

AAMT

Asia-Pacific Association for Machine Translation

Journal



March 2014

No.55

アジア太平洋機械翻訳協会

目 次

巻頭言：	言語文化を守る機械翻訳	熊野 明	1
連載：	機械翻訳事始め[II]	坂本義行	2
レポート：	Japio 世界特許情報検索サービス (Japio-GPG) の紹介	長谷川雅昭, 大塩只明	9
シンポジウム報告：	第 23 回 JTF 翻訳祭セッション開催報告	機械翻訳課題調査委員会	14
製品・サービス紹介：	翻訳キーボードの紹介	株式会社 高電社	17
製品・サービス紹介：	Jソウル9	株式会社 高電社	19
委員会活動報告：	これまでの AAMT Forum メールマガジン	機械翻訳課題調査委員会 WG1,2	22
委員会活動報告：	実務翻訳における機械翻訳の利用に関するアンケート結果報告	機械翻訳課題調査委員会 WG2	29
委員会活動報告：	シンプルな用語集形式 UTX とその活用	機械翻訳課題調査委員会 WG3	33
AAMT 会員のひろば：		Alun Simpson	39
事務局からのお知らせ：	協会活動報告 (2013 年 9 月～2014 年 3 月)	AAMT 事務局	41
編集後記			48

CONTENTS

Foreword:	Machine Translation for Preserving Language Culture	A. Kumano	1
Series:	History of Machine Translation	Y. Sakamoto	2
9Report	Overview of Japio Global Patent Gateway service (Japio-GPG)	M. Hasegawa, T. Oshio	9
Symposium Report	Session at 23rd JTF Translation Festival Tokyo 2013	MT	14
Products & Services	Honyaku Keyboard	Kodensha	17
Products & Services	J Seoul 9	Kodensha	19
Committee Report	AAMT Forum mail magazine	WG1, WG2	22
Committee Report	Report on questionnaire results of using machine translation for business translations	WG2	29
Report	UTX, a Simple Glossary Format, and Its Applications	WG3	33
AAMT Members:		Alun Simpson	39
AAMT Activities:	AAMT Activities (from September 2013 to March 2014)		41
Editor's Note:		T. Utsuro	48

言語文化を守る機械翻訳

東芝ソリューション株式会社

プラットフォームソリューション事業部 熊野 明

多くの人で混雑する週末のショッピングモール。すれ違う人の声に耳を傾けると、日本語と異なる言葉が聞こえる。中国語や韓国語の会話とわかる。

我が国の観光庁は、2020年の東京オリンピック、パラリンピックの開催に向け、外国人に対する対策を始めた。交通インフラや多くの人が利用する施設の案内に、日本語以外の表記を併記するよう指針で求めている。まず必要なのは外国人にわかる英語である。“Kokkai- gijidou”のような単純なローマ字表記は、外国人にとって意味がない。さらに必要に応じて、中国語、韓国語も加える。当然の動きであるが、もっと早く始めるべきだったと感じている人も多いだろう。

もちろんその先には、単なる案内表示だけでなく、日本人と外国人のコミュニケーションのツールも求められる。機械翻訳の技術に期待がかかる。

一方、地球上には数千種類の言語があるが、急速にその数を減らしていると聞く。2週間に一つの言語が失われているらしい。想像し難い速さである。それらの多くは、少数民族で使われている言語であり、保存はおろか記録も難しい。そもそも言語の保存には、それを使いこなす話し手が必要だ。話し手がなくなった時点で、その言語は消滅してしまう。

どうして言語が失われてしまうのか。話し手がいなくなるのか。異なるコミュニティから新たな情報や技術を取り入れるためには、それを伝える手段が必要である。情報共有のために利用可能な言語を選択する。この伝達行為を地球レベルで進めると、その手段は英語を中心とした言語群に収束する。今日急速に進むグローバル化と大量の情報の中で、英語を中心とした情報共有は避けられない。

同様の状況は、日本国内でも見られる。日本では方言を話す人の数が減ってきたと聞く。標準語を共通語として使用するなかで、方言の優位性が少なくなっているからである。しかし、本当にそうだろうか。少数言語や方言でしか表せない文化もあるはずである。

機械翻訳技術は、入力言語と出力言語の関係を分析し、それを利用して人間の翻訳活動を再現する。翻訳の手法によりその形に違いはあるが、入力言語と出力言語に関する知識を言語資源として備えている。この言語資源によって、言語そのものを保存することができると考える。

Google が公開している翻訳サービスは次第に言語の種類を増やし、2014年3月現在で80種類の言語を扱える。このように蓄積した言語資源は、各言語の特徴を表現したものであり、言語を再現するための貴重なデータである。これはそのまま、言語を保存する一手段となりうる。

さらに一歩進み、文化としての言語、言語文化を守るためには、深い技術が必要だろう。少数民族の言語や日本各地の方言を、単なる標本として保存することは、言語文化を守ることにはならない。具体的に表現されている言語現象を分析して、その奥に存在する文化の源を見つける必要がある。

