

ISSN 1883-1818

No.76

June 2022

AAMT Journal

Asia-Pacific Association for Machine Translation

機械翻訳

機械翻訳



目 次

巻頭言

翻訳の日	安達 久博	3
------------	-------	---

解説記事

「日本の大学における教養英語教育と機械翻訳に関する予備的調査」に見る英語教員の現在と未来	小田 登志子	4
MTに関する法的問題について	AAMT	10

事例・研究

機械翻訳と人間翻訳の両方の市場を最大にする方法	ローズ三浦	13
Memsource機械翻訳調査レポート2022年第2四半期	Memsource日本窓口	15
産業翻訳における機械翻訳の継続的改善手法	中山 雄貴	18
最適化された機械翻訳導入モデルによる医薬品開発業務の迅速化への試み	早川 威士、鞠子 幸泰	22
「カスタムMT+フルPE」現状と見通し：医療翻訳分野の事例研究	山田 優、早川 威士	26
特許翻訳におけるドメイン適応型機械翻訳	本間 奨、金子 清隆、 東内 健、江崎 美江、 大崎 慎治	31

イベント報告

AAMT2021, Online 創立30周年記念イベント	出内 将夫	35
第5回自動翻訳シンポジウム 参加報告	東山 翔平	42

編集後記

.....	中澤 敏明	45
-------	-------	----

C O N T E N T S

Foreword

Translation Day	Hisahiro Adachi	3
-----------------------	-----------------	---

Commentary

A review of Yamada et al. (2021): The present and future of general English education in Japan	Toshiko Oda	4
Legal Issues regarding MT	AAMT	10

Case Study

New Translation Market Cannot Evolve Without Interactions Between Humans and Machine	Rose Miura	13
Memsource Machine Translation Report Q2/2022	Memsource Japan	15
Continuous Improvement Method for Machine Translation in Industrial Translation	Takeyoshi Nakayama	18
Optimized Acceptance Model of Machine Translation Technology to Accelerate Process of Drug Development	Yukiyasu Mariko, Takeshi Hayakawa	22
Custom MT + PE in medical domain: A case study.	Masaru Yamada, Takeshi Hayakawa	26
Domain-adaptive machine translation in patent translation	Susumu Honma, Kiyotaka Kaneko, Kenichi Tohnai, Mie Ezaki, Shinji Osaki	31

Event Report

AAMT 2021, Online --- Forefront of machine translation ---	Masao Ideuchi	35
Report on the 5th Symposium on Automated Translation	Shohei Higashiyama	42

Editor's Note

.....	Toshiaki Nakazawa	45
-------	-------------------	----

翻訳の日

安達 久博

一般社団法人日本翻訳連盟 (JTF) 代表理事・会長

日本翻訳連盟 (JTF) は日本における翻訳・通訳者、翻訳・通訳会社、翻訳・通訳支援ツール開発提供会社、クライアントからなる代表的な業界団体です。おかげさまで昨年、創立 40 周年を迎え、世界平和と翻訳・通訳業界の持続可能な発展と健全化を目指しています。創立当初の資料をひも解くとそこにあった言葉は

「だれにも、どこでも役立つ翻訳を Translation, a Pervasive presence」

翻訳は、他国の工業製品を買い手に魅力を感じさせ、その品質保証をより完全なものにします。また、危険な物質の容器に印刷された注意書きによって人の命を救っています。さらに、よくできた翻訳は、メーカーのイメージを高めます。このように、翻訳は、ユーザーとしてのあなたやメーカーとしての企業に貢献し、安全をも保障しているのです。いま世界は、他国に比べて豊かな日本企業の動きを注視していて、ことあるごとに批判の声を浴びせてきます。文化の異なる他国とさらに健全な交流を進めるため、ここで私たちは、単なる「ことば」や「こと」の翻訳から、「こころ」を正しく自然に翻訳するように努め、それによって日本国や日本人がほんとうに外国人から同じ地球上の同胞として愛されるようにならなければなりません。

翻訳の日を記念して、広く産業界一般に翻訳の重要性を訴えるため、実務翻訳者の関心の深いテーマを主としたプログラム構成としました。(第 2 回 JTF 翻訳の日記念行事 1993 年 10 月 2 日開催チラシより)

「翻訳の日」とは 1991 年に国際翻訳家連盟が聖書をラテン語に翻訳した聖ヒエロニムスの命日が 420 年 9 月 30 日であることから「世界翻訳の日」と制定しました。その後、2017 年 5 月に国連総会にて「国どうしを結び、平和と理解、発展を育むうえで専門の翻訳者が果たす役割」を重視する決議が採択され、9 月 30 日が

「国際翻訳デー (International Translation Day)」と公式に認定されました。さらに、国内では 2019 年 5 月 20 日に日本記念日協会より正式に「翻訳の日」が記念日として認定登録されました。

この「翻訳の日」を祝うイベントとして現在の「JTF 翻訳祭」は昨年、第 30 回を迎えることができました。特に、2018 年に京大・芝蘭会館で開催された第 28 回 JTF 翻訳祭の交流会では、アジア太平洋機械翻訳協会の創立者である長尾真京大名誉教授に乾杯のご発声を賜りました。そして JTF の創立者のひとりでもある勝田美保子 JTF 元会長と壇上で握手を交わされたのが印象に残ります。

そして、その翌日には長尾先生の文化勲章受章が決定したとのニュースが流れ、会場は興奮に包まれました。不思議なご縁、巡り合わせを感じました。

また、翌 2019 年には、パシフィコ横浜で開催された第 29 回 JTF 翻訳祭の交流会でも乾杯のご発声を賜りました。悲しいことに、これが長尾先生、勝田元会長とお会いする最後の機会となってしまいました。

長尾先生は当初から、機械翻訳は決して人間が行う翻訳と対立するものではなく、双方がうまく協調、協同してより良い翻訳を世の中に提供し、世界平和に貢献していくべきものとおっしゃっていました。

ニューラル機械翻訳方式になって機械翻訳の翻訳精度は飛躍的に向上しつつあります。しかしながら、翻訳の品質保証を行うわけではなく、翻訳の品質保証は人間にゆだねられています。文化の異なる他国の人々に単なる「ことば」や「こと」の翻訳から、伝えるべき「こころ」を正しく自然に翻訳するように努めることが肝要であり、機械と人間の協創・協力を夢見た先人の思い、理念を DNA として引き継ぎ、気持ちも新たに前に進む原動力として行きたいと思います。

Translation Day

Hisahiro Adachi

Representative Director & President, Japan Translation Federation (JTF)

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Public License. License details: <https://creativecommons.org/licenses/by-sa/4.0/>

「日本の大学における教養英語教育と機械翻訳に関する予備的調査」に見る英語教員の現在と未来

小田 登志子

東京経済大学全学共通教育センター 准教授

1. はじめに

「ニューラル機械翻訳は困ったもんだ」と本誌に投稿したのはトム・ガリー氏である^[1]。ガリー氏と同じ大学英語教員としてこの発言に対して大いに共感を覚える。しかし一方、筆者には別の気持ちがあることも確かである。それは「ニューラル機械翻訳はチャンスかもしれない」という期待である。グローバル化の掛け声もむなしく、英語を使いこなせる日本語話者はそれほど増えていない。筆者が担当する教養英語の授業を履修する学生のほとんどは英語に対してあきらめを感じている。それなら、機械翻訳によって英語が苦手な日本人にブレイクスルーをもたらすことはできないのであろうか。

2016年にGoogle翻訳にニューラル機械翻訳が搭載されて以来、筆者はこの相反する二つの感情を抱えながら2022年の今に至っている。そして、自分と同じような日本全国の大学英語教員は、このニューラル機械翻訳というゲームチェンジャーに対してどのように対峙しようとしているのか関心を持ってきた。

しかし筆者の疑問が解消されることはなかった。2016年以降、マスコミでは機械翻訳が大きく取り上げられ、英語の勉強に苦勞するよりも「助っ人（機械翻訳）を頼ればいい」^[2]といった論調が目立つようになった。しかしそれとは対照的に、英語教員からの反応は皆無に等しいという時期がしばらく続いた。筆者にはこの奇妙な沈黙が英語教員のとまどいを反映しているように思えた。

そうするうちに2020年のコロナ禍が到来し、ようやく沈黙が破られることになった。新型コロナウイル

ス感染症拡大防止のために、日本の大学のほとんどでコンピュータを介した遠隔授業が実施されるようになった結果、学生の機械翻訳の使用が明るみになったからである。「学生の作文がいやにうまい」「学生が提出する英語が似通っている」など、機械翻訳の影響と思われる事例を報告する声が次々と教員からあがるようになった。

自らの疑問を解消するのに良い時期が来たと感じた筆者は、山田ら（2021a）「日本の大学における教養英語教育と機械翻訳に関する予備的調査」というアンケート調査^[3]の実施に参加した。このアンケート調査はコロナ禍に見舞われた2020年度の終わりである2021年3月に実施された。その結果には大学英語教員の機械翻訳に対するとまどいと好奇心が色濃く反映されている。

本稿ではこの山田ら（2021a）によるアンケート結果の解説を行うことを通して、ニューラル機械翻訳によって揺さぶられる英語教員の姿を紹介するとともに、英語教育の今後について考えてみたい。

2. アンケートの概要と主な結果

まず山田ら（2021a）で報告されたアンケート調査の概要を述べる。このアンケートの目的は日本の英語教員の機械翻訳に対する意識を調査することである。アンケートを実施したメンバーが知る限り、教員を対象として行われたこのような調査は日本国内では初めての事例である。対象者は中学・高校・大学で英語を教える教員である。アンケートの参加者は実施メンバーが所属する教育機関や学会を通じて募集し、グー

グルフォームを用いて実施され、のべ 133 名からの回答を得た。

山田ら (2021a) で報告された結果は、この 133 名のうち教養英語を担当する大学教員 60 名の回答を抜粋して集計したものである。(133 名分の回答を調査対象とした報告については別の資料^[4]を参照のこと。) なお、この 60 名の全員が日本語を母語とする英語教員であるため、以下の内容を読む際に留意されたい。

全部で 26 問から成るアンケートの回答を見ると、回答者が担当する学習者は TOEIC400 点程度の英語力の学生が多い。回答者の機械翻訳リテラシーは高いとはいえない。日常生活で機械翻訳を時々あるいはしばしば使用すると答えた回答者は 19 名 (約 32%) であった。そして、アンケートで提示した DeepL による訳例に対して「こんなに機械翻訳が発達しているとは思わなかった」という感想が多数寄せられた。

次にいくつかの設問を選んで解説してみたい。まず学生は授業 (自宅学習を含む) で機械翻訳を (たまに・頻繁に) 使っていると思うかという質問に対して、53 名 (約 88%) がそう思うと回答した (図 1)。

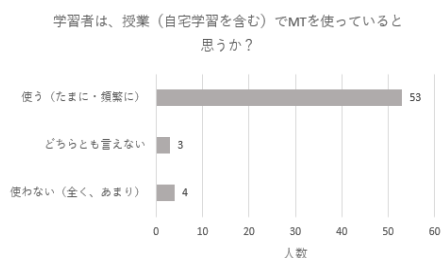


図 1 : 学習者の授業での MT の使用頻度 (教員の主観評価)

では、機械翻訳の実力についてどのようにとらえているのか尋ねたところ、「機械翻訳は学生より (圧倒的に・やや) 優れている」と回答した教員が 44 名 (約 73%) にものぼった (図 2)。したがって、学生が機械翻訳を使えば学生自身の実力以上の英語が出力されることを教員は知っているということになる。

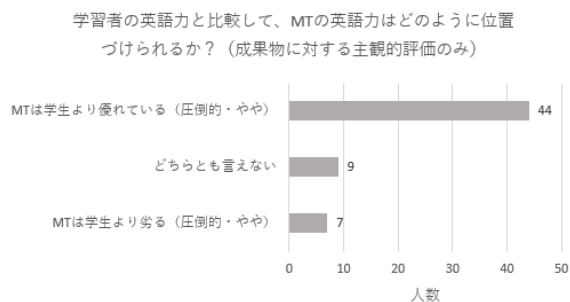


図 2 : MT vs. 学習者の英語力

では、学生が授業 (自宅学習を含む) で機械翻訳を使うことを禁止しているかどうか質問したところ、「禁止も許可もしていない」という回答が 38 名 (約 63%) で最も多かった (図 3)。

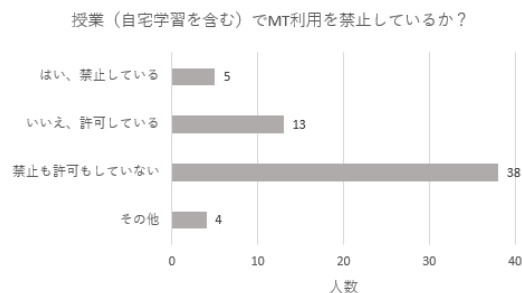


図 3 : 授業での MT 利用の許可と禁止の状況

これは非常にまずい状況である。教員は学生の機械翻訳使用に気づいており、かつその機械翻訳の実力は学生の英語力よりも高いと思っている。にもかかわらず、学生に何も指示をしていない。つまり、学生の英語力を実力以上に評価している可能性を否定できない。このような「野放し」の危険な状態を発見できたことは、このアンケート調査の大きな収穫の一つである。

もちろん機械翻訳を「禁止している」とはっきり答えた教員も 5 名 (約 8.3%) いる。しかしこの回答にも疑問を持たざるを得ない。教室活動はともかく、自宅で行う課題学習で機械翻訳の使用を禁止するのは困難である。「機械翻訳を使って提出した作文は 0 点にする」という発言を学会で聞いたこともあるが、学生が機械翻訳を使用したかどうかどのように判断するのであろうか。学生一人一人の英語力やクセを把握できる少人数の授

業でない限り、学生の機械翻訳使用について確信を持つのはかなり難しい。確たる証拠もないのに学生に対して「機械翻訳を使いましたよね？」と詰め寄ることもできない。そのような事をしたらアカデミック・ハラスメントとみなされてしまいそうである。

では学生に対して「機械翻訳の使用を許可している」と答えた 13 名(約 22%)は問題ないかという、そのように考えるのも時期尚早である。学生が機械翻訳を使用する場合、教育内容を工夫しない限り学生の本当の英語力を把握しにくいからである。筆者も手元にある英語テキストを見てみたが、機械翻訳でやすやすと答えられる問題がある。英作文の問題などは、機械翻訳による出力と模範解答がほぼ一致していたりする。残念ながらこのアンケートでは追跡調査を行っていないため、この 13 名がどのようなポリシーに基づいて機械翻訳の使用を許可しているのか知ることができなかった。

ただし、学生による機械翻訳の使用を（不本意ながら）黙認している教員が「学生が機械翻訳を使用していると気づき、英作文の配点を少なくした」と発言したのを学会で聞いたこともある。筆者自身は、自宅学習における機械翻訳の使用を禁止していない。そして、英作文の宿題は提出の有無のみをチェックし、英語そのものの評価は教室で行われるテストや学生の英語スピーチに基づいて行うようにしている。

ここでアンケート回答者の典型例である教員に話を戻したい。機械翻訳をあまり使用せず、学生の機械翻訳使用に対してほとんど何も言っていない英語教員は機械翻訳に対して関心がないのかというと実はそうではない。機械翻訳を学生の外国語学習に活用してみたいと思うかどうか問うたところ、活用してみたいという回答が 34 名（約 57%）で最も多かった（図 4）。このような機械翻訳に対する関心の高さは筆者の想像を上回るものであった。ただし、機械翻訳に興味を持つ英語教員が率先してアンケートに協力した結果である可能性も否定できない。

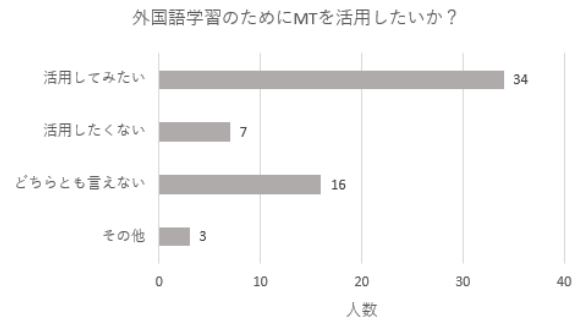


図 4：MT を活用してみたいと考えている教員数

機械翻訳を活用してみたい教室活動としては、ライティング (36 名 60% 複数回答可) が最も多かった（図 5）。スピーキング (23 名 約 38% 複数回答可) についてはプレゼンテーションの作成に絡めて活用してみたいという意見もあった。

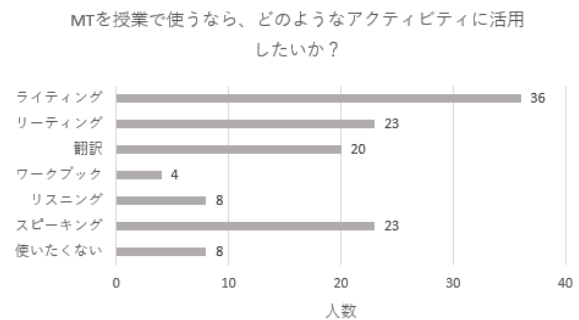


図 5：MT の授業利用における活用アイデア

ただし、関心はあっても実際に授業で機械翻訳を活用するには倫理的な問題をクリアしなければならない。実際、機械翻訳使用は不正行為に当たるのではないかという懸念が教員と学生の双方からしばしば挙がっている。何が不正に相当するかは意見が分かれるところだと思われるため、具体的に「エッセイ課題で、学生がネット上の記事やブログを勝手にコピーし機械翻訳にかけて提出することは著作権の侵害の問題となる可能性がある。この問題を気にするか？」と尋ねたところ、「気にしている」という回答が 51 名（85%）と圧倒的に多かった（図 6）。

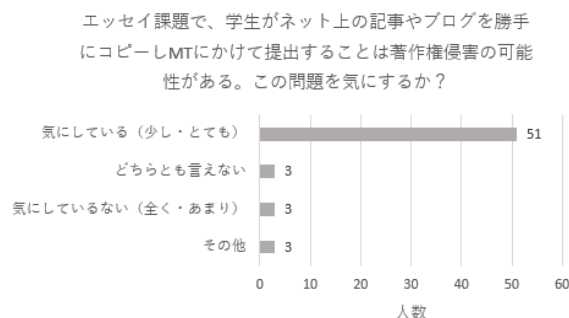


図 6：倫理・著作権の問題意識

このアンケートには最後に自由記述欄を設けたが、主に次のような意見が寄せられた。

- ・学生の機械翻訳使用を防ぐことはできない。
- ・機械翻訳を上手に活用すべき。
- ・機械翻訳のような単語の置き換えだけが外国語教育ではないはず。
- ・剽窃などの倫理的問題が気になる。

