し、お客様の企業が世界中の顧客とつながるための支援をしてきました。当社独自のワールドクラスのプラットフォームを通じて、世界各地で活躍する多数の専門家のネットワークを活用し、さまざまなブランド・企業と提携して文化的に豊かなエクスペリエンスを生み出しています。言語のプロフェッショナルとして、ライオンブリッジは優秀な人材と洗練されたマシンインテリジェンスの能力を活用し、お客様の顧客の心に響くメッセージをお届けします。米国マサチューセッツ州ウォルサムに本社を置くライオンブリッジは、世界24か国にソリューションセンターを設けています。

詳しくはこちら

LIONBRIDGE.COM

LIONBRIDGE

© 2023 Lionbridge. All Rights Reserved.