これまでに保有し、さらにこれから蓄積する大規模な言語資源を活用することにより、言語文化そのものを守ることができないだろうか。少数言語や方言の場合は、データの収集が容易ではない。しかし、十分な言語資源を収集し、最新技術で分析することができれば、言語文化の源として見えてくるだろう。言語消滅の速さとの競争である。

機械翻訳事始め[Ⅱ]

アジア太平洋機械翻訳協会理事 坂本義行

1. 1960年代前半の活動

情報処理学会ができ、その学会誌「情報処理」が発刊され、1960年7月号、学会誌第1号に「計算しない計算機」という題で、工業技術院電気試験所の和田弘氏が論文を書かれている。¹⁾

1. 文字の認識

その内容は、まず最初に文字を読む機械について、電子計算機が符号(code)だけしか読めないで、人間に読める文字が読めない現状は不便です。そこで機械に人間の文字を読ませようとする努力は、いくつかの方法が発表されている。

電気試験所では、800個のゲルマニウムダイオードを用いて、普通の英文タイプにある文字や記号の合計73文字が読める装置を開発した。このような研究は、イギリスのManchester大学、アメリカのM.I.T.やBell研究所でも始められており、音声認識についても京都大学、東北大学、名古屋大学、電気試験所などで研究されており、これらを対象とする研究をpattern認識(recognition)と呼んでいると述べられている。

2. 外国語の翻訳

外国語を機械で翻訳させる研究は逐語訳から始まりました。欧米語はソ連を含めて、同じ系統から発しているもので、一つの文章の中で単語の相互の位置には、あまり大きな違いがありません。

われわれは、semanticな研究よりも、まず構文的(syntactic)な研究が必要だと判断した。

英日翻訳を考え、文章を単純なものだけとした。訳文は仮名文字で、分ち書きして打たせた。この翻訳結果は、前に紹介している。そして、実用化のためには、vocabularyを増すことが必要ですが、研究としては欧米でやっているようなsemanticなものへ移行すべき段階に到達したと思っていますが、それにもmachine

learningの考えを必要とする。さらに、九州大学、教育大学でも研究が始められていると述べられている。

3. 要約機

論文を要約(abstract)する必要があります。できることなら、よい論文は機械を通して一語も減らないが、まずい論文は何も残らないようにしたいものです。

採用した方法は、英文の単語を意味のある単語と意味のない単語に分けて、意味のあるのは名詞だけとし、その頻度を計算し、ある数以上の頻度の単語を選び出す。論文を読み込み、単語がある数以上含まれている文だけを印字させる。

例えば、「猫の首に鈴をつける鼠の相談」という寓話をこの方法で、要約させると、

The cat killed some of the mice,

Put a bell on the cat.

But who is going to put the bell on the cat?

の3文が印字される。注目した単語がcat, mouse, bellの3語であって、上の文ではそれらの語が2個以上含まれていたためです。1,000語程度の文章であれば、「ヤマト」を使うと、約17分で要約されます。

今後はsyntacticな考え、単語のつながり、三段論法などを考慮に入れて改善するつもりですが、ここにもmachine learningの問題が潜んでいる。

4. 検索機

論文を読んで、今後の参考にしたいと感じると、我々は著者名、雑誌名、その番号ページ、発行年月、論文の標題、および内容の要約とか備考などを記録、保存しています。これを機械化して、能率を高めることを企てました。これらの事項をテープに打って機械にいれると、磁気テープに記録、保存されます。必要な時に、それらのどれかの項をこの機械にいれると、該当する論文全部について、上記の事項を自動的に打ち出してくれる検索

機(indexer)を最近完成しました。

将来、何とかして、同じものは一か所にまとめ、古いものは忘れて(整理して)、またひとつのことがいろいろの角度から取り出せるように、例えば、cross-reference というような概念を導入して、要するに合理的な整理方法を案出したいと念じております。

5. むすび

現在のところ計算機が利用されているのは、演繹的(deductive)に書かれた数値計算に限られています。定理を証明するような帰納的(inductive)なことはまだできません。

また、先ほどから述べてきたところの諸例は、いずれも多量のデータを整理して共通な本質ともいうべきものを求めたうえ、判断する性質のこと、すなわち帰納的なことでして、人間の学習課程に似た内容のことです。machine learning と呼ばれる所以です。

本日誕生した情報処理学会が壮年になるころのことを想うと、意見なり思想を表す文章も、この計算機にかかっているでしょう。司法から始まって、行政から政治までも、やがてこれらの機械によって処理されることも夢とばかり片づけられません。

それが実現されるためにはまづ述語論理について、帰納的な理論が確立されることが望まれます。そしてその際にも、言葉(単語)は、element として残りましょう。単語の意味というものは数値と違って先に述べたように曖昧な点がありますが、数学、ことに代数学はわれわれの知っているような単純な数や文字を対象とすることから離れて、複雑な集合のような抽象的な対象を操作する方向に進んでいますから、単語や文章の解明に大いに参考になることと思われます。