以上が山田ら(2021a)で報告されたアンケート調査の概要と主な結果である。予備的な調査であるとはいえ、一定数の英語教員の機械翻訳に対する姿勢が初めて明らかになったことに意義がある。

3. 教養英語教育の今後

アンケート結果からわかるように、教養英語担当者の中には英語教育において機械翻訳を活用してみたいと考える教員が一定数存在する。そして、これらの教員の目下の課題は機械翻訳の効果的な使用方法の習得と倫理面の問題の解消である。

幸いなことに、2021 年度の後半に機械翻訳と外国語教育に関連した発表が増加し^{[5][6][7][8][9]}、シンポジウムなども数多く開かれるようになった。そのような場を通して授業での具体的な機械翻訳使用や、機械翻訳使用のルール作りのアイデアなどについての情報交換が進みつつある。京都大学主催のシンポジウム「転換期の大学言語教育—AI 翻訳とポスト・コロナへの対応」^[10]

や青山学院附置外国語ラボラトリー主催の公開セミナー「機械翻訳と外国語教育—現状・課題・展望」^[11]はその代表例である。そしてさらに、機械翻訳の使用を盛り込んだ教養英語テキスト^[12]が発刊されるに至っている。

これは良い兆候である。外国語教育における機械翻訳の使用に賛成であれ反対であれ、オープンに話し合える雰囲気が形成され、検討材料となる事例が提供されつつあるからである。筆者を含むアンケート実施メンバーの間では、2021 年度の間に環境が激変したという印象が共有されている。少なくとも機械翻訳について話題にしにくいという空気はもはや感じられない。裏を返して言うと、以前は機械翻訳について外国語教員が話題にしにくい雰囲気があった。筆者も「英語教員なのに機械翻訳を学生に使わせるなんて」と責められるのを恐れて、同僚に対して機械翻訳について長らく口を閉ざしてきた。

オープンに議論が行われるようになれば、今後は機械翻訳の利用を認める活動とそうでない活動の住み分けが議論されるようになるだろう。この際、「機械翻訳は report talk には向くが rapport talk には不向きである」^[13]といった外国語教員による報告が参考になるだろう。そして機械翻訳の利用を認める活動の内容は学生の専攻によっても変わってくるかもしれない。例えば、理系の学生と文系の学生ではニーズが異なるように思う。また、学生の英語習熟度によっても異なるかもしれない。これまでに発表された実践報告では「機械翻訳を有効利用できるのは中級以上の学生ではないか」とよく指摘されている。

英語を含む外国語教育全体として考えた場合、今後の喫緊の課題は現在のような「野放し」の状態から脱却することである。そして冒頭で紹介したガリー氏がたびたび発言しているように、教育機関単位で機械翻訳利用の方針を定めることである^[14]。教員によって機械翻訳使用のルールがばらばらでは学生が混乱しかねないし、外国語教員は語学の専門家として確固たる教育理念と方針を示すべきだからである。

その際、授業で機械翻訳の使用を認める・認めないにかかわらず、機械翻訳が世の中で使用されていても英語を学習するのはなぜなのか、教える側はもちろのこと、学習者がはっきりと自覚できるようにすることが重要であろう。そういった自覚を学習者自身を持たない限り、教養英語教育は今まで以上に授業をパスすることだけが目的の空虚なものになりかねない。

もう一步踏み込んでいうと、機械翻訳を使って広がる世界を英語を含む外国語教育にとっての大きなチャンスととらえるべきではないだろうか。事実、これまでの実践報告の中には、機械翻訳を使用して多言語の話者と交流を行ったところ、学習者の外国語学習に対する動機がかえって上がったという指摘もある^[15]。筆者自身も国際交流活動のイベント等で日本人学生が機械翻訳を使って交流に参加する光景を日常的に見ている。ではこういった機械翻訳を使う学生が英語を学ばなくなるかというとなんかそうじゃない。むしろその反対で、話が通じて楽しかったのもっと英語を学びたいと思う学生のほうが多いと感じる。このように、今までは英語その他の外国語が不自由なために国際的な活動から取り残されてきた多くの日本語話者にとって、機械翻訳の発達は朗報のはずである。機械翻訳を片手にだれもが外国語を使用する世界に飛び込むチャンスが訪れたのなら、英語やその他外国語の教員がそのガイド役を務めてもよいのではないだろうか。

4. おわりに

本稿では山田ら（2021a）によるアンケート調査の報告「日本の大学における教養英語教育と機械翻訳に関する予備的調査」の解説と補足を行った。このアンケート結果から、教養英語を担当する教員の機械翻訳に対する無防備ともいえる様子が明るみになると同時に、教員の機械翻訳に対する強い関心が浮き彫りになった。それを反映するかのように、2021年度の後半には機械翻訳と外国語教育に関する発表が相次いだ。今後は、機械翻訳の存在を前提として英語教育を初め

とする外国語教育がどのようにあるべきなのか、外国語教育界全体で模索する必要がある。そのためには本誌の読者である機械翻訳・翻訳の専門家からのフィードバックが重要な役割を果たすことになるだろう^{[16][17]}。

我々は今、教育現場における機械翻訳を「困ったもの」から「チャンス」に変える分岐点に立っているのではないか。

注・参考文献

- [1] トム・ガリー（2020）「ニューラル MT の問題」『機械翻訳』No.72 1-2 頁 アジア太平洋機械翻訳協会
- [2] 石臥薫子・高橋有紀（2018）「英語呪縛からの脱却」『AERA』12-21 頁 朝日新聞出版（2018 年 3 月 5 日）
- [3] 山田優・ラングリッツ久佳・小田登志子・守田智裕・田村颯登・平岡裕資・入江敏子（2021a）「日本の大学における教養教育英語と機械翻訳に関する予備的調査」『通訳翻訳研究への招待』No.23 139-156 頁 日本通訳翻訳学会
- [4] 山田優・ラングリッツ久佳・小田登志子・守田智裕・入江敏子（2021b）「英語教育現場における機械翻訳の活用に関する実態調査と分析結果」『The JACET International Convention Proceedings: The JACET 60th Commemorative International Convention』243-244 頁 大学英語教育学会
- [5] Kennedy, Olivia. (2021) Unexpected student writing strategies during the Covid-19 pandemic. *Paper presented at JALTCALL2021*. (2021 年 6 月 6 日)
- [6] 幸重美津子・薦田和美（2021）大学英語授業における機械翻訳 (MT) の活用——実践的英語使用者としての自立学習者育成に向けて The JACET 60th Commemorative International Convention (2021 年 8 月 28 日)
- [7] 高橋秀彰（2021）機械翻訳と中上級レベルの英語学習者のパフォーマンス比較から考える外国語教育政策の可能性 日本通訳翻訳学会関西支部第 57 回例会 (2021 年 12 月 11 日)

[8] 坂本輝世 (2021) 英語コミュニケーションへの意欲と機械翻訳 日本英語表現学会第 50 回全国大会 (2021 年 12 月 18 日)

[9] 佐竹幸信 (2021) 機械翻訳が英語ライティング学習に与える効用について 日本英語表現学会第 50 回全国大会 (2021 年 12 月 19 日)

[10] 京都大学 (2021) 京都大学創立 125 周年記念シンポジウム 転換期の大学言語教育——AI 翻訳とポスト・コロナへの対応 (2021 年 12 月 4 日)

[11] 青山学院大学附置外国語ラボラトリー (2022) 公開セミナー 機械翻訳と外国語教育——現状・課題・展望 (2022 年 3 月 4 日)

[12] 幸重美津子・葛田和美・西山幹枝・Tom Gally (2022) 『AI 翻訳で英語コミュニケーション』三修社

[13] 2021 年 12 月 11 日に開催された日本通訳翻訳学会関西支部第 57 回例会における高橋秀彰による「機械翻訳と中上級レベルの英語学習者のパフォーマンス比較から考える外国語教育政策の可能性」で述べられた内容の一部である。この発表では日本人大学生が翻訳機を用いて英語母語話者と会話を行った様子の談話分析の結果が報告されたが、機械翻訳を用いた会話では笑いが全くなかったなど、話し相手との人間関係を築きにくい状況が報告された。

[14] トム・ガリー (2021) 語学教育と MT・機械翻訳の問題——第二言語教育の立場から アジア太平洋機械翻訳協会第 2 回定時社員総会招待講演 (2021 年 6 月 23 日)

[15] 成田潤也 (2019) 「機械翻訳を介しての外国語と国語の横断的学習に関する研究」『2019 年度第 14 回児童教育実践についての研究助成研究成果報告書』博報財団

[16] 機械翻訳の専門家が英語教員向けに行った講演の例として井佐原均 (2019) ディープラーニングで進化する機械翻訳 何ができて、何ができないか JACET 関西支部 2019 年度第 1 回支部講演会 (2019 年 6 月 22 日) がある。

[17] 翻訳の専門家が英語教員向けに行った講演の例

として山田優 (2021) 機械翻訳の進歩に接して外国語の教師が心得ておくべきこと 公開セミナー 機械翻訳と外国語教育——現状・課題・展望 青山学院大学附置外国語ラボラトリー (2022 年 3 月 4 日) がある。

謝辞

この報告は JAITS (日本通訳翻訳学会) プロジェクト「機械翻訳と外国語教育について考える」(2019 年度～2021 年度) の助成を受けて行われたものである。

MT に関する法的問題について

AAMT

MT を利用することによって起こりうる法的な問題と、その場合の解釈をまとめました。MT 利用時の参考にしてください。

■ MT を利用することによって発生する法律問題

1 みずからの情報発信のために MT を利用する場合

みずからの情報発信のために MT を利用する場合に生じる法律問題は、「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、当該翻訳結果に誤りがあり、当該誤りに起因して第三者が損害を被った場合に、MT 利用者はどのような責任を負担するか」というものである。

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る。

この場合においては、単純に当該情報発信者（MT 利用者）が誤った情報を発信したというだけであって、それにより第三者が損害を被ったのであれば、当該情報発信者（MT 利用者）は法的責任を負う。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さない。

2 他者から翻訳の依頼を受けて MT を利用する場合

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、（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点である。

（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る。

秘密保持条項は、端的に言うところ「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、という条項であるが、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」に該当する。

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となる。

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、有償無償は無関係と思われ、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るのではないと思われる。

具体的には、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」したことにはならないと思われる。

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等に限定されている場合も、「第三者に開示」していないと解釈できるとと思われる。

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 利用者の間の翻訳契約上許容されていなければ、秘密保持義務違反に該当する可能性が高い。

（２）MT を利用した翻訳結果に誤りがあった場合に MT 利用者はクライアントに対してどのような責任を負うか。

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さない。したがって、本ガイドラインにおいては最低限の言及だけ行う。

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 条等）を負う。

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になる。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る。

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る。

■ 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」という問題である。



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る。

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」であるが、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ない。

ただし、これは「MT 提供者・MT 利用者間の契約」の問題であり、「MT 利用者・クライアント間の翻訳契約」の問題ではない。「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る。

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）」の問題であるが、この点については結論から言うと著作権侵害にはならない。

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ある。

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る。

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 号である。

（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あるが、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ある。

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」ということになる。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・改変等する行為」「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」などが可能である。

機械翻訳と人間翻訳の両方の市場を最大にする方法

ローズ 三浦

エッジ・トランスレーション

1. 人間翻訳はこのまま衰退していくのだろうか？

近年、機械翻訳の発展は目覚ましい。著者はプロの翻訳者として日常的に人間翻訳業務に携わっているが、同業者の中には機械翻訳の台頭によって失業するのではないかと危機感を覚える者が多数いる。著者が翻訳者を対象として 2020 年 5 月に Twitter 上で実施したアンケート調査によると、「自分の訳は DeepL よりレベルが高いか」という質問に対して、なんと 4 割弱が「時と場合による」と回答した。

しかし、翻訳というものはそう単純ではない。本稿では、「翻訳市場は非ゼロ和ゲーム」であって、「機械翻訳と人間翻訳のマスの取り合いではなく、マス数そのものを増やせる可能性がある」という仮説を検証し、機械翻訳と人間翻訳の両方の市場を最大化する方法について、考察する。

2. 機械も「間違える」

まず、前提として、人間と同様に機械翻訳も「間違える」という認識が必要である。機械翻訳が「間違える」例として、ある医療分野の案件の訳例では、重複や、原文にない単語が出現している。2 行目の「grade 3 or 4 hypertension, myocardial infarction, acute coronary artery syndrome or ventricular arrhythmia, and NYHA grade II or higher」という文と、3 行目の「Grade 3 or 4 hypertension, myocardial infarction, acute coronary artery syndrome, or ventricular arrhythmia, and NYHA classification Grade II or higher）」という文が重複し、Clozapine という原文に

ない単語が湧き出している。全文は、AAMT 2021, Online における著者の発表資料を参照されたい (<https://www.slideshare.net/ssuser8d5970/ss-250808319>)。YouTube 動画で詳しく解説している [1](#)。

3. 機械翻訳に誤訳させないコツ

機械翻訳に誤訳させないコツとして、次のものを提案する。

- 文章全体の一貫性を保つ
- 曖昧さを避け、明確に書く
- ユーモアや文化的背景に基づく表現を避ける
- 短くて簡潔な文を書く
- 文法構造を単純化する
- なるべく能動態を使う
- 繰り返し使えるフレーズを書く
- 不要な文言を削除する
- 決まり文句や口語的な言い回しを避ける
- 必要に応じて定冠詞 (the) を使う
- 代名詞を多用せず、なるべく同じ名詞を繰り返す
- 常に情報の正確さを保つ

4. 「読解」と「発信」でチェック段階を分ける

まず、「読解」（読者が原文情報の大意を把握することを目的とする場合）については、機械翻訳をチェックなしで使用可とする。

「発信」を目的とする場合は、機械翻訳を人間がチェックする必要があり、誤訳が生じた場合に影響が大きいものほどチェックを厳しくする（影響度に応じて差をつける）。

最も影響が大きいものは、人命にかかわる内容（医療や労働安全衛生など）であり、これらについては、当該分野に精通した者がチェックする必要がある。逆に影響が最も小さいものは、一般的な内容で、英語をできる者がチェックすれば可とする。

このように、機械翻訳のリスクを抑えつつ、スピード・効率を最大化できるようにチェック体制を整えることを提案する。

5. 情報伝達または訴求

「情報伝達」については、基本的には前述の「3. 機械翻訳に誤訳させないコツ」を基に、機械翻訳 + PE を行う。ただし、難易度が高いといった特別な場合には人間翻訳を行う。

「訴求」については、基本的には人間翻訳を行う。例として、「Easy to pick up, addictive side-scroller action」という英文を DeepL に掛けると、「簡単な操作で、やみつきになる横スクロールアクション!」という訳文が生成されるが（※発表時点の翻訳結果）、人間翻訳では「簡単につかめて楽しい操作、超ハマる横スクロール型のアクション!」と訳せる。また、「SupercolorpixelisticXDRidocious」（iPhone 13 のキャッチコピー）という英文を和訳させると、DeepL は「スーパークラリックピクシスティックXDRidocious」という訳文を生成し（※発表時点の翻訳結果）、人間は「魔法みたいなディスプレイ体験」と訳す（公式訳は「スーパーキラキラカラフルクッキリディスプレイ」）。その他の訳例は、AAMT 2021, Online における著者の発表資料を参照されたい²。

6. MT 利用ガイドライン委員会

AAMT は、専門の委員会を設けて、MT 利用ガイドラインを作成し公表するための作業を行っている。著者は、その委員の一人である。機械翻訳を人間の『補佐』として位置づけた上で MT 利用ガイドラインを規

定すれば、両者の共存が可能であろう。詳しくは、著者の AAMT 2021, Online での発表資料を参照されたい²。

参考文献：

1. ローズ三浦. ローズ三浦のメディカル翻訳講座③.
https://www.youtube.com/watch?v=KG9CpOaVbBM&ab_channel=%E3%83%AD%E3%83%BCE%E3%82%BA%E4%B8%89%E6%B5%A6%E3%81%A1%E3%82%83%E3%82%93%E3%81%AD%E3%82%8B
2. ローズ三浦. 機械翻訳と人間翻訳の両方の市場を最大にする方法.
<https://www.slideshare.net/ssuser8d5970/ss-250808319>

Memsource 機械翻訳調査レポート 2022 年第 2 四半期

Memsource 日本窓口

1. はじめに

Memsource（メムソース）はクラウド型の翻訳管理システムおよび翻訳支援ツールです。2012 年にリリースされ、世界で 35 万以上のユーザーに利用されています。ツール内の機械翻訳ソリューションである Memsource Translate では、30 種類以上の機械翻訳（MT）エンジンとの連携が可能のため、ユーザーは機械翻訳を活用して翻訳作業を行えます。

Memsource では四半期ごとに、ユーザーが実際の業務等で Memsource を通じて行った翻訳のデータから、機械翻訳の利用動向を分析する調査を行なっています。今回はその最新の結果を報告します。

2. 調査方法

ポストエディットスコアによる MT 品質の測定

ポストエディットとは、MT エンジンから出力された文章を翻訳者が修正・改善するプロセスです。つまり、ポストエディットが多く行われた場合は元の MT の質が不十分であり、あまり行われなかった場合は MT の質が高かったと推測されます。

この考えに基づき、翻訳者が加えた変更の割合に応じて、セグメント（文章の区切り）ごとに 0～100 のスコアづけを行います。MT の出力がそのまま受け入れられた場合はスコア＝100、翻訳者によって完全に書き換えられた場合はスコア＝0 となり、スコアが高いほど、MT の質が高いことが示唆されます。これを「ポストエディットスコア」と呼びます。

対象となるデータ

本調査では、Memsource のシステム内で収集した、匿名化されたポストエディットデータを使用しています。

対象となるのは、MT の出力を翻訳者がポストエディットしたセグメント、および機械翻訳を利用できる状況で翻訳者が一から翻訳を行ったセグメントです。翻訳メモリのマッチングによって翻訳されたセグメントや、翻訳不可能なセグメントは除外します。