ことに、日本語を科学的に考察し改善することは必要で、これを怠ると機械文明の進歩にもかかわらず、日本人の生活には向上しない面が残されることになりはしないかと案ぜられます。

半世紀以上も前、トランジスターでやっと計算機ができ、プログラミング言語としては、機械語しかなかつ

た時代に、言語処理の未来を予測した論文を投稿された先輩、和田氏の先見性に(私、坂本は)驚かされました。

2. 1962 年

前掲の2年後、「情報処理」²⁾のVol.3, No.3, 1962年5月号に和田氏が機械翻訳の現状を解説されています。

1. 現状

まず、前号で私が報告したような1960年以前の世界での研究・開発の状況報告を行い。

問題の所在として、文中の単語の位置がヨーロッパ言語間では、殆ど同一なのに比べ、日本語では、大きく異なると述べ、

機械翻訳においては両者は総合的に研究されねばならない。研究は

1. 外国語を整理する
2. 整理された内容を機械に教える
3. 訳文を合成する

の3段階に分けられる。各国の研究状況を見ると、逐語訳から一歩進んで、構文の分析に関心が持たれる時代になっている。自国語を合成することにはまだあまり力が注がれていないようだ。

2. 言語を計算機へ

2. 1. 処理方法

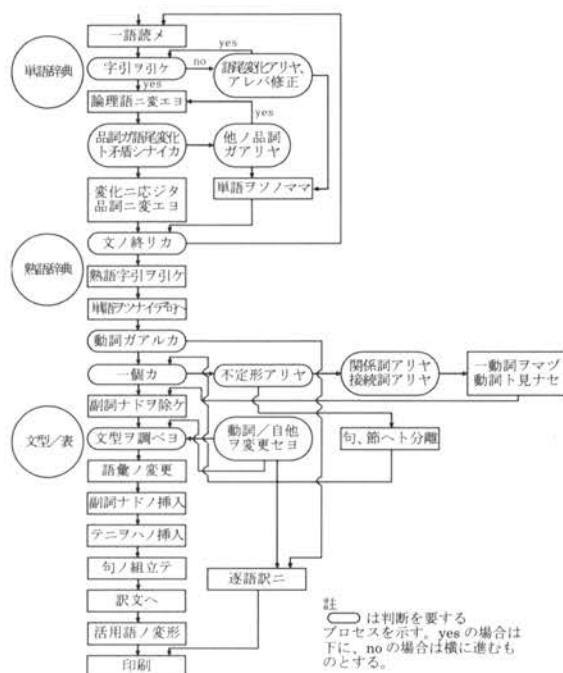
文字の表示法として、アルファニュメリックの2進法表示と引き算によるその照合方法について述べ。つぎに単語辞書について、派生語を含めて見出しをどうするか。さらに、その配列方法。使用頻度順に配列するのも一つの方法であろうと述べられている。

語尾処理には、サブルーチンという手法を用いると機械に便利だとも述べられている。

文型の決定では、単文の処理で、単文は主語、述語を含んでいて、その述語の性質で目的語あるいは補語を必要とする。このことは四則演算が $a \times b$ のように必ず2個の数値を操作するのと違う点である。文型を機械的に確認させる方法として、table を用いることによって、容易に照合がとれる。ただし、多品詞語には苦しめられるとも述べられている。

2. 2. プログラム

処理手順を具体的に表示するために、流れ図をつくる必要がある。これを第1図に示す。さらに、語尾変化を処理するブロックを subroutine を使って行う。最後に、このような詳細図を機械の持つ「命令」の系列を使って表示しなければならない。命令を機械語ともいう。



第1図

われわれが翻訳（電気試験所の「ヤマト」）に当たって使いたい命令形の単語は機械に備わっている命令には備わっていないし、これらでは表現できないように感じられる。これが機械翻訳の一つの隘路になっているようだ。しかし、これを機械でやらせようとするからには、命令で表せるような工夫が必要だ。

その要領を述べよう。” $a \times b$ ”なる演算は、

$$\begin{array}{r} 23 \\ \times 14 \\ \hline 92 \\ 23 \\ \hline 322 \end{array}$$

が示すとおり、九九、左への桁送り、加算、という操作を積み重ねた結果である。このように単純な操作を集合

として記述することを算法(algorithm)という。従って命令“乗算”がなくても、上記 algorithm ができれば、機械に乗算を行わせられる。普通は“乗算”と書けば、操作する電気回路が機内に組み込まれている。このような操作はプログラム software でも実行できるが、「ヤマト」では hardware で組み込んでいる。

翻訳でプログラムを作りたいときに、欲しい機械語がなくても前述の「字引を引け」のような疑似命令(pseudo code)を提供することは、プログラミングにとって大切な課題である。

3. 現実の文章

一文中の各単語の品詞の決定では、1品詞のもの、多品詞のもの、それに単語が辞書にない場合と品詞の決定には多くの困難がある。

辞書づくりでは、単語の見出し語をどのように登録するかに問題があるとし、Washington 大学の Relfler は、全てを登録する形で、170,000 語を登録していると、California 大学の Lamb は、単語を分解し、lexeme（語幹、接頭辞、接尾辞、語尾変化）に分類し、15,000 を用意し、普通単語はこれらの組合せで作られて約 1,141 万語までできるはずだし、同大学が化学を分担している関係から、化学用語は、2,400 万語まで見出せる筈だという。

見出し語(entry)がどうであつたらよいかは、多くの問題があるから、形態論(morphology)の議論が続くであろう。

文章は、文からなり、文は句(phrase)、節(clause)という概念で処理している。句には、形式上動詞を含まないもののほかに、不定詞、動名詞、分詞構文などがある。それらは一面、文中での働きに応じて、名詞句、形容詞句などと分類される。句は単語が連結されたもので、このことは句が一つの結合語とでもいえる因子(string, package, fused unit などとよばれる)となって、他の単語と同格な立場で単文を構成していると思ふ。

節も同じで、coordinate clause は接続詞でつながれているから、重文と同じであり、単文を f と表わせば、 $f_1 + f_2$ の形になる。

これに対して subordinate clause は日本人を苦しめやすいが、その clause は形式的には単文の形を備えており、句読点、接続詞、関係詞がその始まる位置を示している。そこでこの節を見いだせれば、節にも働きによって名詞節、形容詞節、副詞節などになるから、この新因子は他の単語とともに、全体としての単文の規則に従うことになる。

構文分析の例として、Georgetown 大学では、第 2 図のような手順でおこなっており、ソ連では、第 3 図のような手順で行われており、Indiana 大学では、多品詞語の品詞の決定することに重点をおいて、動詞の関係している品詞の確定できないものを第 4 図のような 6 種に絞って、その Subroutine をつくっている。さらに、Pennsylvania 大学では、文章に含まれている無数の表現を整理して、それらの翻訳形式を用意すれば、辞典とおなじようにそれを look up させることで翻訳の一解法になると述べられている。

1. 字引を引く
2. 形態の分析
 - a. 熟語の調査
 - b. Syntagmatic の分析
 - c. 挿入句の除外
 - d. 分離
3. 構文分析

第 2 図

1. 動詞の分析
2. 句読点の分析
3. 構文の分析
4. 名詞の分析
5. 形容詞の分析
6. 語順の変更

第 3 図

最後に和田の試案というものが紹介されている。

文章を構成する element は単文であると見做す。

関係詞、if, then, as, because…などがこれに含まれる。Comma の如きもこれに含める。

- | | |
|-----------------|-------------------|
| 1. n/v (現在) | point, face |
| 2. a/v (現在) | clean, complete |
| 3. n/v (過去分詞) | cut, thought |
| 4. a/v (過去分詞) | given, interested |
| 5. a/n/v (現在分詞) | meaning, running |
| 6. 特長のあるもの | mine, like, well |

第 4 図

次に修飾句などを棚上げすること、単文の場合とおなじ、このようにして簡素化された文に対して、別に用意してある基本文型を、

1. 文中に接続語を見出して、これを文型とあわせるか。
2. 文中に主格を見出して、これを文型と合わせるか。
3. 文中に動詞を見出して、これを文型と合わせるか、などで照合させてみて、文型の残余の部分と文章のこれに対応するところが、文型の全領域で合致するかどうか調べる。
この調査には品詞の決まっていない語が文型の要求する品詞をふくむかどうかが含まれているであろう。

以上の操作を繰り返すことによって、文章全体の構文を決定できれば、品詞も決まることになる。なお、句については、動詞を含む句の基本形を用意する必要がある。

4. 関連ある研究

4. 1. 中間言語

機械翻訳は当初英語から日本語へというように、特定の国語間で進められてきた。しかし、どんな国語からでも翻訳できるように拡張したくなる日もあろう。その場合に、共通点も多いのだから、それらをまとめて記述できれば労が少なく済むはずである。あるいは各国が自国語を分析した資料を作って相手国と交換すれば、直ちに MT ができることが望ましい。

MT はさきに記述したとおり、辞典の問題、形態の問題、および構文の解析によって基本文型(pivot language と呼ぶ人もある)まですれば、あとは自国語の合成である。これらの問題をどんな形で記述するかに当たって、参照すべき一つの基準があつて、それが

algorithm に便利な姿であると便利である。Esperanto があるように、このような基準となる言語を、中間言語 (intermediate language) と呼ぶ。

これをどんなものにすればよいかはまだ多くの研究に俟たねばなるまいが、次に述べる述語論理がその方向を示唆してくれれば有難い。

4. 2. 述語論理

命題論理は、文と文との結合方式を明らかにしてくれたが、述語論理は一つの文の構造を研究して、量化した概念を取り扱えるようになってきた。また文が主語と述語だけの他に、いくつかの目的語を含んだ単文についてそれらの言葉の間の関係を追及してくれそうな傾向にある。