ドメインの自動分類

MT の出力品質は、医薬、法律、ビジネス、エンターテインメント、IT など、翻訳するコンテンツの分野によって変わります。このような分野を「ドメイン」と呼びます。Memsource では、文中のキーワードを AI アルゴリズムが分析し、自動的にドメイン进行分类しています。例えば、Study、patients、treatment、dose、mg、clinical といったキーワードから、医薬ドメインのコンテンツであると判断します。

3. 調査結果

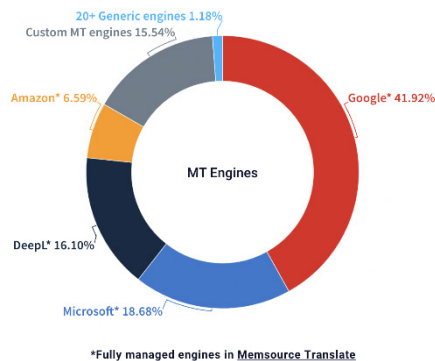
対象セグメント数

今回対象となった総セグメント数は 5400 万以上、言語ペア数は 60 以上でした。対象セグメント数は調査ごとに増加を続け、今回も過去最高を記録しました。

使用された MT エンジンの種類

Memsource は 30 種類以上の MT エンジンをサポートしています。図からわかるように、Google、Microsoft、DeepL、Amazon といった汎用エンジンが使用の大半を占める一方、独自のカスタム機械翻訳エンジンの使

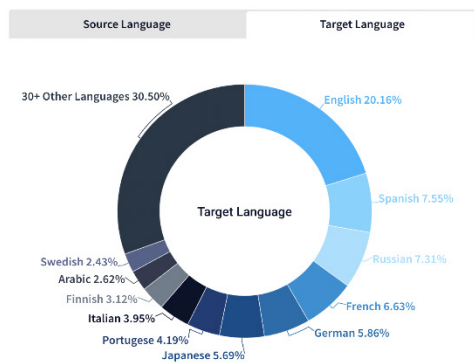
用も約 15% ありました。なお、使用されたエンジンの 80% 以上が、複雑な設定の必要がない Memsources の「フル接続対応」エンジンでした。



使用されている MT エンジンの種類

MT が良く使われる言語ペア

トップ 3 の言語ペアは、英語・スペイン語、英語・ロシア語、英語・仏語でした。ソース言語、ターゲット言語とも英語が第 1 位で、とくにソース言語では全体の 64.83% を英語が占めています。日本語はソース言語としては 4.99% (2 位)、ターゲット言語としては 5.69% (6 位) でした。



MT が使われた各言語の割合 (ターゲット言語)

言語ペア別の MT 品質

75% の言語ペアで平均スコアが 70 以上でした。2021 年第 4 四半期と比較して、全体平均では約 1 ポイント上昇しています。従来からの傾向として、スペイン語、イタリア語、ポルトガル語といった関係が密

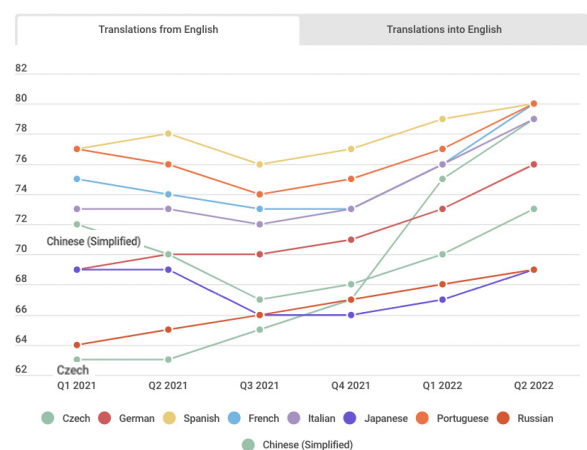
接な言語ペアで高いスコアが見られます。全般的に欧米系の言語でスコアが高い一方、アジア系 (日本語、中国語) では比較的 low、ロシア語もやや低い傾向が見られます。スコアが高い言語では MT 出力の 7~9 割が活用されており、英語・日本語のペアでも、平均で 7 割近くの MT 出力が活用されていることがわかります。

	Q2 2022					Q1 2022				
Source below	Czech	German	English	Spanish	French	Italian	Japanese	Portuguese	Russian	Chinese (Simplified)
Czech		91	76			73				
German	75		79	75	79	75			64	
English	79	76		80	80	79	69	80	69	73
Spanish		71	81		78	90		87		
French		68	77	76		73				
Italian		78	77	83	78				66	
Japanese			68							53
Portuguese			79	79						
Russian			77		61					
Chinese (Simplified)			54			48				

言語ペア別の平均スコア

MT 品質の変化

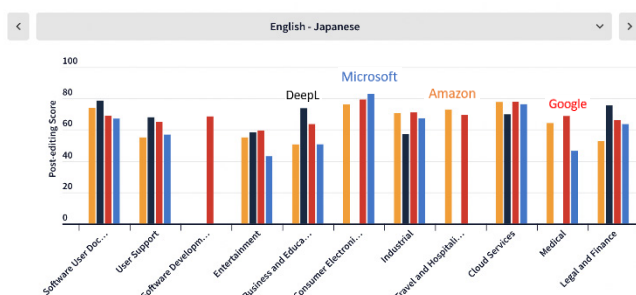
2021 年第 1 四半期から現在までの期間で、英語をソース言語とする翻訳において、全体の平均スコアが 5.1 ポイント上昇。英語をターゲット言語とする翻訳では 4.3 ポイント上昇しています。日本語に限ると、どちらも横ばいか微増か程度の変化です。



2021 年第 1 四半期~2022 年第 2 四半期のスコアの変化 (ソース言語: 英語)

高品質を得るために必要なエンジン数

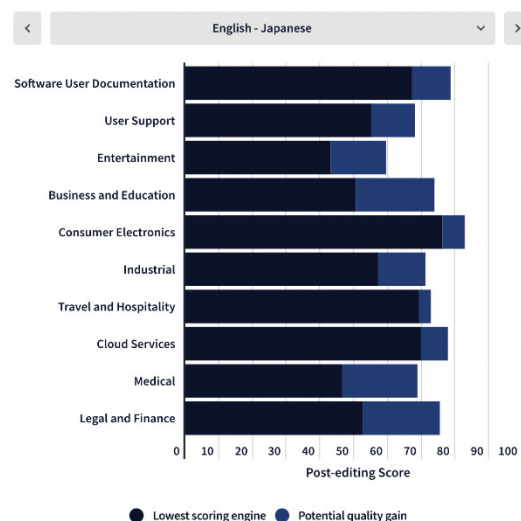
グラフから、平均スコアが最高であるエンジンの種類は、ドメインごとに異なることがわかります。全ドメインで最高の結果を得るには、平均で3種類のエンジンを使い分ける必要があります。英語・日本語ペアでは、例えばビジネス・教育では DeepL、コンシューマーエレクトロニクスでは Microsoft、旅行・ホスピタリティでは Amazon、医薬では Google がそれぞれ最高の平均スコアを記録しています。



英語・日本語ペアにおける、ドメイン別の各 MT エンジンのスコア

エンジン選びが品質に与える影響

選択するエンジンの違いによって生じる品質の変化を知るために、各ドメインで平均スコアが最も低いエンジンと最も高いエンジンのスコア差を求めると、例えば、英語・スペイン語では 10 ポイント、英語・仏語では 11.5 ポイント、英語・日本語では 14 ポイント、英語・中国語簡体字では 21.5 ポイントの差が見られました。日本語ではビジネス・教育、医薬、法律・金融の分野で差が大きくなっています。



各ドメインにおける、平均スコアが最も低いエンジンと最も高いエンジンのスコア差

4. まとめ

欧米系の言語ペアでは MT 出力の 7~9 割を、英語と日本語のペアでも MT 出力の 7 割近くを活用できていることがわかりました。また、MT エンジンの品質は、言語ペアやドメインによって変化する傾向も確認されました。全ドメインで最高の結果を得るには、平均で3種類のエンジンを使い分ける必要があることがわかりました。

産業翻訳における機械翻訳の継続的改善手法

中山 雄貴

株式会社ヒューマンサイエンス

1. はじめに

本稿では、企業の翻訳部門や翻訳会社を対象に、機械翻訳を導入したあとで翻訳業務を継続的に改善する手法について説明する。

機械翻訳を導入してもその作業者は翻訳者であり、翻訳者が翻訳とポストエディットの両方を担当する。翻訳者の作業は、翻訳メモリからの訳文の流用とポストエディットに分けられる。それぞれの内容と改善方法について説明する。

2. 翻訳メモリからの訳文の流用

まず、翻訳メモリの使用状況について説明する。図1に、ある IT 文書を新規に翻訳する際の翻訳メモリからの流用率を示す。新規の場合でも定型表現が多いと半分近くの訳文が翻訳メモリからの流用だった。定型表現がない文書でも、いくらかの訳文は過去の訳文から流用可能であり、また、現在の文書の中の翻訳済みの訳文から流用することも可能だった（レペティションと呼ばれる）。

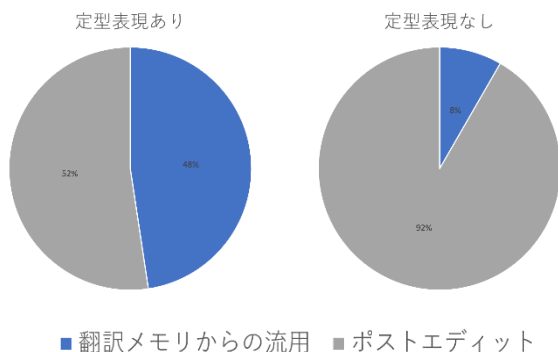


図1：流用文の割合（新規文書の場合）

図2に、改訂文書の場合の流用率を示す。あるソフトウェアのアップデートに伴い、ヘルプ文書を改訂したときに、マイナーアップデートでは99%が翻訳メモリからの流用だった。メジャーアップデートでも89%が流用だった。

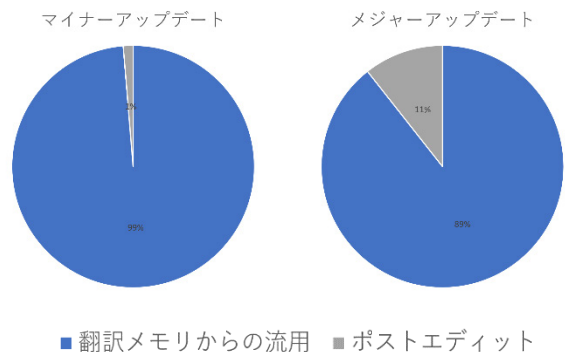


図2：流用文の割合（改訂文書の場合）

翻訳業務に機械翻訳を導入しても、翻訳メモリの重要性が変わらないことが分かる。翻訳メモリの訳文の品質が低いと流用時に修正が必要になり、翻訳効率には上がらないため、翻訳メモリの訳文の品質を維持することが重要である。では、どうやって翻訳メモリを維持管理したら良いのだろうか。

3. 翻訳メモリの種類とライフサイクル

翻訳メモリには翻訳用、レビュー用、マスターの3種類を用意する。

翻訳用は、一時的なもので、翻訳プロジェクトごとに作成し、プロジェクトが終わると破棄する。翻訳者が読み書きする。レビュー用も、一時的なもので、プロジェクトごとに作成、破棄する。レビューアが読み書きする。マスターは、永続的なもので、プロジェク

トを超えて使用する。翻訳者とレビューアは読み込み専用（参照専用）で使用する。プロジェクトマネージャがプロジェクト完了時に書き込む。後述するが、このマスター翻訳メモリは機械翻訳のモデルを学習するためにも使用する。

図3に翻訳メモリのライフサイクルを示す。

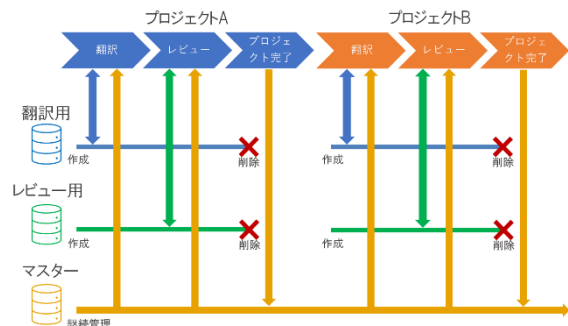


図3：翻訳メモリのライフサイクルと読み書き

翻訳プロジェクトAとBがあるとする。プロジェクトAの最初に翻訳用とレビュー用の翻訳メモリを新規に作成する。翻訳のステップでは、翻訳者が、翻訳用の翻訳メモリを読み書きし、マスターの翻訳メモリを参照専用で使用する。レビューのステップでは、レビューアが、レビュー用の翻訳メモリを読み書きし、マスターの翻訳メモリを参照専用で使用する。プロジェクトAの完了時に、翻訳用とレビュー用の翻訳メモリは破棄する。プロジェクトマネージャが、最終的に確定した訳文のみをマスターの翻訳メモリに書き込む。

その後、プロジェクトBが始まる時には、改めて翻訳用とレビュー用の翻訳メモリを作成する。マスターの翻訳メモリは前のプロジェクトAから引き継ぐ。翻訳とレビューでは、それぞれ新しく作成した翻訳メモリを読み書きし、マスターの翻訳メモリを参照専用で使用する。プロジェクトBの完了時に、翻訳用とレビュー用の翻訳メモリは破棄し、最終的に確定した訳文のみをマスターに書き込む。

このようにすることで、マスターの翻訳メモリには確定済みで高品質の訳文が継続的に蓄積されることになる。

製品ラインが複数ある場合や、求められている品質が異なる場合は、それぞれ用にマスターの翻訳メモリを用意することが推奨される。

4. ポストエディット

ポストエディットでは、原文の1文ごとに以下の作業を行う。

- ・原文を読み、理解する
- ・機械翻訳の訳文を読む
- ・正確さを確認して、修正する
- ・流暢さを確認して、修正する
- ・用語を確認して、修正する
- ・スタイルを確認して、修正する

ここで、スタイルとは、翻訳用スタイルガイドに記載されたルールを指す。全角文字と半角文字の間にスペースを入れる、記号には半角を使う、敬体（ですます調）を使う、カタカナの長音を省略する、といったものがある。

あるIT文書の翻訳プロジェクトでそれぞれの作業にかかった時間を計測した。ポストエディットの作業時間の中に占める割合を図4に示す。

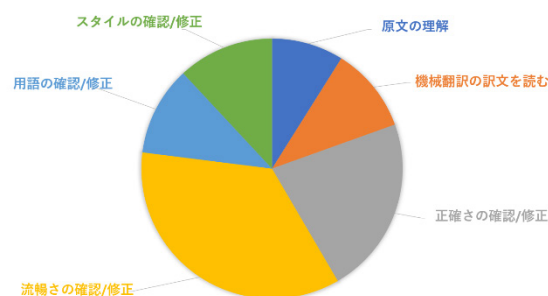


図4：ポストエディット作業の詳細とその割合

これらの作業について改善する方法を個別に検討する。まず、用語とスタイルについて説明する。

5. 用語とスタイル

機械翻訳の訳文には用語やスタイルの誤りが多数含まれている。図5の上部にGoogle翻訳が出力した訳

文、下部にその修正例を示す。ファイルパスの中に全角コロンや不要なスペースが含まれている。また、会社名、部署名、製品名が間違っている。

The file is saved in C:\Users\username\Downloads.	AI	ファイルはC:¥Users¥username¥Downloadsに保存されます。
HUMAN SCIENCE provides machine translation solutions.	AI	ヒューマンサイエンスは、機械翻訳ソリューションを提供しています。
Please contact our Business Development department for details.	AI	詳細については、事業開発部門にお問い合わせください。
PE booster is a software to support post editors.	AI	PEブースターは、投稿編集者をサポートするソフトウェアです。



The file is saved in C:\Users\username\Downloads.	AI	ファイルはC:¥Users¥username¥Downloadsに保存されます。
HUMAN SCIENCE provides machine translation solutions.	AI	HUMAN SCIENCEは、機械翻訳ソリューションを提供しています。
Please contact our Business Development department for details.	AI	詳細については、事業推進部にお問い合わせください。
PE booster is a software to support post editors.	AI	PE Boosterは、ポストエディターをサポートするソフトウェアです。

図5：用語とスタイルの修正例

このような用語とスタイルの修正は単純作業であり、自動化することが可能だ。機械翻訳サービスの中には用語集機能やスタイル置換機能を備えているものもある。用語集機能を使うと、原文の特定の用語を対象に訳語を指定して訳文に出力できる。スタイル置換機能を使うと、訳文側で文字列置換や正規表現置換ができる。弊社では、DeepLやGoogleなどの機械翻訳エンジンに追加できる汎用的な用語集機能とスタイル置換機能を開発した。これらの機能はMemsourceとTradosで利用でき、社内の翻訳業務で活用している。また、企業の翻訳部門や翻訳会社向けにそれぞれMTrans（エムトランス）for MemsourceとMTrans for Tradosというサービスとして提供しており、REST APIとして利用することも可能だ。

用語集機能とスタイル置換機能の効果を図6に示す。109文の英日翻訳の中で、用語とスタイルの修正だけで25分もの時間が取られていたが、用語集機能とスタイル置換機能を利用することでこの時間を削減できた。副次的な効果として、翻訳者が、単純な修正作業から解放され、正確さと流暢さに集中できるようになったため、訳文の品質が向上すると同時に、翻訳者が抱えていたポストエディット作業に対する嫌悪感を抑えることもできた。

Google	Google + MTrans	
109文中95文 87.2%	誤りが存在した文の数	0文 0%
661	誤りの数	0
25分5秒	修正にかかった時間	0秒
単純作業が多く手間に感じるⓈ ポストエディターの感想 翻訳に集中できるⓈ		

図6：用語集機能とスタイル置換機能の効果

図6のような効果を得るには、用語集機能やスタイル置換機能を適切に設定することが重要だが、その設定を一回で完成させることは難しく、設定内容を継続的に改善することが必要になる。そこで役立つのが、ポストエディット前後の訳文の差分データだ。翻訳プロジェクトの終了時に、ポストエディット前後の差分を作成し、この差分を見て、用語集やスタイル置換条件を更新する。

Memsourceでは差分の作成機能が提供されている。具体的な方法についてはMemsourceヘルプの「Evaluating MTQE Results」に記載されている(<https://help.memsource.com/hc/en-us/articles/360012527380-MT-Quality-Estimation> 参照)。TradosではプラグインのPost-Edit Compareが無料配布されている。RWS AppStoreからダウンロードできる(<https://appstore.rws.com/jp/language/app/post-edit-compare/610/> 参照)。