まだ、未完成であるが、文章が数式化できる暁には、その基準となる素質を備えていると信ぜられる。

また、Basic English にもふれ、L. A. Richards が 1943 年に提案した方法では、18 の述語で表現できるといわれ、これらを記号で表すとすれば、述語論理の手法を使えば、文章は名詞、修飾語と記号とで表示することもできるのではあるまいか、とも述べている。

4. 3. 自動プログラミング

電子計算機の科学計算、事務計算に必要な program は、初期は machine language を使って克明に作られていたが、記号 program を経て、いまや Algol, Cobol などによる自動 programming の時代に入ろうとしている。各メーカーはこのための compiler をつくる競争をしているといってもよからう。従って機械翻訳の研究と compiler の研究は近い。

このようにプログラムを楽に各思想が普及してくると、機械翻訳の分野でもこれに似たものが開発されるようになることは当然であろう。例えば、MIT の COMIT は単に翻訳だけでなく、数字でない文字をいろいろの目的で計算機で取扱う際に便利のように完成されたもので、要求があれば配布できるようになっている。

私 (坂本) も電気試験所で、最初に学んだプログラミング言語は、COMIT であった。

最後に、あとがきとして、我が国でも、MT が話題

になるようになって、早くも 5 年になる。この方面に関心を持つ人が次第に増えてきていることは喜ばしい。これらの人はいずれも MT に寄与する専門知識を持っている。言語学者、論理学者、技術者にしても、同じ分野の人々だけで固まったのではグループとはなっても MT を推進させる力になりにくい。

MT をなるべく速く実用化に近づけるようにするには、それらの専門家の方の研究の方向に向く必要がある。そして事実その意志もあるように見受けられる。

そこで具体的な方策となるが、それには専門を異にする人々が自己の立場で話することから一步前進して、共通の立場としてプログラムを中心とした話へと移ることは良いと思われる。もちろん言語学、機械、理論などに分類していることは差支えないが

と述べられている。

3. 日米機械翻訳セミナー

3.1. 第 1 回の日米機械翻訳セミナーのニュース

第 1 回の日米機械翻訳セミナーの内容が「情報処理」の 1964 年 5 月号のニュース欄に以下のように紹介されている。

機械翻訳セミナー (日米科学協力委員会) 東京で開かる
1964 年 4 月 20 日から日米科学協力委員会第 2 部会 (科学技術情報資料交換会) の主催による機械翻訳セミナーが都道府県会館 (東京) で開かれた。日本側議長丹羽保次郎、米国側議長 Franz L. ALT をはじめ日本側 26 名、米国側 11 人の参加者があった。

論文発表が 3 日半、将来の協力計画の討論が 1 日、その他見学などがあった。発表者は日本側 13 人、米国側 8 人で、両国における機械翻訳研究の現状調査の報告と、個別的な研究の報告とがあった。米国側の報告がメタ言語理論の観点からなされたものが多かったのに比し、日本側の報告は自然言語 (とくに日本語) を取り上げたものが多かった。その他問題向言語、文字読取機、特殊計算機などについても触れられた。

おわりに 8 項目の勧告が採択された。これにもとづき、当学会またはその機械翻訳研究会が第 2 部会とのあ

いだけでなんらかの了解に達するものと期待されている。勧告には高速度計算機の容易な利用、機械翻訳関係の日本語論文に英文の要約をつけること、カードやテープの形の両国語文章資料の交換、プログラムの交換などがあげられている。

日本側発表論文の予稿集は希望者に配布される（取扱者・西村・電試）。米国側発表論文および資料はおもな研究機関に配布保管されているはずである。

3.2. 論文紹介

「情報処理」の1964年の9月号と1965年の1月号にその内容の概要が日本語で紹介されている。

最初に、丹羽保次郎氏の開会の挨拶が以下のように述べられている。

丹羽保次郎

Welcome Address

科学協力に関する日米委員会は1961年の池田総理とケネディー大統領との会談の結果生まれたもので、その第一回会合において5項目を選び、両国政府に勧告した。その中に科学情報と資料との交換の促進があるが、このために両国にそれぞれのパネルが置かれ、密接に連絡して仕事をすすめている。米国側パネルの主査は、最初はIBM社のDr. E. R. Pioreであったが、今はNBSのDr. A. V. Astinである。日本側は丹羽がつとめている。科学情報と資料のパネルでは、論文の抄録、目録索引に関する専門家の会議や一次出版物編集者の会合を企画した。この度の機械翻訳に関するセミナーは日米両国における機械翻訳の研究の現状を検討し、今後の協力につき協議する目的で、第2回の日米科学委員会で提案せられた。