6. 正確さと流暢さ

正確さと流暢さの作業を改善するには、機械翻訳エンジンの選択とモデル学習の2つの方法がある。

前者については、各社の機械翻訳エンジンを比較し、自社の翻訳分野に強いエンジンを選択する。一般にDeepLの品質が高いと言われているが、これは複数の文をまとめて段落単位で翻訳した場合である。翻訳支援ツールとDeepLを組み合わせると文単位の翻訳となるため、DeepLの品質を最大限に活用できない場合がある。ただし、文単位でも長文の場合は、DeepLの訳文は流暢なことが多い。実際の業務で使用する翻訳

支援ツールと文書を使ってエンジンを比較検証することが推奨される。

「継続的改善」という視点で見ると、モデル学習を備えていない機械翻訳エンジンは、改善を自社でコントロールできない点が問題になる。

後者のモデル学習とは、汎用の機械翻訳モデルに対して独自の原文・訳文のペアを学習させ、独自の訳文を生成できるようにすること。代表的なサービスとして Google AutoML Translation がある。執筆時点で DeepL ではモデル学習することはできない。

モデル学習のメリットとしては、学習した分野について、訳文の精度が向上する。Google AutoML Translation で検証したところ、わずか 800 文の学習データで、BLEU 値が 3.1 ポイント向上した（図 7 参照）。

モデル学習前	39.7
モデル学習後	42.8 ↑3.1

分野: IT
言語: 英日
学習データ分量: 800 文

図 7 : モデル学習前後の BLEU 値

モデル学習には、前述のマスター翻訳メモリに保存されている原文と訳文のデータを使用する。質の良い原文と訳文が高品質なモデルに繋がるため、マスター翻訳メモリに質の良い原文と訳文のペアを蓄積することが重要になる。

図 8 に示すように、モデル学習は継続的な作業である。日々の翻訳案件の中で、ポストエディットした訳文をマスターの翻訳メモリに蓄積する。これを使ってモデル学習を行う。このループを定期的に実行する。

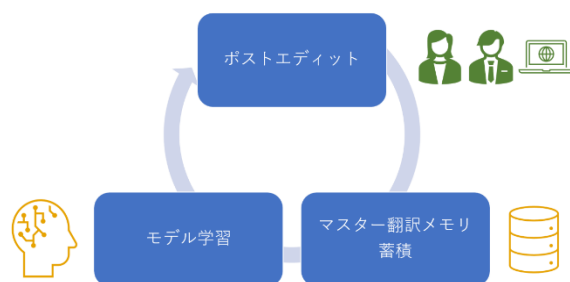


図 8 : モデル学習のサイクル

注意点としては、学習済みのモデルを使っても、学習データの訳文、用語、スタイルがそのまま出力されるわけではないため、翻訳メモリ、用語集機能、スタイル置換機能を引き続き利用する必要がある。

例文を図 9 に示す。「学習データの訳文」と「モデル学習後の訳文」を比較すると、学習データの訳文とは異なる訳文が出力されている。用語では、「Layer 2」の訳が学習データの「第 2 層」ではなく「レイヤー2」になっている。スタイルでは、学習データに含まれていた全角文字と半角文字の間のスペース（「▲」で示す）や、括弧が半角であることも訳文に反映されていない。

原文	The NIC exists on the 'Data Link Layer' (Layer 2).
学習データの訳文	NIC▲は「データリンク層」▲(第▲2▲層)▲に位置します。 (「▲」は半角スペースを示しています)
モデル学習前の訳文	NICは、「データリンク層」(層2)に存在します。
モデル学習後の訳文	NICは、「データリンク層」(レイヤー2)に存在します。

図 9 : モデル学習の注意点

7. まとめ

機械翻訳導入後に翻訳業務を継続的に改善するには以下の3点を行う。

- ・翻訳メモリからの流用作業を改善するために、翻訳用、レビュー用、マスターの翻訳メモリを用意し、マスター翻訳メモリを永続的に維持管理する。
- ・ポストエディットの用語とスタイルの修正作業を改善するために、機械翻訳サービスの用語集機能とスタイル置換機能を導入した上で、ポストエディット前後の訳文の差分を見て、用語集機能とスタイル置換機能の設定を更新する。
- ・ポストエディットの正確さと流暢さの修正作業を改善するために、マスター翻訳メモリを使ってモデル学習する。

最適化された機械翻訳導入モデルによる医薬品開発業務の迅速化への試み

早川威士

株式会社アスカコーポレーション

業務推進部

鞠子幸泰

中外製薬株式会社

信頼性保証企画部 翻訳マネジメントグループ

1. 医薬品開発業務への機械翻訳 (MT) 導入

人工知能 (AI) 技術はすでに様々な形態で実用化され、いまや産業の発展を強力に推進するツールとなっている。ただし、AI は出力に対して確実性や正確性が約束されるものではなく、導入した効果の測定が難しいという課題がある。AI の有効活用には、期待することや目的を明確にし、導入効果を可視化できるような枠組みが必要である。

医薬品業界はグローバル化が進展した産業であり、多言語コミュニケーション手段として翻訳が重要な役割を持つ。翻訳には高い知識水準が要求される一方で時間短縮も課題となっており、機械翻訳 (MT) 技術には、翻訳品質の安定化と翻訳作業の迅速化の双方が求められている。しかしながら、現在までの MT 技術は、医薬品関連文書に求められる高いレベルの翻訳品質に追いついておらず、直ちに人手の翻訳を MT に置換できるわけではない。そのため、人によるポストエディット (PE) を行うことを前提とした最適な MTPE 手法を設計・開発する必要がある、MT の有用性は PE の品質と作業スピードを加味して評価すべきと考えられる。

我々は本研究において、MT を用いた翻訳モデルを最適化することにより、MT 技術が翻訳品質の安定化と迅速化の双方にいかに関与するかを検討した。具体的には、人手翻訳による翻訳プロセスから、MTPE に移行することで、翻訳に要する期間がどれだけ短縮できるかを比較した。また、翻訳モデルの最適化として、(1) 翻訳品質を改善するためのドメイン (医薬品開発) に特化した MT の開発、(2) 過不足なく適切な PE を

実施するための品質水準の設定、の 2 つの手法を取り入れた。

2. ドメイン特化型 MT の開発および評価

本研究の対象となっている医薬品開発業務は極めて専門性が高い領域であることから、非特異的なコーパスで学習させた MT (汎用 MT) は本業務の用途に適さないと考えられる。そこで、転移学習の一手法である domain adaptation を用いて、ドメイン特化型の MT を構築し本研究の MTPE に使用するものとした。

学習データは著者の企業が保有する日本語と英語がペアになった業務用文書約 24 万文を使用し、基盤 MT として「みんなの自動翻訳@TexTra® (国立研究開発法人情報通信研究機構)」に対し domain adaptation を実施した。

上記で構築した MT (カスタマイズ MT) が実際に PE の品質および作業スピードの改善に寄与するかを評価するため、カスタマイズ MT と汎用 MT を用いた英語から日本語の翻訳出力に対してそれぞれ PE を実施し、編集時間と編集率を記録して比較した。編集時間は 1 文ごとに PE に要した時間 (秒) を計測し、平均値を算出した。編集率は PE 前訳文と PE 後訳文のレーベンシュタイン距離を正規化¹した値とした。また一般的な評価指標として、BLEU²および TER (translation error ratio)³を算出した。以上の指標の統計学的検定には paired Student's t-test を用いた。さらに、医薬品開発業務における専門用語や固有の表現がどの程度正確に再現されているかを評価するため、

¹ 2 文のレーベンシュタイン距離を長い方の文の文字列長で除すことにより求めた。

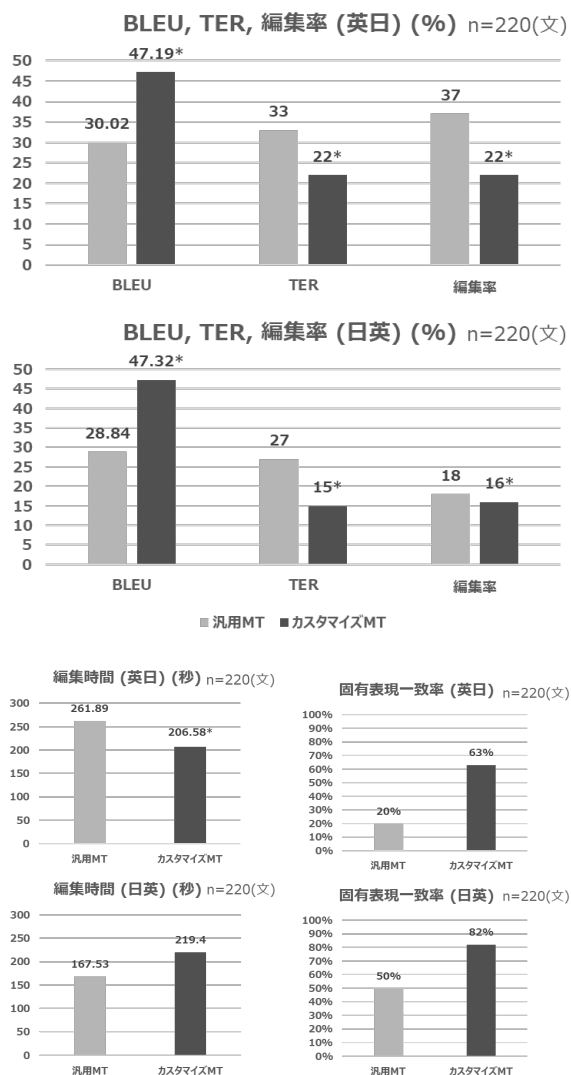
² Papineni, Kishore, et al., 2002

³ Snover et al., 2006

これらの表現の正答率（固有表現一致率）を比較した。いずれの評価も医薬品開発業務に関連するテキストより n=220 文を用いた。

PE の評価の詳細を図 1 に示す。まず編集率ではカスタマイズ MT が汎用 MT に対して優れており、英日翻訳では有意差を示した（22% vs 37%, $p<0.01$ ）。編集時間については、英日翻訳では有意にカスタマイズ MT が優れていた一方で、日英翻訳では有意差には至らないものの汎用 MT のほうが短かった（英日 206.58 秒 vs 261.89 秒, $p<0.01$; 日英 219.40 秒 vs 167.53 秒, $p=0.67$ ）。

図 1. カスタマイズ MT と汎用 MT の比較



* 統計学的有意差を示す (paired Student's t-test)。

BLEU スコアの平均値はカスタマイズ MT が英日、日英いずれも約 47%を超える高値となり、汎用 MT と比べ精度の改善が示された（英日 47.19% vs 30.02%; 日英 47.32% vs 28.84%、いずれも $p<0.01$ ）。TER も同様にカスタマイズ MT が英日、日英いずれも汎用 MT より優れていた（英日 22% vs 33%; 日英 15% vs 27%、いずれも $p<0.01$ ）。

固有表現の一致率については、英日では 43 ポイント、日英では 32 ポイント、カスタマイズ MT が汎用 MT を上回り、domain adaptation の効果を裏付ける結果となった（英日 63% vs 20%、日英 82% vs 50%）。

これらの結果から、概ねカスタマイズ MT の方が優れた翻訳結果を出力していると考えられ、実際の PE 作業においても編集率が減少しており、PE の労力を軽減していることが示された。一方で、日英翻訳の編集時間についてはカスタマイズ MT の方が時間を要し、編集率やその他指標の示唆する評価とは相反する結果となった。

3. PE 品質水準の設定

PE の作業時間を短縮し、MTPE による翻訳モデルを最適化するために、PE の品質水準を設定し事前にポストエディターと共有した。これにより、ユーザーが期待する翻訳品質を上回る過剰な編集／重要性の低い編集にかかる作業時間、および編集作業の判断にかかる時間を短縮することを狙っている。すなわち、ここで定義した品質水準は、品質に関する要件であるとともに PE の効率性を高める効果を期待したものであるといえる。

品質水準の内容は、(1) 要求事項と (2) 非要求事項をリストアップすることで構成され、PE でやるべきこととやらなくてよいことを明示化している。要求事項は、「正確性」、「ユーザビリティ」、「リーダビリティ」および「申し送り」の 4 項目から構成され、業務利用のための翻訳文書として比較的重要な項目として絞り

込んだ。一方で、非要求事項には表現の好みや一貫性など、一般に翻訳品質の構成要素ではあるものの使用を考慮すると比較的優先度の低い事項を挙げている。図2に品質水準の項目の一例を示す。

図2. 品質水準の項目の一例

・ 要求事項 ➢ 正確性：原文のプロトコルの情報が、過不足なく翻訳版に反映されている。 - 治験を実施する上で問題となる数値や単位の誤差がない。 - 否定／肯定の逆転がない。 - 原文と翻訳版の図表のデータの内容に齟齬がない。 - 原文中の治験実施に必要な情報について、不足や原文にない情報の追加がない。 ・ 非要求事項 ➢ リーダビリティ：逸脱を起こす読みにくさがない - 意味や可読性に影響しない表現の好みや適用しない。 - 意味や可読性に影響しない語順や「てにをは」、文末表現にこだわらない。 - 文章の内容理解に支障を来さない限り、逐語訳を許容する。

4. MTPE と人手翻訳の実証的研究

前章までで記述した翻訳モデルの妥当性を確認するため、実証的研究として人手翻訳によるプロセスとMTPEとで、翻訳に要する時間がどれだけ異なるかを比較した。

本研究では、構築した翻訳モデルを医薬品開発の実務に適用させるため、薬事規制に基づいて実施される臨床試験に実際に使用した「治験実施計画書」を対象文書として選択した。また、異なる2つの治験実施計画書の翻訳を2回に分けて行い、MTPEの時間短縮効果に再現性があるかを確認するものとした。時間短縮効果の目標は、人手翻訳に対し、MTPEで30%以上の時間短縮効果がみられることとした。同時に品質水準の達成度について、要求事項4項目に対するエラー数をカウントして評価した。1回目の結果で人手翻訳よりMTPEの方がエラー数が多い場合、翻訳およびPE担当者に対してフィードバックを実施し、2回目でのエラー抑制効果を確認した。

1回目の比較では、人手翻訳に対し、MTPEでは48%の時間短縮効果を確認することができた(図3、226時間 vs 106時間)。しかしながら、品質水準の達成度では、4項目いずれもエラーが発生しており、中でも「正確性」においてはMTPEの方がエラー数が多

い結果となった(図4、19件 vs 26件)。このため2回目の実験では、エラー数を削減してより品質水準を高度に達成することも考慮することとした。

図3. 人手翻訳とMTPEの治験実施計画書の翻訳に要した時間の比較

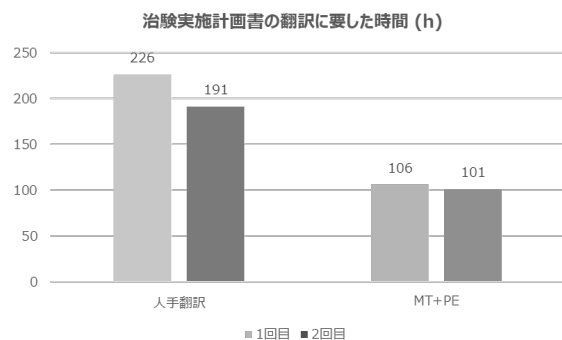
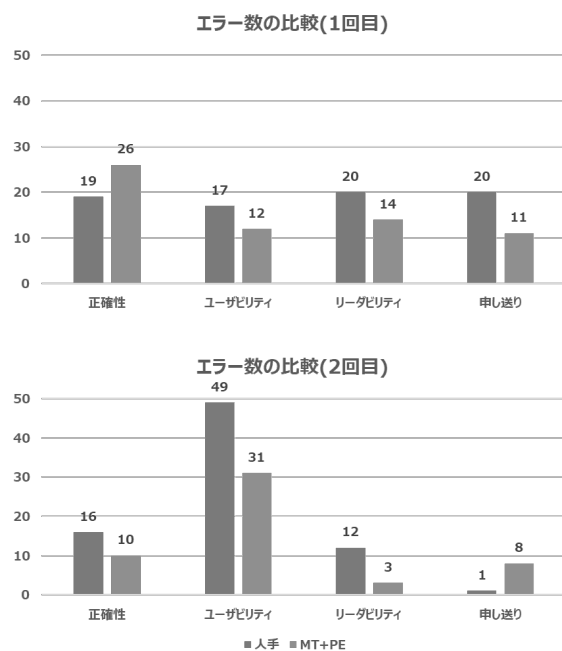


図4. 人手翻訳とMTPEのエラー数の比較結果(1回目及び2回目)



2回目の比較では、1回目引き続き、人手翻訳に対してMTPEで43%の時間短縮効果を確認した(図3、191時間 vs 101時間)。なお、翻訳作業の実時間についても、1回目から短縮された。品質水準の達成度については、「正確性」「リーダビリティ」では1回目に比べてエラー数が減少するとともに、人手翻訳に比べ

て MTPE のエラー数が少ない結果となった。「申し送り」では 1 回目に比べてエラー数は減少したが、人手翻訳より MTPE の方が多い結果となった。また「ユーザビリティ」ではエラー数は増加したが、MTPE の方が少なかった (図 4)。2 回めでエラー数を抑制できた背景としては、1 回目の作業を踏まえた改善としてポストエディターへのフィードバックや作業指示書の更新やなどの対策の効果が大きいと考えられ、PE の品質改善にはコミュニケーションの改善が有効である可能性が示唆された。

5. 結論

本研究では、人手翻訳と比べて品質を落とすことなく、さらなる翻訳スピードを追求することを目的として、科学的手法とエビデンスに基づく MT の導入モデルの構築を試みた。この導入モデルはカスタマイズ MT の開発・評価と品質水準の設定という 2 つの要素により構成された。カスタマイズ MT の開発にあたっては、MT の品質を測る指標 (BLEU、TER および編集率) と PE 作業時間を計測することで、MT の精度改善を確認した。また品質水準は必要以上に高い品質目標を立てるのではなく、水準を明示的に文書化することにより、合理的な品質を効率よく達成することを主眼にデザインした。これらのモデルの効果を検証するために、2 回にわたって実際の翻訳プロセスによる比較を行い、時間短縮効果を確認できた。

今後は、実際の業務における実測データをもとに、エビデンスの再現性を確認するよう検討したいと考えている。

「カスタム MT+フル PE」現状と見通し：医療翻訳分野の事例研究

山田 優¹ 早川 威士²¹立教大学 ²株式会社アスカコーポレーション

1. はじめに

医学・医療分野の翻訳の特徴は、高度な専門性がその1つと考えられる。例えば、語彙の特殊性や構文の複雑さなどがあげられる。また、医学・医療分野は、人類医療の資産となりうる研究や、法規制に基づく新薬の承認申請、患者の生命を左右する情報を扱うことから、正確性の要求品質が高い。機械翻訳（MT）のドメインアダプテーションないしカスタム MT という技術を使うことで、専門領域に特化した MT エンジンを作ることができる。その際に、「医学・医療分野（メディカル分野）」という括りでは専門性を絞りきれているとはいえない。より細分化された分野・用途に応じた文書特化型 MT エンジンの構築が必要となる。

さて、このようにして作ったカスタム MT を産業翻訳、とくに医学・医療翻訳の分野で、ポストエディットに用いるとした場合に、プロ翻訳者は、それをどのように見るか。作業効率は向上するのか。品質はどうか。MT の産業への実装という観点から興味深い問いである。

このような背景から、本研究では医学・医療分野の中でも詳細分野に特化した MT エンジンを作成し、そのカスタム MT を活用した実務翻訳における MT+フルポストエディット（以下、フル PE）の実用性を評価する。

2. カスタム MT の構築

本研究では、医療の文書特化型エンジンと医療汎用型エンジンの2種類の MT モデルを構築した。基盤となる MT モデルには情報通信研究機構（NICT）の汎

用 NT を使用した。まず、医学・医療分野から広く収集した教師データ 40 万文で訓練を行い、医療汎用型エンジンを構築した。続いて、英日方向については、医学論文のテキスト 13 万文を追加して訓練（学習の重みを強めるため data augmentation）を行い、医学論文に特化した MT を構築した（英日の文書特化型エンジン）。日英方向については、医薬品臨床試験の計画書から抜粋したテキスト 19 万文を追加して訓練を行い、日英の文書特化型エンジンを構築した。いずれも教師データには公開済みデータを原文として収集し、それを翻訳したものを使用した。

表 1 は、ある原文に対する英日の医療文書特化型エンジンと医療汎用型エンジンの出力の例である。例えば、原文の“chronic kidney disease”は、文書特化型エンジンでは「慢性腎疾患」、汎用型エンジンでは「CKD」と出力されているなど、用語の違いがみられる。

原文	Randomised, double-blind, placebo-controlled and parallel dose group trial to investigate efficacy and safety of multiple doses of oral XXX-12345 over 14 weeks, alone and in combination with empagliflozin, in patients with diabetic and non-diabetic chronic kidney disease
文書特化型エンジン	糖尿病性及び非糖尿病性慢性腎疾患患者を対象として、XXX-12345 の経口反復投与(単独投与及びエンパグリフロジンとの併用投与)の効果及び安全性を14週間にわたり検討する、無作為化、二重盲検、プラセボ対照、並行群間比較試験
医療汎用型エンジン	糖尿病及び非糖尿病性CKD患者を対象に、経口XXX-12345を単独又はエンパグリフロジンと併用で14週間にわたり反復投与したときの有効性及び安全性を検討する無作為化二重盲検プラセボ対照並行群試験

表 1 カスタム MT の出力例

BLEU スコアは、英日の場合、文書特化型エンジンで 51.11、医療汎用型エンジンで 47.56 であった。日英の場合、文書特化型エンジンで 44.50、医療汎用型エンジンで 40.82 であった。つまり、MT を医学・医療分野の中の特定の文書種に特化させることで

Title	
Randomised, double-blind, placebo-controlled and parallel dose group trial to investigate efficacy and safety of multiple doses of oral XXX-12345 over 14 weeks, alone and in combination with empagliflozin, in patients with diabetic and non-diabetic chronic kidney disease	

原文	Randomised, double-blind, placebo-controlled and parallel dose group trial to investigate efficacy and safety of multiple doses of oral XXX-12345 over 14 weeks, alone and in combination with empagliflozin, in patients with diabetic and non-diabetic chronic kidney disease
○ MT1	糖尿病性及び非糖尿病性慢性腎疾患患者を対象として、XXX-12345 の経口反復投与(単独投与及びエンバグリフロジンとの併用投与)の効果及び安全を 14 週間にわたり検討する。ランダム化、二重盲検、プラセボ対照、並行群間比較試験
☑ MT2	糖尿病性及び非糖尿病性慢性腎疾患患者を対象に、経口 XXX-12345 を 14 週間にわたり単独で及びエンバグリフロジンとの併用で反復投与したときの有効性及び安全性を検討する無作為化二重盲検プラセボ対照並行群間試験
選択理由	MT2 のほうが修正が単純であると判断した (文章の構造を変える必要が無い)。

カテゴリ一覧	使用する?
正確さ - 漏脱	☐
正確さ - 抜けと余分	☐
正確さ - 未翻訳	☐
正確さ - その他	☐
流暢さ	
流暢さ - 誤入力	☐
流暢さ - 誤字	☐
流暢さ - 同音異義語誤り	☐
流暢さ - 文法誤り	☐
流暢さ - 誤用	☐
流暢さ - 不適表現誤り	☐
流暢さ - 不統一	☐
流暢さ - 読解不能	☐
流暢さ - その他	☐
用語	
用語 - 指定用語違反	☐
用語 - 特定分野用語違反	☐
用語 - 用語不統一	☐
用語 - その他	☐
スタイル	
スタイル - 指定スタイル違反	☐
スタイル - 特定分野スタイル違反	☐
スタイル - スタイル不統一	☐
スタイル - その他	☐

コメントの追加 [A1]: 流暢さ - その他

あいまい

コメントの追加 [A2]: 用語 - 用語不統一

原文は chronic kidney disease だが、MT では CKD と略語に変換されている。

削除: CKD

図 1 実験作業環境の例

BLEU スコアは両言語方向において向上した。

本研究の目的は、このようにして構築した特化性の異なる 2 種類の MT をプロ翻訳者にフル PE してもらい、エラー・主観評価・修正量などの観点から、カスタム MT+フル PE の「実力」を評価することである。

3. 実験デザイン

実験参加者は英日、日英それぞれでプロ翻訳者 3 名 (計 6 名) であった。翻訳経験年数は最低 2 年以上、最高で 16 年以上のベテランである。ポストエディットの経験年数は 2 年未満がほとんどだった。

評価用テキストは、英日では医学文書 50 文、日英では治験文書 50 文を用意した。これらを文書特化型と医薬汎用型の 2 種類の MT につけ、参加者には MT の選択とフル PE を行ってもらった。実験に先立って、参加者には ISO 18587 (ISO, 2017) [1]に基づきフル PE の要求事項を説明した。

実験手順は以下の通りである。

1. 2 種類の MT 出力からいずれの出力が PE しやすいかを判断し、MT を選択する。(評価はセグメント (おおむね 1 文) 毎に行った。どちらの MT が特化型かは参加者には伏せた状態)。
2. 選択理由を書く。

3. 選んだ MT をフル PE する。

4. 修正箇所エラーカテゴリの付与をする (エラーカテゴリは JTF 品質評価ガイドライン (日本翻訳連盟, 2018) [2]の一部を使用)。

5. 修正箇所について修正理由を書く。

全てのテキストのフル PE と上記の作業が終了した後に、主観品質評価・効率性・認知負荷・満足度などの質問項目を含むアンケートを実施した。

図 1 は実験作業環境の例である。作業はワードマクロを組み込んだ MS Word 上で行った。上部に 1 つの原文テキスト全体が表示され、その下にはセグメントごとに原文、2 種類の MT 出力と MT の選択ボタン、そして MT の選択理由を記入するボックスが表示される。各修正箇所にはエラーカテゴリのオプションが表示され、カテゴリを選択すると修正理由をコメントできる。

4. 結果

4.1 MT の選好性

英日方向の実験において、文書特化型エンジン (MT1) と医薬汎用型エンジン (MT2) のうち選ばれたセグメントの割合を比べたところ、MT1 は 64%、MT2 は 36%であった。 χ^2 検定の結果、MT1 と MT2

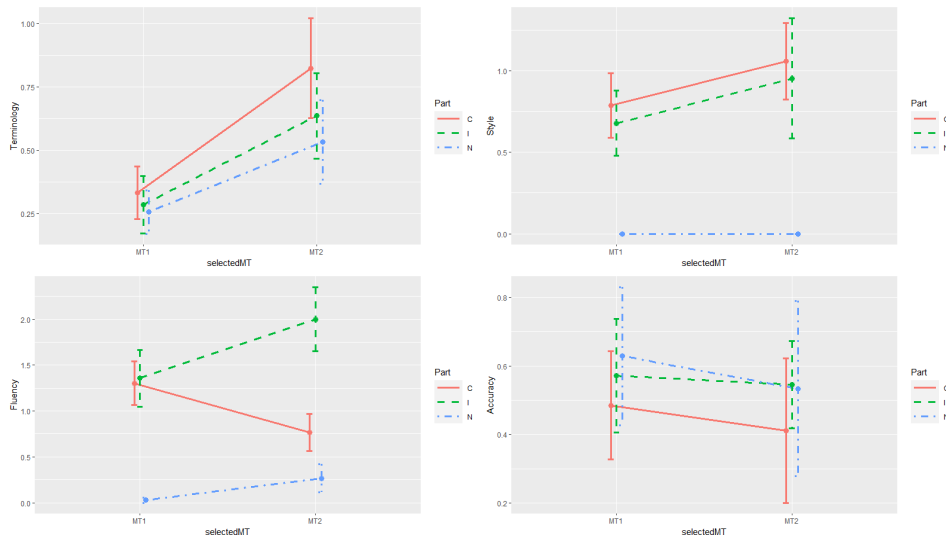


図 2 MT 間のエラー数の違い（英日）

の間には統計的に有意な違いがあった ($\chi^2(1)=11.76, p<.001, \phi=.28$)。

同様に日英方向の実験において、文書特化型エンジン (MT1) と医薬汎用型エンジン (MT2) の選ばれた割合を比べたところ、MT1 は 43%、MT2 は 57% であった。しかし、 χ^2 検定の結果、MT1 と MT2 の間に統計的に有意な違いはなかった。

4.2 MT の選択とエラーの関係

まず英日方向のフル PE についての分析結果を報告する。図 2 は、文書特化型エンジン (MT1) と医薬汎用型エンジン (MT2) でのエラー数の違いを示す。参加者ごとに線の種類を変え、エラータイプ（用語・スタイル・流暢さ・正確さ）ごとにグラフを分けた。丸い点はエラー数をセグメント数で割った平均値で、エラーバーは標準誤差を表す。

図 2 で、例えば、左上の用語エラーについては、MT1 のほうが MT2 よりも平均エラー数が少ない傾向がはっきりとみえる。具体的には、用語エラー (terminology) において、MT1 では 1 セグメント当たりのエラー数は、実験参加者 N において、平均 0.25 個である。しかし、MT2 では、0.50 強と約 2 倍に増えているのが観察できる。これに対して、スタイル、流暢さ、正確さエラーについては参加者全員に共通し

てみられるような傾向はない。

これらのグラフの傾向をみて、どの種類のエラーの増減が MT1 か MT2 を選択するののかの決定に寄与しているのかを調べるためのモデルを立てた。被説明変数を MT1 と MT2 の選択、説明変数を各セグメントにおける各エラーカテゴリー（用語・スタイル・流暢さ・正確さの 4 種類）のエラーの数、そして変量効果を 3 名の参加者とした一般化線形混合モデル (generalized linear mixed effect model) を立てた。

その結果、表 2 に示すように用語エラーのみに統計的に有意な差が見られた。つまり、MT1 と MT2 の選択に、用語エラー（の数）が深く関わっている。しかしスタイル・流暢さ・正確さエラーはあまり大きな影響を与えていないことが明らかとなった。またこれに加えて、被説明変数と変量効果は同じで、説明変数を各セグメントにおける Edit Distance Rate (EDR≒修正量) としたモデルも検討したが、有意にはならなかった。

日英方向のデータについても同様の分析を行った。しかし、参加者全員に共通する傾向はどのエラーカテゴリーにもみられなかった。また、一般化線形混合モデルも全てのエラーカテゴリーで有意な結果にはならなかった。つまり、MT1 と MT2 の違いに深く関連するエラーカテゴリーを特定することはできなかつ

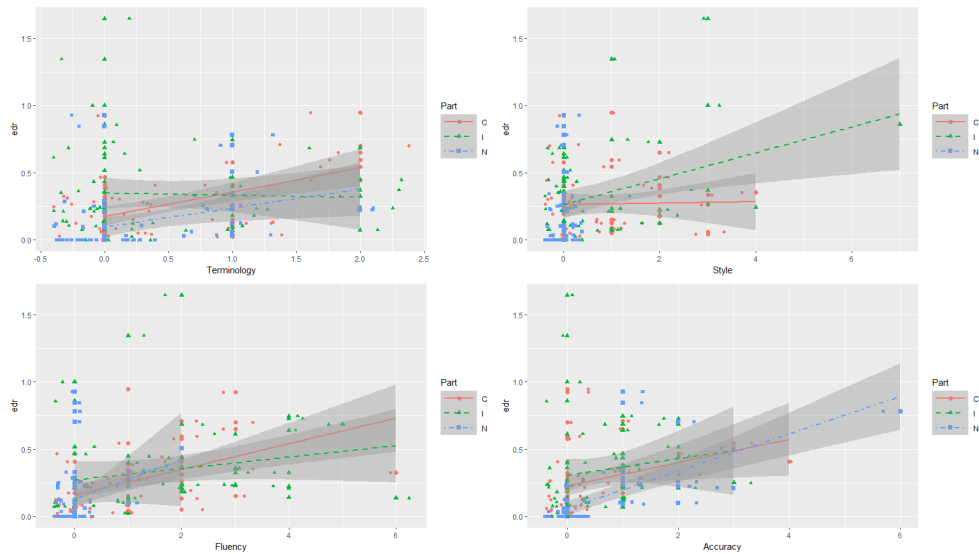


図 3 EDR とエラー数の関係（英日）

た。

selectedMT				
Predictors	Odds Ratios	std. Error	CI	p
(Intercept)	0.38	0.08	0.25 – 0.58	<0.001
Terminology	2.36	0.63	1.40 – 3.98	0.001
Random Effects				
σ^2	3.29			
τ_{00} Part	0.00			
N Part	3			
Observations	150			
Marginal R ² / Conditional R ²	0.089 / NA			

表 2 一般化線形混合モデル（英日）

4.3 修正量とエラーの関係

まず、英日方向の結果を示す。図 3 は EDR とエラー数の関係を示す。y 軸は EDR で、x 軸は 1 セグメントにおける各エラーカテゴリー（用語・スタイル・流暢さ・正確さ）のエラー数である。散布図の点は結果がわかりやすいように多少 x 軸方向にずらしで表示している（そのためマイナスにみえるデータがある）が、回帰直線は散布図に依存しないように表示しているので解釈には影響しない。3 種類の線は参加者を表す。下 2 つの流暢さ、正確さエラーのグラフをみると、どの参加者でも MT が出すエラー数が多いセグメントほど EDR（フル PE での修正量）も大きくなる傾向がある。それに対して、上 2 つのスタ

イルエラーと用語エラーのグラフでは、参加者全員に共通するような傾向はみられない。

これらのグラフの結果にたいして、被説明変数を各セグメントの EDR、説明変数を各セグメントにおける各エラーカテゴリー（用語・スタイル・流暢さ・正確さ）のエラーの数、変量効果を 3 名の参加者とした線形混合モデル（linear mixed effect model）をたてた。まず、説明変数がひとつのみのモデルを 4 つのエラーカテゴリーそれぞれで検討した結果、4 つ全てで統計的に有意な結果になった。その結果を踏まえ、説明変数を複数組み合わせたモデルをたて、ステップワイズ法でモデル選択をした結果、説明変数を「スタイル+流暢さ+正確さ」としたモデルが選ばれた（表 3）。この結果から、スタイル、流暢さ、正確さエラーが EDR に大きな影響を与えていることがわかった。

edr				
Predictors	Estimates	std. Error	CI	p
(Intercept)	0.10	0.03	0.05 – 0.16	<0.001
Fluency	0.06	0.01	0.04 – 0.09	<0.001
Accuracy	0.09	0.02	0.05 – 0.13	<0.001
Style	0.07	0.02	0.03 – 0.10	<0.001
Random Effects				
σ^2	0.06			
τ_{00} Part	0.00			
N Part	3			
Observations	150			
Marginal R ² / Conditional R ²	0.293 / NA			

表 3 線形混合モデル（英日）

この結果は、すなわち、翻訳者が MT 出力をたくさんフル PE しなければならない時というのは、その MT 出力の中に、スタイルエラー、流暢さエラー、正確さエラーが多く含まれている時であることを表す。この中に、用語エラーが含まれていないということは、逆にいうと、プロ翻訳者が良い MT エンジン (MT1 か 2) を選ぶ理由は、用語が正しく訳出されているからであると推察される。

日英方向のデータについても同様の分析を行った。EDR と各エラーカテゴリーのエラー数の関係を観察したところ、3 名の参加者全員に共通して、用語・流暢さ・正確さエラーが多いセグメントほど EDR が高くなっている傾向がみられた。しかし、スタイルエラーについてはそのような傾向はないようであった。この結果について英日の場合と同様に、被説明変数を EDR、説明変数を各エラーカテゴリーのエラーの数、変量効果を 3 名の参加者とした線形混合モデルをたてて検証したところ、説明変数を「用語+流暢さ+正確さ」としたモデルが選ばれた (表 4)。したがって、日英においては、用語、流暢さ、正確さエラーが EDR に大きく影響していることが予想される。

Predictors	edr			
	Estimates	std. Error	CI	p
(Intercept)	0.11	0.04	0.02 – 0.20	0.017
Terminology	0.10	0.03	0.05 – 0.15	<0.001
Fluency	0.09	0.02	0.04 – 0.14	<0.001
Accuracy	0.07	0.02	0.04 – 0.10	<0.001
Random Effects				
σ^2	0.03			
τ_{00} Part	0.00			
ICC	0.13			
N Part	3			
Observations	150			
Marginal R ² / Conditional R ²	0.244 / 0.341			

表 4 線形混合モデル (日英)