日本における機械翻訳の研究は、今まで物理学者、数学者、電子学者、言語学者、心理学者などの数グループにより、大学や、研究所などでおこなわれていた。これらの研究の連絡、討議のため、1962年に機械翻訳研究委員会が発足し、日本における機械翻訳の研究者を網羅する体制ができ、今までに10回のシンポジウムを催した。今回の会合についても、この委員会に追うとこ

ろが少なくない。

日本の国語、国学、国文は欧米のそれに比し、非常に異なっている。この事実は論文の和欧または欧和の機械翻訳の研究をいっそうむづかしいものにするが、一方それだけ機械翻訳研究の国際協力の必要性を強くするものといえよう。今回の会合が日米両国における機械翻訳の研究と両国の協力の推進に寄与することを望んでやまないものである。

以下、論文のタイトルは、

1. Franz L. Alt: A Survey of Automatic Translation in the United States
2. D.G. Hays: Summary of U.S. contributions
3. H.H. Josselson: Linguistic Basis of mechanical Translation. Contributions of Standard Linguistic Theory
4. W. S. Wang: The Linguistic Basis of Translation: a Transformational Approach
5. S. M. Lamb: Stratificational Linguistics as a Basis for Mechanical Translation
6. E. D. Pendergraft : Hardware for Mechanical Translation: Relation between Hardware and Linguistics
7. Arnold C. Satterthwait : Hardware for Mechanical Translation : Problem-Oriented Programing Language
8. E. N. Adams: Input Reading for Language Processing system
9. Osamu Watanabe : Verb Patterns for the Basic of Sentence Generation
10. Akihiro Nozaki : Mechanical Translation of Pre-edited Japanese
11. Tuneo Tamati : A Survey of MT Oriented Hardware in Japan
12. Toshiyuki Sakai: Models and Strategies for MT
13. Bunpei Koro : English-Japanese Translation Formulas

14. Hiroshi Wada : Status and Problems of Mechanical Translation in Japan
15. Makoto Nagao : An Approach to the General Theory of Natural Language
16. Seiichiro Ohnishi : Psych-Linguistic Studies of Mechanical Translation in Japan
17. Toshihiko Kurihara : Design of the Special Purpose Computer KT-2
18. Hirohiko Nishimura : Automatic Segmentation of Japanese Text
19. Isao Imai : Japanese Grammar of the Zeroth Approximation
20. Shokichi Iyanaga : Procedure for pre-editing Japanese

である。

さらに、第 2 回 日米機械翻訳セミナーが米国で行われ、その報告と座談会が「情報処理」に掲載されている。それは次号に。

参考文献

- 1) 和田 弘 : 「計算しない計算機」、p11~15, 「情報処理」、Vol.1, No.1, 1960.7.
- 2) 和田 弘 : 「機械翻訳」、p144~154, 「情報処理」、Vol.3, No.3, 1962.5.
- 3) 「ニュース」、p178, 「情報処理」、Vol.5, No.3, 1964.5.
- 4) 「機械翻訳セミナー紹介」、p299~302, 「情報処理」、Vol.5, No.5, 1964.9.
- 5) 「機械翻訳セミナー紹介」、p47~55, 「情報処理」、Vol.6, No.1, 1965.1.

Japio 世界特許情報検索サービス (Japio-GPG) の紹介

一般財団法人 日本特許情報機構

長谷川雅昭、大塩只明

1 はじめに

一般財団法人日本特許情報機構（以下「Japio」という）では、2012年7月から、主要国・機関の特許文献を日本語と英語の両方で横断的に検索できる「Japio 世界特許情報検索サービス」（以下「Japio-GPG」という）を提供しています。

今回、AAMT ジャーナルで、Japio-GPG のご紹介の機会を頂きましたので、その特長、概要、使用している機械翻訳技術、今後の予定について、ご説明します。



[Japio-GPG]

2 Japio-GPG の特長

（特長1）世界の特許情報を日本語でワンクリック横断検索

Japio-GPG の最大の特長は、主要国・機関の特許文献（日本、米国、欧州、中国）と、国際特許出願（以下「PCT」という）をはじめとする世界の特許情報を日本語（機械翻訳含む）により、一括で横断検索できることです。具体的には、発明の名称、要約文、第1請求項（第1クレーム）を英語^{*1}、中国語^{*1}に加え、日本語（機械翻訳を含む）で検索できます。

※1 欧州、PCTは英語文献のみ、中国語は中国文献のみ対応。



[Japio-GPG の特長]

（特長2）中国特許公開・登録・実用新案も日本語で検索

Japio は、独立行政法人情報通信研究機構（以下「NICT」という）との中日機械翻訳技術の共同研究の結果、中国特許文献の高精度な機械翻訳技術を2012年秋に確立することができました。^{注1}そして、その半年後の2013年春には、この高精度な機械翻訳技術を活用し、中国特許・実用新案の発明の名称、要約、第1請求項をあらかじめ日本語に機械翻訳することで、Japio-GPG サービスの機能アップを実現しました。^{注2}