5. まとめ

本研究では、医学・医薬分野において特定の文書種に特化した MT を構築し、プロの翻訳者にフル PE を行ってもらうことで、エラー・主観評価・修正量の観

点から MT の「実力」の分析を試みた。その結果、以下のような結論が得られた。

1. MT エンジン医学・医薬分野の中でも細分化された文書種に特化させることで、BLEU スコアが英日・日英の両言語方向において向上した。
2. 英日方向においては、文書特化型エンジンが医薬汎用型エンジンよりも有意に多く選ばれた。
3. 英日方向においては、用語エラーが MT の違いに大きく関わっていることがわかった。
4. 英日方向においては、用語エラー以外が EDR の予測に大きく影響していた。これは 3 を裏付ける結果であると考えられる (ただし、今回は MT の違いに EDR 自体があまり影響を与えていないという結果であったことに注意が必要である)。
5. 日英方向においては、カスタム MT の品質の差が、フル PE 作業に影響していることが、はっきりとは確認できなかった。

謝辞

本研究の一部は、日本学術振興会科研費補助金基盤研究(S)「翻訳規範とコンピテンスの可操作化を通して 翻訳プロセス・モデルと統合環境の構築」(研究課題番号:19H05660)の支援を受けて行われた。

参考文献

- [1] International Standard Organization. (2017). *Translation services – Post-editing of machine translation output – Requirements* (ISO Standard No. 18587:2017). <https://www.iso.org/standard/62970.html>
- [2] 日本翻訳連盟 (2018) 「JTF 翻訳品質評価ガイドライン」 [Online] https://www.jtf.jp/tq/pdf/jtf_translation_quality_guidelines_v1.pdf

特許翻訳におけるドメイン適応型機械翻訳

本間 奨¹⁾、金子 清隆²⁾、東内 健一³⁾、江崎 美江⁴⁾、大崎 慎治⁵⁾

1)日本特許翻訳（株）、2)コニカミノルタ（株）、3)リコーテクノロジーリサーチ（株）、4)ブラザー工業（株）、5)京セラドキュメントソリューションズ（株）

1. はじめに

AI 翻訳（ニューラル機械翻訳）技術の進展が著しい。特許業界では特許庁をはじめ¹⁾、特許出願明細書の翻訳に AI 翻訳を取り入れ翻訳の迅速化並びに生産性を向上する動きが加速してきている。²⁾、³⁾ 本報告では、従来の Google や DeepL などの汎用型エンジンに対して、特許や医薬などの産業翻訳で要求される特定分野の特定文型や用語に適切に対応可能なドメイン適応型機械翻訳について解説を行う。続いて、電子写真分野におけるドメイン適応型機械翻訳の検証事例を紹介する。最後に今後の展開についてまとめた。

2. 特許翻訳における汎用型と分野特化型とドメイン適応型機械翻訳について

1) 汎用型翻訳システム

適用分野が広く、また扱える言語も広く選択できる。代表的な例として①Google 翻訳、②DeepL 翻訳、③国立研究開発法人情報通信研究機構（以下 NICT）汎用 NT、④みらい翻訳などの翻訳システムが知られている。

2) 分野特化型翻訳システム

特許や医薬、法務などの各分野に特化したタイプで、扱える言語は限定され、精度は汎用型とドメイン適応型機械翻訳システムの間である。ドメイン適応型翻訳エンジンのベースエンジンとしても用いられる。具体的な例として① NICT 特許 NT、②ロゼッタ T-3MT、③ Mirai Translator「契約書・法務モデル」などがある。

3) ドメイン適応型機械翻訳システム

訓練データと同じ言語および同種の原文に限定されるが最も高精度である。顧客の保有する対訳や翻訳メ

モリをもとに独自翻訳システムを構築する。ベースエンジンに対して、対訳による追加学習（アダプテーション）したエンジンに基づく翻訳システム。対訳文の精度が高い場合、この訓練データと同種の原文については最高精度の翻訳結果を得ることが可能であるが適用範囲は限定される。

図 1 に国際特許分類 G06F、出願人富士通での請求項に関する精度評価結果を示した。約 4,000 文対から構築された NICT によるドメイン適応型エンジンが BLEU、RIBES 評価で最高精度を示すことがわかる。

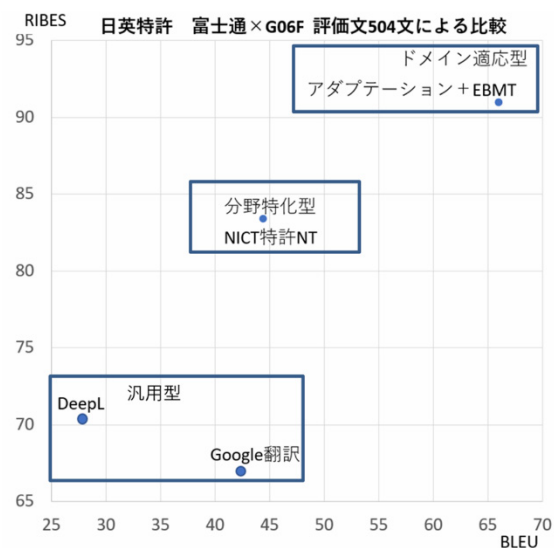


図 1 富士通×G06F 請求項ドメインの公報の対訳 504 文対による汎用型、分野特化型、ドメイン適応型機械翻訳システムの精度比較

Domain-adaptive machine translation in patent translation

Susumu Honma, Kiyotaka Kaneko, Kenichi Tohnai, Mie Ezaki, Shinji Osaki

1)Nihon Patent Translation Co. Ltd. 2) KONICA MINOLTA, INC. 3)Ricoh Technology Research, Inc. 4)BROTHER INDUSTRIES, LTD. 5) KYOCERA Document Solutions Inc.

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Public License. License details: <https://creativecommons.org/licenses/by-sa/4.0/>

3. ドメイン適応型機械翻訳の事例紹介

筆者達が所属する一般社団法人ビジネス機械・情報システム産業協会（以下 JBMIA）知的財産委員会では、ビジネス機械・情報システム分野の特許明細書を高精度に日英翻訳するニューラル自動翻訳実現に向けフイージビリティ・スタディに取り組んでいる。⁴⁾

特許明細書の翻訳では、高精度な訳文に加えて訳語の統一が重要である。そこで、訳揺れの少ない高精度な翻訳システムの実現可能性を検証すべく、日本特許翻訳（株）の協力を得て 2021 年 12 月にドメイン適応型機械翻訳システムを構築した。ここでは、翻訳の高精度化と訳語統制を実現するためのポイントと、先行評価に基づく検証結果について報告を行う。

1) 訓練データ用対訳コーパス作成のポイント

①ドメインの絞込み

少量の訓練データでも、アダプテーション効果を確実に得られるように、対象ドメインを「電子写真装置の基本構成部分」に限定した。

②対訳コーパスの整備

一件の特許明細書の中では、原文も訳文も用語統一されている。しかし、各社から収集した対訳コーパスでは、各社毎に異なる用語で統一されているため、全体として原文と訳文共に用語統一されていない。更に、好ましい訳文とするために、意訳文や原文にはない主語あるいは目的語を補った訳文も含まれているが、これらは訓練データとして好ましいものではない。

そこで、訳揺れを抑え高精度な翻訳を実現するために、対訳コーパスのガイドラインを定めて全文対を点検整備し、6,700 文対の訓練データ用対訳コーパスを作成した。

ガイドラインの主要項目を以下に記す。

1. 原文と訳文の態を同じとする
2. 原文と訳文とは、内容の省略も追加もしない

	原文	訳文
編集前	感光ドラム上に形成されたトナー像は、記録紙または中間転写ベルト上に転写され、定着器の熱や圧力により記録紙に定着される。	The toner image formed on the photosensitive drum is transferred onto a recording sheet or an intermediate transfer belt, fixed to the recording sheet by heat or pressure from a fixing unit.
編集後	感光体ドラム上に形成されたトナー像は、用紙あるいは中間転写ベルト上に転写され、そして定着器からの熱や圧力により用紙に定着される。	The toner image formed on the photosensitive drum is transferred onto a sheet or an intermediate transfer belt and fixed to the sheet by heat and/or pressure from a fixing device.

表 3-1 対訳コーパスの整備

3. 原文と訳文は各一文ずつにする
4. 訳文で不要な括弧は原文から削除する
5. 原文は漢数字を使う（参照符号、数値を除く）
6. カタカナ表記の長音記号（ー）はルールに従う
7. 専門用語は JBMIA 標準用語集に従う

図 3-1 に、ガイドラインによる用語統制の考え方を示す。表 3-1 にガイドラインに従って整備した訓練データ用対訳コーパスの例を示す。

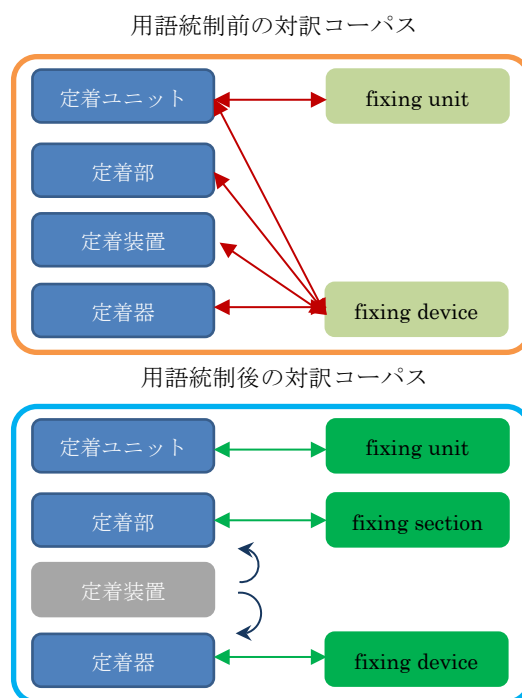


図 3-1 ガイドラインによる用語統制の考え方

2) ドメイン適応型機械翻訳システムの翻訳精度

NICT 特許 NT（以下 NT）をベースエンジンとし、6,700 文対の対訳を追加学習したドメイン適応型エンジン（以下 TM）は、図 3-2 に示すように翻訳精度が大きく向上し、外国特許出願用の明細書翻訳に利用することで、翻訳の効率化を期待できる。

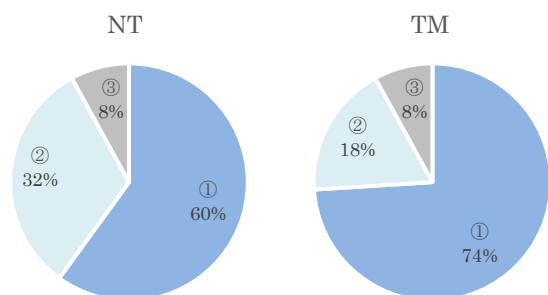


図 3-2 翻訳精度比較 (200 文を評価)

評価用原文 200 文を用いて、NT と TM の翻訳精度を評価した。なお、評価用原文も訓練データと同じドメインを対象にした。評価者によるバラつきを防ぐため、JBMTA が作成した翻訳品質評価シートに基づき、各訳文について、訳漏れ、過剰訳、誤訳、文法、一般用語、専門用語、流暢さ等を個別に検討後、US 出願した場合の補正の要否に基づいて、表 3-2 に記載した三段階評価を行った。

評価	判定ガイド (US 出願用の翻訳文として判定)
①	このままで使用可能 (出願後の補正不要)
②	軽微な修正が必要 ※そのままで意味は通じる (必要に応じて補正要)
③	軽微でない修正が必要 ※意味が通じない (補正が必須)

表 3-2 評価基準

表 3-3 に示すように、TM の BLEU 値と RIBES も向上しており、特に BLEU 値の向上が際立っている。これらの結果からも、ドメイン適応による高精度化が図られていると考えられる。

翻訳エンジン	BLEU 値	RIBES
NT	49.6	88.2
TM	88.3	96.2

表 3-3 翻訳エンジン性能比較

3) 用語統制効果

翻訳精度の向上と共に、意図したとおりに訳語が選択され、且つ訳語の揺れも抑制できた。

一例として、a) の評価用原文に含まれている JBMTA 標準用語集の用語 (下線付き) に着目し、b) NT の訳文と c) TM の訳文で使用された訳語の調査結果を表 3-4 に示す。TM 訳文では JBMTA 標準用語集で定めた訳語に統一されていることが確認できた。

同様の手法で、複数の評価用原文について調査した結果を表 3-5 および表 3-6 に示す。何れも、TM では訳揺れが生じておらず、指定した訳語で訳出されていることを確認できた。なお、各表のタイトルは、JBMTA 標準用語集で規定した用語と訳語の対を表し、数値は各エンジンの訳文における訳語の出現率を示す。

a) 評価用原文

給紙ユニット 3 は、給紙カセット 31 と、ピックアップローラー 32 と、給紙ローラー 44 と、分離ローラー 45 とを含む。

b) NT 訳

The paper feed unit 3 includes a paper feed cassette 31, a pickup roller 32, a paper feed roller 44, and a separation roller 45.

c) TM 訳

The sheet feed unit 3 includes a sheet feed cassette 31, a pickup roller 32, a sheet feed roller 44, and a separation roller 45.

用語	指定訳語	NT	TM
給紙ユニット	sheet feed unit	NO	YES
給紙カセット	sheet feed cassette	NO	YES
給紙ローラー	sheet feed roller	NO	YES
分離ローラー	separation roller	YES	YES

表 3-4 用語統制効果 (1)

用語: 画像形成部、指定訳語: image forming section			
訳文中で使用された訳語	NT	TM	
image forming section	8%	100%	
image forming unit	84%	0%	
image forming portion	8%	0%	

表 3-5 用語統制効果 (2)

用語: 給紙ユニット、指定訳語: sheet feed unit			
訳文中で使用された訳語	NT	TM	
sheet feed unit	0%	100%	
paper feed unit	66%	0%	
sheet feeding unit	17%	0%	
paper feeding unit	17%	0%	

表 3-6 用語統制効果 (3)

4. 今後の展開

先行評価結果では、訳揺れがなく且つ指定した用語で訳出される高精度な翻訳システムを、ガイドラインに沿った統一的な対訳コーパスを用いたドメイン適応型機械翻訳システムで実現出来る可能性が示された。

然しながら、先行評価の原文は、訓練データと同じ「電子写真装置の基本構成部分」のドメインに限定し、且つ、明晰な日本語文章としていた。今後は、他のドメインや、様々な日本語文章を用いた評価を行い、ドメイン適応型機械翻訳システムを外国出願用明細書翻訳に用いたときの有用性を確認すると共に、対象とするドメインの拡張についても検討を進める。

また、JBMIA では特許明細書の日本語原文を明晰化する活動（産業日本語の活用）も推進している。明晰な日本語文章であれば、ドメイン適応型機械翻訳システムを利用して翻訳することで、ポストエディットの負荷を大幅に軽減可能となることを立証し、他の産業分野でもこの翻訳システムが採用されることを目指している。そのためにも、官民一体で産業日本語とドメイン適応型エンジンの普及を加速させ、日本企業の国際競争力強化を牽引すると共に、特許明細書の誤訳訂正等に対する柔軟な法的対応についても国際協調の下で実現が図られるよう働き掛けを行いたい。

参考文献

- (1) 東芝デジタルソリューションズ株式会社
ニュースリリース 2018 年 7 月 10 日
- (2) (一社) 日本知的財産協会「知財管理」Vol.70
No.9 2020 pp.1253～1264
- (3) (一社) 日本弁理士会 パテント Vol.75 No. 2
2022 「特許翻訳におけるドメイン適応型機械翻訳」 pp.1～6
- (4) (一社) ビジネス機械・情報システム産業協会
ニュースリリース 2019 年 12 月 9 日
<https://www.jbmia.or.jp/whatsnew/detail.php?id=1255>

AAMT 2021, Online

創立 30 周年記念イベント

出内 将夫
情報通信研究機構

1. はじめに

2021 年 12 月 8 日～9 日に開催された AAMT 2021, Online のイベント報告をします。コロナ禍の影響により、2020 年に続いてオンライン開催となる中、AAMT 創立 30 周年を記念して 2 日に渡って開催されました。記念講演 3 件、長尾賞/学生奨励賞受賞者講演 3 件、年次大会発表 6 件、公募セッション発表 6 件と密度の高いイベントでした。また、Web サイト上に企業展示コーナーが設けられ、商品紹介も行われていました。

本イベントの講演/発表資料は <https://aamt.info/aamt-2021-online/handout-20211208/> からアクセスできますので、ぜひ、これらの資料をご参照ください。

2. 記念講演

記念講演は、以下の 3 件でした。

◎AAMT と MT の軌跡と未来：

黒橋 禎夫氏／京都大学大学院 教授

◎翻訳業界と MT との過去、現在と未来：

安達 久博氏／（一社）日本翻訳連盟 代表理事・会長

◎長尾先生の言語観と例からの機械翻訳：

辻井 潤一氏／（国研）産業技術総合研究所 フェロー
同 人工知能研究センター 研究センター長

下記のまとめは、講演資料と発表内容より、報告者が主観的にまとめたものであることにご注意ください。

◎AAMT と MT の軌跡と未来：黒橋 禎夫氏

機械翻訳(MT: Machine Translation)は自然言語処理の発展にも深く関係しており、1940 年～1960 年中

頃までの黎明期、1990 年までの忍耐期を経て、発展期に入り、2015 年から飛躍期に入った。ニューラル機械翻訳(NMT: Neural Machine Translation)に至るまでには、用例翻訳や統計機械翻訳(SMT: Statistical Machine Translation)と技術の変遷があり、現在は Transformer[1]を用いた NMT が広く使われている。

1991 年に AAMT の前身である日本機械翻訳協会(JAMT: Japan Association for Machine Translation)が設立され、翌年に AAMT に改称された。AAMT では、毎年 AAMT 長尾賞や学生奨励賞を選考するほか、アジア翻訳ワークショップ(WAT: Workshop on Asian Translation)の活動を支援している。今後は、MT 活用がマルチモーダルや異文化の対話支援、実用翻訳でさらに進み、AAMT の活動がアジア太平洋地域に拡大していくと考えている。

◎翻訳業界と MT との過去、現在と未来：安達 久博氏

AAMT の前身である JAMT は、大学関係者や翻訳会社、電機メーカーを主なメンバーとして発足した。初期の AAMT ジャーナルの広告からは、1992 年時点で、すでに MT サービスの提供やポストエディット(PE: Post Edit)が行われていたことが伺える。AAMT を設立された長尾先生には、日本翻訳連盟(JTF: Japan Translation Federation)の活動でも大変お世話になり、「MT と人間の翻訳がうまく協調し、より良い翻訳を世の中に提供していくべき」という言葉をいただいた。JTF では、MT を効果的に使うために対訳や用語集の整備を進めており、2019 年に用語バンク委員会を設立した。今後、様々な業界に役立つ用語集を提供していく予定である。

◎長尾先生の言語観と例からの機械翻訳:辻井 潤一氏

AAMT 設立前から言語を工学的に扱う取り組みは始まっており、データから言語を帰納的に捉える構造主義、人間が心で言語がどう捉えているかを重視する認知主義といった観点で研究が進められてきた。長尾先生は、構造主義的、認知主義的な傾向の双方をお持ちであった。

また、長尾先生は、社会への貢献を重視されていた。1980 年代に行われた計算機で自然言語を扱うための種々のプロジェクトのうち、MU プロジェクトを主導されていた。1987 年に箱根で立ち上げた MT Summit も長尾先生が中心となられ、政府や民間企業など多様な立場の方が参加されていたのが印象的だった。

長尾先生は、「言語は様々な要因の重なりで発現している複雑なもの」であり、言語を「膨大な有限集合」として捉えておられた。特に類推や帰納、メタファーに興味をお持ちだった。長尾先生は 1984 年に「例による翻訳」[2]を発表されており、他の種々の研究より数年早く、問題の本質を捉えられていた。MT は言語的文脈の翻訳ではかなりうまく行っているが、言語的文脈だけでは読み取れない情報も考慮するのが次のステップではないかと考えられる。

3. 長尾賞/学生奨励賞受賞者講演

長尾賞/学生奨励賞受賞者講演では、第 8 回長尾賞学生奨励賞 1 件と第 16 回 AAMT 長尾賞 2 件の、計 3 件の受賞講演が行われました。

第 8 回長尾賞学生奨励賞 (1 件)

◎ Learning to Describe Unknown Phrases with Local and Global Contexts :

石渡 祥之佑氏/株式会社 Mantra

第 16 回 AAMT 長尾賞 (2 件)

◎WMT-2020 ニュース翻訳タスクに参加して :

Team Tohoku-AIP-NTT at WMT-2020

鈴木 潤氏/東北大学

◎Towards Fully Automated Manga Translation :

石渡 祥之佑氏/株式会社 Mantra

それぞれの講演について、前述の講演資料や講演内容から、次のとおりまとめます。

◎ Learning to Describe Unknown Phrases with Local and Global Contexts : 石渡 祥之佑氏

受賞した博士論文の研究のうち、最も新しいトピックである、未知のフレーズを説明するテーマ[3]について発表する。まず、翻訳対象に未知のフレーズがあった際に、そのフレーズと周辺の文脈を入力とし、フレーズの説明文を生成するタスクを定義し、データセットを構築した。そして、フレーズの近傍の語を考慮するローカルな文脈と、膨大な用例を考慮するグローバルな文脈の双方を扱える手法を提案した。データセットは、Wikipedia 記事のリンクに関連付けられた Wikidata のエントリを元に構築した。フレーズの説明文生成では、単語を対象とした既存研究で用いられていた語ベクトルや、表層の文字に加え、周辺の語とそのアテンションを考慮する手法を提案した。既存のデータセットと提案データセットの双方を用いて、生成文の BLEU 値を評価し、すべてのデータセットで提案手法が既存手法を上回った。

◎WMT-2020 ニュース翻訳タスクに参加して :

Team Tohoku-AIP-NTT at WMT-2020 : 鈴木 潤氏

WMT-2020 のニュース翻訳タスクに参加したシステム[4]の概要と、システム開発で得られた知見、コンペティション参加の意義について発表する。WMT は機械翻訳で最も有名かつ競争の激しいコンペティションで、参加したのはニュース翻訳タスクのうち、英独、独英、英日、日英の 4 トラックである。参加したシステムは、英独と独英の公式な自動評価と人手評価で 1 位を獲得したほか、他の論文で言語現象別に評価した論文でも最良の評価を得た。参加システムでは、逆翻訳によるデータ拡張、ハイパーパラメータの調整、ファインチューニング、アンサンブル、リランキングの処

理を組み合わせた。開発途中の試行錯誤段階で、データ選択処理の導入を検討したが、BLEU スコアが上がらず、導入を断念した。コンペティションに参加することで、自他の技術レベル把握や、最先端研究に必要な技能や実験環境が得られ、様々な技術が培える。最高性能の追及だけでなく、他の実験に役立てられるので、コンペティションへの積極的な参加をお勧めしたい。

◎Towards Fully Automated Manga Translation :

石渡 祥之佑氏

マンガは世界中で楽しまれているが、正規のサービスは行き渡っていない状況である。そこで Mantra では、マンガの海外展開や外国語教育用コンテンツとして、マンガ翻訳に取り組んでいる。マンガ翻訳には、評価用データが無く、文脈の考慮、絵の考慮が必要という難しさがある。まず、評価セットの構築に取り組み、5 つのジャンルに跨るマンガ翻訳のデータセットを作成した[5]。文脈の考慮では、複数の吹き出しに跨る文の翻訳を行うため、1 コマに含まれる複数の吹き出しを文脈とした文脈翻訳を行っている。コマや吹き出しの領域検出と順序推定では、画像処理を活用している。マンガでは会話文が多く、話者情報が必要となるため、コマの画像認識結果を visual タグとして NMT に入力している。現状では、画像認識誤りが翻訳精度低下を招いており、テキストの文脈のみ考慮した方が高精度である。今後は、より広範囲な文脈の考慮や、吹き出しの形状による心情の考慮、吹き出しの形から話者を特定するなどに取り組む予定である。

4. 年次大会

年次大会発表は次の 6 件でした。

◎令和 2 年度特許出願技術動向調査「機械翻訳」:

成瀬 博之氏／特許庁

◎機械翻訳サミット参加報告:

田中 英輝氏／情報通信研究機構

◎ The 8th Workshop on Asian Translation (WAT2021) 報告:

中澤 敏明氏／東京大学

◎Red Hat が目指す翻訳の未来 - 全冊公開により見えてきた課題とソリューション:

燃脇 綾子氏／レッドハット株式会社

◎MT 利用ガイドライン委員会の取り組みと今後の予定:

山田 優氏／立教大学

◎MT 提供および利用の法的課題について:

柿沼 太一氏／STORIA 法律事務所

それぞれの発表について、前述の資料と発表内容より次のとおりまとめます。

◎令和 2 年度特許出願技術動向調査「機械翻訳」:成瀬 博之氏

特許庁では、日本の研究開発の方向を見定めるために、注目度の高い技術テーマを対象に、出願動向等を調査している。過去約 10 年間の特許と論文を対象とした調査の結果を紹介する。

MT の市場は拡大傾向で、2025 年にかけてアジア太平洋地域の成長率が高くなる見込みである。MT に関連する政策動向として、総務省のグローバルコミュニケーション計画 2025 があり、2025 年を目途に「文脈・話者の意図等を補う同時通訳」の実現を目指している。特許出願傾向では、出願先国として米国が主戦場となっており、中国籍が出願人の特許が 2016 年以降急増している。論文発表件数は 2014 年以降急増しており、欧州国籍の比率が首位となっている。

調査の総括として、①機械処理しやすいデータの共有や MT のリテラシー向上といったデジタルトランスフォーメーションの推進、②分野適用やセキュリティ確保などの MT 活用サービスの付加価値向上、③MT を活用した翻訳業務の生産性向上、④同時通訳の技術開発、の 4 点を提言する。

◎機械翻訳サミット参加報告：田中 英輝氏

2021年8月16日から20日にかけて、遠隔会議形式で開催されたMT Summitについて、ユーザー・プロバイダトラックを中心に報告する。ユーザー・プロバイダトラックではMTシステムの出力を扱う報告が多数を占めた。

報告の一つに、MTシステムを用いて英語ニュースを日本語に翻訳して配信するサービスの大規模配信実験の報告があった[6]。配信実験では、ユーザーからリクエストがあればPEも行う機能も併せて提供していた。被験者のアンケートからは、MTの直接配信でも7割以上の記事が適切に翻訳されており、139名の被験者のうち約6割から継続利用希望の回答があった。

別の報告では、人間翻訳(HT: Human Translation)と、MT後のPE作業の関係を分析していた[7]。分析の結果、HTの翻訳品質が高い作業者ほどPEの翻訳品質も高いという相関関係が見られ、作業者にとってHTとPEには同様のスキルが求められると分析されていた。また、作業者のレベルによってはPEを行うことで、MTより翻訳品質を悪化させることもあるという報告があった。

◎ The 8th Workshop on Asian Translation (WAT2021) 報告：中澤 敏明氏

2021年に開催された、アジア言語を対象とした機械翻訳評価ワークショップであるWATについて報告する。2021年のWATでは、過去最高となる国内外26チームの参加があった。招待講演では101言語のMT評価用データセットや、Graham Neubig先生によるACMの論文解説講演があった。WATは、年々ドメインや言語を拡充しており、参加チームはタスクを選択して参加する。参加後は、MTの結果を自動評価サーバーにアップロードして、評価結果を参照できる。

WAT2021でのタスクをいくつか紹介する。特許翻訳のタスクでは、日英・日中・日韓の双方向の言語方向で特許を対象に翻訳精度を競った。追加リソースを使う条件下で自動評価の一位を獲得したのは Bering

Labで、追加リソースを使わない条件で自動評価が一番高かったのは東京都立大学だった。他のタスクとしては、NMTが苦手とする訳語統一を条件とした制限翻訳、画像情報を加味して翻訳するマルチモーダル翻訳があった。

WATは、アジアの翻訳研究活性化やデータ整備等を目的として2014年に始まり、一定の成果を得ている。今後も言語対増加や、実用上必要となるタスク設計やデータの準備を行っていく。

◎Red Hatが目指す翻訳の未来 - 全冊公開により見えてきた課題とソリューション：燃脇 綾子氏

Red Hatは、DeepLを利用して製品マニュアルを全冊和訳して公開した。MTによるマニュアル公開には賛否両論があり、翻訳提供側からは「MTには誤訳が含まれるため適切でない。」という声がある一方、サポートチームからは「MTでも公開することでユーザーの手間を減らせる。」という声があった。

現在は、一部手動部分はあるが、ほぼ自動で最新の英語版マニュアルを翻訳した日本語版マニュアルを公開している。毎週日曜に英語版マニュアルの更新を自動検出し、TM適用後にMTを行い、日本語版を公開している。MT部分に対して優先順位付きの作業リストが作られ、翻訳者はリスト順にLightPE作業を行う。まずは一通り人的作業を経ることを目標としている。

2021年7月時点で、日本語版マニュアル公開のリードタイム中央値は、MT利用前の110日から4日に改善され、毎週の英語マニュアルの公開/更新150万語に対応している。マニュアルの未翻訳が原因で発生していた、別バージョンの製品マニュアルに対する誤訳報告が減り、サポートの手間が省けた。MT導入後の反応も好意的なものが多く、MTによってもたらされる未来はあかると考えている。

◎MT利用ガイドライン委員会の取り組みと今後の予定：山田 優氏

健全かつ正しい MT の利用推進を目的として、MT 利用ガイドラインを作成している。MT 利用ガイドラインでは、メインメッセージとして「MT は間違えるため、MT を過信せず、人手による精査・確認が必要である。」ということを伝えた上で、適切な MT 利用を促進する。

MT 利用ガイドラインは、翻訳の原理、パラメータ、ガイドラインのガイドライン、という構成となる予定である。翻訳の原理では、翻訳の原理原則について説明する。具体的には、翻訳品質の評価基準や、翻訳のプロセス、翻訳対象の分野、翻訳に関する法令、MT に関する技術に言及する予定である。パラメータでは、MT の利用シーンや使い方のコツを説明する。ガイドラインのガイドラインは、翻訳の原理とパラメータに基づくサマリとしてまとめる予定である。

2022 年 3 月末に MT ガイドラインの日本語版を FIX し、4 月で英訳、5 月に出版準備を行い、6 月のリリースを予定している。

◎MT 提供および利用の法的課題について：柿沼 太一氏

誤訳があった場合の翻訳者の法的責任、翻訳会社がクライアントから受領した翻訳対象文書の扱い、MT の学習用データとして用いることができるデータ、について、具体的設例を交えて説明する。

誤訳があった場合の翻訳者の法的責任については、当事者間の契約に取り決めが無い場合、民法に基づき発生した損失に応じて上限なく契約不適合責任が発生する。そのため、当事者間の契約において条件や賠償義務の限定を行うことが望ましい。参考として DeepL の利用規約でも同様の限定が行われている。

次に翻訳会社がクライアントから受領した翻訳対象文書の扱いについて述べる。翻訳原稿がクライアントの秘密情報に該当する場合、第三者に開示するにはクライアントの承諾が必要となる。そのため、翻訳会社が外部 MT サービスを利用する場合に、その外部 MT サービス提供者が入力データを外部 MT サービス提供者自身のためにアクセスする利用条件だと、翻訳会社

はクライアントに対して契約違反となる可能性がある。

MT の学習用データとして用いることができるデータについては、データの種類とデータの提供・取得方法の組み合わせにより、契約や法律の規制を受けるか判断する必要がある。データ取得時に契約があれば、契約の規制を受ける。ただし、Web ページ上に注意事項などが記載されているのみで、同意が成立する行為を行わずにダウンロード可能なデータであれば、契約は成立していないと考えられる。法律については、MT の学習は情報解析にあたり、著作権の例外規定に該当するため、著作権法上は問題とならない。

5. 公募セッション

公募セッション発表は次の 6 件でした。

◎機械翻訳と人間翻訳の両方の市場を最大にする方法：

ローズ 三浦氏／エッジ・トランスレーション

◎Memsources の翻訳データから探る、機械翻訳の最新動向：

石井 潤一氏／Memsources a.s.

◎産業翻訳における機械翻訳の継続的改善方法：

中山 雄貴氏／株式会社ヒューマンサイエンス

◎最適化された機械翻訳導入モデルによる医薬品開発業務の迅速化への試み：

鞠子 幸泰氏／中外製薬株式会社

早川 威士氏／株式会社アスカコーポレーション

◎「カスタム MT+フル PE」現状と見通し：医療翻訳文やの事例研究：

山田 優氏／立教大学

早川 威士氏／株式会社アスカコーポレーション

◎NICT アダプテーション+EBMT エンジンの実力とそれを用いた統合翻訳環境：

本間 奨氏／日本特許翻訳株式会社

それぞれの発表について、前述の発表資料より次にまとめます。

◎機械翻訳と人間翻訳の両方の市場を最大にする方法：ローズ 三浦氏

昨今の MT は精度が急速に向上しており、翻訳者の中にも危機感を抱いている人がいるが、MT と HT はマスの取り合いではない。発信のための翻訳で、訴求が目的の場合、MT より HT が適しており、誤訳の影響度が大きい翻訳では人手のチェックが欠かせない。一方、読解のための翻訳では、MT をチェックなしで使用できる。また、発信のための翻訳でも、内容が一般的であれば、最低限のチェックで充分である。MT ガイドラインで、MT を人の補佐として位置付け、規定することで、MT と HT が共存できる。

◎Memsource の翻訳データから探る、機械翻訳の最新動向：石井 潤一氏

Memsource では、6 か月ごとに Memsource のシステム上で行われた MT 後に PE を行った作業に対し、MT の品質指標として、MT 結果と PE 結果の編集距離に基づくスコアを算出している。2021 年 1 月から 6 月の調査結果では、一期前に比べ多くの言語ペアでスコアが上昇していた。MT エンジンとしては、Google、Microsoft、DeepL、Amazon が多く使われていた。言語ペアの人気度では、英語をソース言語とする翻訳が多数を占める中、英語への翻訳で日本語がトップであり、英日と日英がそれぞれ 2 位と 4 位だった。特定の言語ペアやドメインに対して、どのエンジンを選択するかで、出力品質が大きく変わっていたが、MT の出力品質は全般的に高いと言える。

◎産業翻訳における機械翻訳の継続的改善方法：中山 雄貴氏

産業翻訳における PE 作業者の主な作業として、翻訳メモリ (TM: Translation Memory) からの訳文流用作業、PE 作業の 2 つがある。TM からの訳文流用については、案件において MT 導入後も TM を流用する機会が多いことが分かっており、TM の管理が重要となる。TM のマスターには、作業完了後の訳文のみが

登録されるよう、案件の遂行中は翻訳やレビューの作業用の TM を分けて運用すべきである。さらに、求められる品質ごとにマスターを分けることも有用である。PE 作業においては、用語・スタイルの修正作業を自動化していくことで、PE の時間を 20%以上削減できる。PE 前後の訳文の差分から用語集やスタイル置換条件を更新し、PE 後の訳文を用いたモデル学習を行うことで、PE の作業負荷を減らせる。

◎最適化された機械翻訳導入モデルによる医薬品開発業務の迅速化への試み：鞠子 幸泰氏／早川 威士氏

医薬品開発における翻訳において、HT と比べて品質を落とさずに、MT と PE により翻訳スピードを向上する目的として、MT エンジンのカスタマイズと、PE の品質水準策定を行った。カスタム MT エンジンは、過去に翻訳した 24 万文を用いて作成した。カスタム MT エンジンは、基盤エンジンに比べ、自動評価や PE の編集率と編集時間、固有表現の一致度、といった観点で性能が改善していた。PE の品質水準策定では、2 回の実証的研究を通して、HT と比べ、MT と PE で時間短縮効果があることが確認できた。また、1 回目の PE 結果を踏まえてポストエディターとコミュニケーションを取ることで、2 回目の PE では 1 回目の PE より翻訳品質が改善できた。

◎「カスタム MT+フル PE」現状と見通し：医療翻訳文やの事例研究：山田 優氏／早川 威士氏

医学・医薬分野の翻訳では、高いレベルの専門性や要求品質が求められる。医学・医薬分野において、文書特化型と医薬汎用型の 2 種類のカスタム MT エンジンを作成し比較評価を行った。文書特化型は 10~20 万文対、医薬汎用型は 40 万文対を追加訓練に用いた。2 つのカスタム MT エンジンの出力をプロの翻訳者 3 名が PE した上で、MT の各種誤り、作業効率、満足度など多種の指標で MT の出力を評価した。BLEU や編集距離など多くの評価指標で文書特化型が高評価だったが、流暢さやスタイルは医薬汎用型の方が高評

価だった。さらに、プロの翻訳者がどちらの MT 出力を
選好するかについて、どの評価指標が寄与したかを
分析した。BLEU や編集距離の MT 選好度への寄与度
は低く、用語エラー数の少なさが MT 選好度に有意に
寄与していた。

◎NICT アダプテーション+EBMT エンジンの実力と それを用いた統合翻訳環境：本間 奨氏

ドメイン適応型 MT は汎用型 MT より正しい訳語が
出力されるが、ドメイン適応型 MT を誰でも簡単に構
築できるようにするために統合翻訳環境を開発した。
NICT から 2020 年 10 月にリリースされたアダプテー
ション+EBMT という方式を用いて、統合翻訳環境で
特許翻訳用にドメイン適応型 MT を作成したところ
1,000 文対程度でも高精度なカスタム MT エンジンが
作成できる。特許分野の複数ドメインで、汎用型の MT
エンジンに比べ、BLEU や RIBES で 20 以上高いス
コアが出ることを確認した。専門家が選んだ 6,750 文
対でアダプテーションした電子写真ドメイン特化型
MT エンジンの実験では、用語学習効果があることや、
BLEU が 80 を超える賢い TM として機能することを
確認できた。

6. おわりに

昨年に引き続き、オンライン開催となったにもかか
わらず、創立 30 周年を記念した盛況なイベントとな
りました。テーマごとにルームを分けた参加者同士の
オンライン交流会や、ゴールドサポーター企業様によ
るランチタイムウェビナーなど、新たな試みも行われ
ました。今後の年次大会にもご期待ください。

参考文献

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J.,
Jones, L., Gomez, A. N., et al. (2017). Attention is all you
need. *Advances in Neural Information Processing
Systems*, 5998-6008.