注1)

http://www.japio.or.jp/press/files/20121105pre_rel.pdf

注2)

http://www.japio.or.jp/press/files/20130328-2pre_rel.pdf

（特長3）日本語ダイレクト検索・表示

あらかじめ機械翻訳された日本語を直接検索することにより、界の特許情報を日本の特許公報を検索するのと同じ感覚で簡単に検索できます。

また、機械翻訳された日本語を検索回答表示にも使用していますので、迅速に閲覧できます。

(特長4) 膨大な語数の分野別機械翻訳辞書を使用
Japio独自のノウハウにより厳選した膨大な語数の分野別機械翻訳辞書を使用する事で、高精度な機械翻訳を実現しています。

(特長5) 和文抄録・PAJの使用
機械翻訳された日本語に加えて、和文抄録(中国公開特許和文抄録、米国公開特許和文抄録、米国特許和文抄録、欧州公開特許和文抄録)も蓄積しています。これにより、米国・欧州・中国特許庁の特許文献においては、より充実した検索・閲覧を行うことができます。また、日本公開特許の英文抄録は、PAJを使用しています。

3 Japio-GPGの概要

(1) 蓄積データ

蓄積データは、主要国・機関の特許文献、独立行政法人工業所有権情報・研修館(以下「INPIT」という)により作成された和文抄録、欧州特許庁(EPO)が提供している書誌情報等を含むデータベース(以下「DOCDB」という)等を使用しています。以下に概要を説明します。

・日本、米国、中国、欧州※2の特許文献、PCT※2については、発明の名称、要約文、全請求項※3を機械翻訳した日本語及び英語※4を蓄積しています。

※2 欧州、PCTは英語文献のみ対応。

※3 中国文献は第1請求項のみ対応。

※4 機械翻訳の英語は日本実用新案の発明の名称、要約文のみ。

・米国、欧州、中国の特許情報については、日本国特許庁、INPITにより作成された和文抄録(米国特許明細書と文抄録、米国公開特許明細書と文抄録、欧州公開特許明細書と文抄録、中国公開特許和

文抄録)も蓄積しています。

データの提供範囲と使用データ		
収録国	蓄積範囲 (公開発行年)	使用データ
日本 (JP)	1976～	DOCDB、代表図面、PAJ、日本公開特許公報データ、公表公報データ、日本特許公報データ、日本実用新案公報データ、機械翻訳データ
アメリカ(US)	1976～	DOCDB、代表図面、米国公開特許公報データ、米国登録特許公報データ、米国公開特許和文抄録データ、米国特許和文抄録テキストデータ、機械翻訳データ
欧州 (EP)	1978～	DOCDB、代表図面、欧州特許公開出願公報データ、欧州登録特許公報データ、欧州公開特許和文抄録テキストデータ、機械翻訳データ
中国 (CN)	1985～	DOCDB、代表図面、中国公報データ、中国公開特許和文抄録テキストデータ、機械翻訳データ
韓国 (KR)	1978～	DOCDB、代表図面、機械翻訳データ
PCT (WO)	1978～	DOCDB、代表図面、PCT国際出願データ、日本公表・再公表公報データ、機械翻訳データ
その他 (BR, RU, IN等) 90カ国(地域)、4機関	1976～ (注:国により異なります。)	DOCDB

[データの提供範囲]

・主要国を含む100の国・地域と機関の特許情報が収録されているDOCDBの発行国、文献種別(公報種別)、出願日、出願番号、文献番号、文献発行日、発明者、出願人、特許分類(IPC、CPC、ECLA)の書誌情報、発明の名称、要約を蓄積しています。また、発明の名称、要約文については、英語の場合は機械翻訳により日本語を作成し、蓄積しています。

[検索対象国設定画面]

(2) 検索

検索項目

検索項目は、書誌情報(文献発行国、公報種別、出願番号、公報番号、優先権主張番号、出願日、

公報発行日、出願人、発明者)、分類 (IPC、CPC、ECLA)、発明の名称 (英語、日本語・MT 日本語、中国語※⁵)、要約文 (英語、日本語・MT 日本語、中国語※⁵)、第 1 請求項 (英語、日本語・MT 日本語、中国語※⁵) です。

※ 5 中国語は中国文献のみ対応。

検索方法

検索方法は、予め検索画面に検索項目が窓表示された「項目検索」と、質問式 (クエリ) を入力する「詳細検索」の 2 種類があります。

項目検索では、項目間は AND 条件 (分類項目のみ OR 検索) で、同一項目では AND、OR、NOT 検索を指定します。また、括弧も使用可能です。

詳細検索では、項目名、AND、OR、NOT、括弧を用いて質問式 (クエリ) を作成・入力します。

[検索画面]

検索回答表示

検索回答表示は、「一覧表示」、「詳細表示」、「公報表示」の 3 種類があります。

ア. 一覧表示は、一画面 10 件で発明の名称、出願人、発明者、出願番号、出願日、文献番号・発行日、分類 (IPC、CPC)、優先権主張番号、パテントファミリー、代表図面を一覧形式で表示します。