[2] Nagao, M. (1984). A framework of a mechanical
translation between Japanese and English by analogy
principle, *Proc. of the international NATO symposium on
Artificial and human intelligence* (pp. 173-180). Lyon,
France: Elsevier North-Holland, Inc.

[3] Ishiwatari, S., Hayashi, H., Yoshinaga, N., Neubig,
G., Sato, S., Toyoda, M., et al. (2019, jun). Learning to
Describe Unknown Phrases with Local and Global
Contexts, Minneapolis, Minnesota.

[4] Kiyono, S., Ito, T., Konno, R., Morishita, M., &
Suzuki, J. (2020, nov). Tohoku-AIP-NTT at WMT 2020
News Translation Task, *Online*.

[5] Hinami, R., Ishiwatari, S., Yasuda, K., & Matsui, Y.
(2021). Towards Fully Automated Manga Translation.
Proc. of the AAAI Conference on Artificial Intelligence.

[6] Yasuda, K., Yamada, I., Okazaki, N., Tanaka, H.,
Asaka, H., Anzai, T., et al. (2021). Field Experiments of
Real Time Foreign News Distribution Powered by MT.
*Virtual: Association for Machine Translation in the
Americas*.

[7] Yang Garland, E., & Gao, R. (2021, aug). A Rising
Tide Lifts All Boats? Quality Correlation between Human
Translation and Machine Assisted Translation, *Virtual*.

第5回自動翻訳シンポジウム 参加報告

東山 翔平

国立研究開発法人情報通信研究機構

1. はじめに

2022年3月11日に、第5回自動翻訳シンポジウム～2025年に向けたグローバルコミュニケーション技術～がオンラインで開催され、無料で配信された。招待者による講演とパネルディスカッションが行われ、並行して翻訳事業に携わる国内企業16社によるオンライン展示も行われた。566名が視聴し、オンライン展示サイトには280名が訪問し、盛況であった。本稿では、講演2件とパネリスト発表、パネルディスカッションの内容を報告する。

2. 基調講演：丸山宏氏「Software 2.0 とデジタルトランスフォーメーション」

株式会社 Preferred Networks／東京大学／花王株式会社の丸山宏氏による基調講演の内容を紹介する。

前半は、近年の深層学習技術の発展と相俟って、ソフトウェア開発で広がっている「Software 2.0」の考え方についての説明であった。これは、「つくるプログラミングから探すプログラミングへ」の変化、つまり、人がゼロからプログラムを書くことから、プログラム全体の空間から欲しいプログラムを探索することへの変化を意味する。

プログラムとは、「入力に対して出力を対応付ける関数」とみなすことができ、プログラムの仕様 S とはプログラムが満たすべき命題に相当する。機械学習モデルをプログラムとして見ると、入出力例の集合である訓練データ T に対し、誤差 L を最小化するパラメータ θ を探索する計算が「探す」プログラミングに相当し、仕様 S は誤差 L として与えることになる。特に、深層

ニューラルネットは、非常に多くのパラメータを持ち、任意の多次元非線形関数を近似することができる。仕様、つまり誤差を与える方法として、人手によるアノテーションを利用する教師あり学習の他、観測データの確率分布を推定する生成モデル、シミュレータを用いた深層強化学習などがある。

後半は、IT イノベーションの第2フェーズに相当するデジタルトランスフォーメーション (DX) についての説明であった。第1フェーズは、人が行っていたプロセスを置き換えて効率化する段階を指し、給与計算を始め多くの事例がある。第2フェーズは、IT で初めて可能になるビジネスを創発する段階である。たとえば、通常の小売業がよく売れる商品のみを扱うものであるのに対し、Amazon.com では、めったに売れない商品の合計売上が80%以上を占め、欲しいものが検索で見つかることから成立しているビジネスである。他にも、Uber や Airbnb によるマッチングビジネスも IT なしでは成立しないものである。

丸山氏は花王にて、実在の人物のバイタルデータや生活習慣に関するデータを基に、ありそうな人体情報を同時確率分布として学習する仮想人体生成モデルを構築する取り組みを行っている。API でのサービス提供を予定しており、一部の属性値を基に他の属性値を推定するなどの活用方法があり、ヘルスケアや医療などに応用することでイノベーションにつながると期待している。

DX では様々な不確実性がチャレンジとなるが、短いサイクルで変化への対応を繰り返すアジャイル開発の考え方により、不確実性に対処でき、DX が世の中を変えていくのではないかと丸山氏は結んだ。

3. 講演：隅田英一郎氏「自動翻訳の素材を蓄える翻訳バンク」

情報通信研究機構（NICT）の隅田英一郎氏による講演の内容を紹介する。

自動翻訳は研究開始から 75 年経ち、現在では使えるレベルになっている。NICT で公開している音声翻訳アプリ VoiceTra、テキスト翻訳サイト「みんなの自動翻訳@TexTra」を含め、様々な翻訳サービス、アプリ、専用機が提供されている。

2014 年以降、深層学習を用いた自動翻訳システムである Neural Machine Translation（NMT）により、長い間越えられなかった実用の壁を越えた。2020 年時点で、自動翻訳システムは TOEIC スコアで 900 点相当に達し、多くの人にとって役に立つレベルとなっている。さらに現在は、用例ベースの方法を実装した NMT が研究されており、隅田氏はこれを第 3 世代の NMT と呼んだ。

NMT は、入力と出力の対である対訳データを真似するシステムであり、質の高いデータをいかにして集めるかが非常に重要である。「翻訳バンク」の枠組みにより、信頼できる組織から寄付いただく形で高品質なデータを集め、分野ごとに高精度なシステムを作っている。

翻訳バンクの活動開始から 4 年以上経ち、これまでの「点」から「領域」というステージに至っている。1 つ目の事例はコミュニティとの連携で、Linux Foundation Japan 等との連携により、OSS ドキュメントの翻訳に活用されている。2 つ目の事例はビジネスから文化へと拡大している例で、日越茶道・文化交流会と連携している。3 つ目の事例は官民連携で、NICT と金融庁が連携して金融業界団体からまとめたデータを確保し、金融分野向け NMT を構築することで、日本語・英語間の翻訳において汎用 NMT に比べて精度向上を実現している。

4. パネリスト発表：丸山宏氏、川竹一氏、菅谷史昭氏

パネルディスカッションの前に、前述の丸山宏氏、ソースネクスト株式会社の川竹一氏、マインドワード株式会社の菅谷史昭氏の 3 名のパネリストによる発表が行われた。各発表の内容を簡潔に紹介する。

（1）丸山宏氏「本 1 冊機械翻訳で訳してみた」

丸山氏の著書「新 企業の研究者をめざす皆さんへ」を、機械翻訳を利用して英訳した経験について。同書は、過去に出版した本の内容と新たに書き下ろした文章からなり、これを機械翻訳して、丸山氏と英語母語話者による後編集を行い、英訳版として公開した。構造が明確な日本語の文章を書くことで、機械翻訳がうまくいき、そのままでもほとんど通じる英文となった。一方、辞書の併用による用語のコントロールや、イタリック体などのフォーマットの保持についてはうまくいかなかった。

（2）川竹一氏「POCKETALK」

ソースネクスト株式会社の子会社としてポケットトーク株式会社を設立し、「言葉の壁をなくす」をミッションとしている。日本はアジア圏からの観光需要が高いが、日本人の英語力（英語能力国際順位）は低下しており、テクノロジーでの解決を目指している。ポケットトーク事業は米国でも実績が伸びており、英語非母語話者により利用されている。これまでの通訳機に加え、ビデオ会議システムと連携する字幕表示ソフト、スマホアプリ、法人向けのポケットトーク管理用 Web アプリのサービスも提供している。

（3）菅谷史昭氏「利用者の選択肢が増える自動翻訳サービス」

自動翻訳の性能が向上し、サービス提供者の選択肢も増え、個人が読解に用いる「分かる翻訳」から、出版補助などの「デキル翻訳」として使われ始めている。分野に特化した専用エンジンは、製薬や特許文書など、人手をかけずに短時間で大量に翻訳したい場面で有益である。10～100 万文程度の追加訓練データを集め、

求める品質に応じてクレンジングも行い、アダプテーションを行うことで高精度な専用エンジンを構築できる。今後は、業界から企業、個人へと、小さい粒度での専用エンジンの実現もあり得るのではないかと。

5. パネルディスカッション「2025 年に向けた多言語翻訳の将来展望」

NICT の内山将夫氏をモデレータとして、前述のパネリスト3名とのパネルディスカッションが行われた。以下、中心的な内容を紹介する。

「自動翻訳を意識して書くと執筆速度はどうか」英訳を意識することで日本語の文章の論理構造が明確になり、むしろ書きやすくなった。(丸山氏)

「自動翻訳精度とフォーマット保存等の周辺機能ではどちらが足りていないか」数式や図表、たとえば図中の日本語テキストも翻訳してほしいなど、フォーマット関係での不足を感じる。(丸山氏)

「2025 年に『言葉の壁』がどの程度解消していそうか」100%を目標にすると数%程度かもしれない。自動翻訳の使用者が増え、他言語の人と話すといった形で人の行動が変わっていくと、データも増えていく循環が生まれる。(川竹氏)

「2025 年での企業・個人向け翻訳エンジンの実現可能性」個人やグループで十〜十数冊相当の似たトピックの対訳があると、遠くない将来にも実現性が出てくる。自分でカスタマイズした部分を別の翻訳エンジンへ移植しやすいと嬉しい。(菅谷氏)

「利用者視点での現状の翻訳エンジンへの評価・評判」Web テキストの翻訳などは多くの人が日常的に使っており、使えるものになっている。音声翻訳でも、意識して明瞭・明確に話すと実用レベルである。一方、その場の使い切りではなく記録に残すための文章の翻訳では状況が異なる。専門用語や、新製品の名前などの固有名は難しく、頻繁に使われない言葉を翻訳できるようになるにはまだ時間がかかる。確実さの点に懸念があり、数値や条文などの重要な箇所が正しく翻訳

されていることを保証できることに対する事業者からの要望がある。(各登壇者)

6. おわりに

最後に、報告者による所感を述べる。先進的な IT ビジネスの取り組みや産業界での翻訳事業の展開についての各講演・発表はいずれも興味深く、自動翻訳の普及・実用化が、個人や企業に加えて、業界団体・コミュニティへと広がっていることを確認できた。

利用者視点での自動翻訳の評判に関する議論で出た各種課題については、自然言語処理の研究コミュニティでよく挙がる課題と重なっている印象を受けた。流暢な言語生成（翻訳）ができ、概ね意味が伝わればよい状況での実用レベルに到達してきたことで、固有名、専門用語、数値表現といった、決まった形式を保持することが求められる言語表現を、どう正確に処理するかという観点の重要性が浮き彫りになっていると感じている。利用者・産業界からの課題や要望を研究コミュニティにフィードバックし、双方の現状認識や目標の共有をしていくことが今後もますます重要になると考えている。

編集後記

中澤 敏明

AAMT 編集委員会

本稿執筆時点で、個人が開発した日英の機械翻訳エンジンが非常に高精度であるということで、Twitterなどで話題になっています。このサービス自体がどうこうというよりも、個人でもこのレベルの翻訳エンジンが開発できる環境になっている、という事実をしつかりと認識しておくことが大切だと思います。AAMT ジャーナルがそのための情報源の一つとなれば幸いです。

巻頭言では日本翻訳連盟会長である安達久博様より「翻訳の日」という題でご寄稿いただきました。9月30日は国際翻訳デーとして2017年の国連総会にて公式に認定されており、国内でも2019年に翻訳の日として正式に認定されました。近年の機械翻訳技術の目覚ましい発展を好機と捉え、文化の異なる他国の人々に単なる「ことば」や「こと」の翻訳から、伝えるべき「ところ」を正しく自然に翻訳するように努めることが肝要であるとのお言葉をいただきました。

小田登志子様からは『日本の大学における教養英語教育と機械翻訳に関する予備的調査』に見る英語教員の現在と未来という題で、大学英語教員を対象に行われた教養英語教育と機械翻訳に関するアンケート調査の結果の解説をいただきました。アンケート結果から、英語教員の間で機械翻訳への関心は非常に高いものの、どのように機械翻訳と向き合うべきかという点については準備が不十分であり、外国語教育界全体で機械翻訳の存在を前提とした外国語教育の在り方を模索する必要があるとのご示唆をいただきました。

AAMTからは「MTに関する法的問題について」という題で、MT利用のさまざまなケースを想定して、どのような法的リスクが存在するかを解説しております。MTを正しく利用するためにも、ぜひご一読いた

だきたいと思います。

ローズ三浦様からは「機械翻訳と人間翻訳の両方の市場を最大にする方法」という題でご寄稿いただきました。翻訳の市場は「機械翻訳と人間翻訳のマスの取り合いではなく、マス数そのものを増やせる可能性がある」という仮説をたて、これを実現するためのさまざまなポイントをご提案いただきました。

Memsources社からは「Memsources 機械翻訳調査レポート 2022 年第2四半期」という題で、Memsources上での機械翻訳の利用動向の最新の調査結果をご報告いただきました。MTの品質は少しずつ向上しており、欧米系の言語間ではMT出力の7から9割、日英でも7割近くを活用できているとのことでした。一方、最良の結果を得るためには翻訳対象文書に応じて複数のMTエンジンを使い分ける必要があるとも述べられており、興味深い結果となっていました。

中山雄貴様からは「産業翻訳における機械翻訳の継続的改善手法」という題で、機械翻訳導入後にいかにして継続的に翻訳業務を改善したら良いかというテーマでのご提案をいただきました。ポイントとしては翻訳メモリを適切に蓄積していくための工夫、用語集やスタイル置換の適切な設定、蓄積した翻訳メモリを利用した機械翻訳モデルの更新の3点があるとのことでした。

鞠子幸泰様と早川威士様からは「最適化された機械翻訳導入モデルによる医薬品開発業務の迅速化への試み」という題で、正確性を含む品質の保証が欠かせない医薬品分野における機械翻訳の活用可能性についてご報告いただきました。翻訳対象文書ごとに適切にカスタマイズされたMTを利用することと、ポストエディットの品質基準を明確化することにより翻訳作業

Editor's note

Toshiaki Nakazawa

AAMT Editorial Board

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Public License.
License details: <https://creativecommons.org/licenses/by-sa/4.0/>

は効率化できるという結論が得られており、有用な知見が得られたと思います。

山田優様と早川威士様からも同様に『『カスタム MT +フル PE』現状と見通し：医療翻訳分野の事例研究』という題で、高品質が要求される医学・医薬分野の翻訳における詳細分野に特化した MT エンジンを活用した MT+ポストエディットの実用性についてご報告いただきました。カスタム MT の品質の差がフル PE の作業量に与える影響をはっきりとは確認できなかったとのことで、今後の展開にも期待したいと思います。

本間奨様からは「特許翻訳におけるドメイン適応型機械翻訳」という題でご寄稿いただきました。ドメインをかなり絞り、MT のアダプテーションに用いる訓練対訳文をしっかりとコントロールすれば、高精度なドメイン適応型 MT が実現できる可能性があると報告されています。

出内将夫様からは AAMT2021 Online のイベント報告をいただきました。AAMT2021 は AAMT 創立 30 周年を記念して二日間に渡り開催され、記念講演 3 件、長尾賞/学生奨励賞受賞者講演 3 件、年次大会発表 6 件、公募セッション発表 6 件と盛りだくさんの内容であり、非常に盛況でした。

東山翔平様からは第 5 回自動翻訳シンポジウムの参加報告をいただきました。2 件の講演とパネルディスカッション、企業展示が行われ、非常に盛況だったとのことです。機械翻訳においては流暢性が飛躍的に向上し、読みやすい文が生成できるようになった一方で、固有名詞・専門用語・数値など訳し方が決まっているものをいかに正確に訳すかという点が重要になってきていると述べられています。

冒頭でも述べましたが、AI 技術の進歩は非常に早く、最新の情報にアンテナを張っておかないとすぐに世界の流れから置いて行かれてしまう状況になっています。AAMT および AAMT ジャーナルからもさまざまな情報の発信を通して、機械翻訳の普及と発展に貢献していきたいと思います。

AAMTジャーナル「機械翻訳」No. 76

- 【発行日】2022年6月15日
【発行】アジア太平洋機械翻訳協会 (AAMT)
ホームページ: <https://aamt.info/>
【住所】〒619-0289
京都府相楽郡精華町光台3-5
国立研究開発法人情報通信研究機構 先進的翻訳技術研究室内
【編集委員会】内山将夫 後藤功雄 中澤敏明 新田順也 園尾聡 森口功造 隅田英一郎
仲山裕子 山田優 石川弘美 早川威士 出内将夫
【表紙デザイン】泉谷東十郎
【題字】長尾真
【事務局】石川弘美
【印刷所】株式会社 プリントパック

Asia-Pacific Association for Machine Translation (AAMT)
c/o National Institute of Information and Communications Technology
3-5 Hikaridai Seika-cho Soraku-gun Kyoto 619-0289, Japan