イ. 詳細表示は、1 画面 1 件で全ての検索項目に加え、全請求項※⁶、パテントファミリー、代表図面を表示します。

ウ. 公報表示は、特許公報のフロントページまたは全ページを PDF 形式で表示します。

※ 6 中国文献は第 1 請求項のみ対応。



[検索時の操作のながれ]

検索・表示支援機能

ア. 用語検索支援

日本語や英語の技術用語の関連語を抽出・表示し、検索式の作成支援を行います。これは、Japio 独自の言語資源とノウハウを活用したもので、技術用語における日本語・英語の異表記語・関連語の把握と漏れの無い検索に役立ちます。

イ. 出願人検索支援

出願人名の日英中の異表記が参照できます。また、それらを Japio-GPG 用の検索式として自動生成し、コピー&貼り付けで検索に利用することにより、漏れの無い出願人検索が実現できます。

ウ. 絞込インデックス

検索結果に含まれる公報発行国や公報種別、IPC、キーワード、発明者、出願人の各項目に対し、絞込を希望する項目をワンクリックすることにより、検索結果の絞込が瞬時にできます。

検索用語等が文献頻度付きで表示されますので、容易に検索結果の絞込や分析ができます。

エ. 代表ファミリー表示

同じ発明が複数の国に出願されたいわゆる「パテントファミリー」を有する場合、検索回答が重複して表示されます。この重複表示を避けるため、「代表ファミリー表示」においては、パテントファミリーの中の代表的な言語（例えば日本語）を選択し、代表文献だけを表示します。この機能によりスクリーニングする回答文献を減らし検索効率を上げることができます。

その他

Japio-GPG は、以下の機能も備えています。

- ・しおり付与、表示
- ・検索式の登録、表示、再利用
- ・ファミリー分析
- ・検索結果のソート
- ・ダウンロード(書誌等最大 1,000 件)
- ・ファイル出力(全項目、1 件単位)

4 Japio-GPG の機械翻訳

(1) 2 種類の機械翻訳システム

Japio-GPG では、2 種類の機械翻訳システムを使用しています。一つは英語特許文献及び英訳された発明の名称、要約文、第 1 請求項を英日翻訳するための「ルールベース方式機械翻訳 (RBMT) システム」、もう一つは中国特許文献の発明の名称、要約文、第 1 請求項を中日翻訳するための「統計方式機械翻訳 (SMT) システム」です。以下に各々を簡単に紹介します。

(2) 英日機械翻訳 (RBMT) システム

Japio-GPG の英日機械翻訳システムは、特許文献用 RBMT エンジンに、Japio 独自の分類別の英日辞書を組み合わせることで、翻訳精度のさらなる向上を図っています。

Japio は、対応特許など実際の特許文献から採取した技術用語対訳データベースを保有しており、ここから分類別英日辞書を作成しました。これらの辞書には、各分類において最も使用頻度の高い訳語を採用し、技術分野に則した的確な翻訳結果が得られるように工夫しています。また、複合語を多数採用し、単語レベルでは適訳語の特定が困難な用語への対応も行っています。

(3) 統計方式機械翻訳 (SMT) システム

Japio は、NICT との共同研究により、中国と日本の対応特許から大量の中日対訳文対の生成に成功しており、この豊富な学習データを SMT システムに用いることで、高精度な翻訳を実現しました。

また、SMT は、学習データを統計的に処理して前後の語とのつながりの良い訳語を選定する方式ですので、分類別の辞書がなくても適切な技術用語が選定される確率が高いという利点があります。

(4) 2 種類の機械翻訳結果の併用による効果

Japio-GPG では、これら 2 種類の機械翻訳システムにより、例えば中国文献に関しては、まずは中国で公開された中国語文献の中日 SMT 翻訳結果、そして、その数ヶ月後には、その文献に関する DOCDB からの英文抄録情報を用いて、英日 RBMT 翻訳するという 2 種類の日本語情報を利用することができるため、両者を併用することで、より正確な文献内容の把握が可能となります。

5 Japio-GPG 全文検索への拡大

Japio-GPG サービスは、平成 24 年 7 月のサービス開始以降、皆様から多くのご期待、ご要望をいただいて参りました。これらのご期待、ご要望に応えるため、今春、検索対象を主要国の特許明細書全文に拡大し、機械翻訳を利用した日本語による横断検索サービスへの拡充を予定しています。

全文の検索対象は、日本（日本語、MT 英語）、中国（MT 日本語）、米国（MT 日本語、英語）、欧州（MT 日本語、英語）、PCT（MT 日本語、英語）の予定です。

Japio は、関係の皆様のご協力のもと、グローバルな特許検索環境の改善に寄与するために、特許明細書全文の検索対象国を広げるとともに、多言語の機械翻訳の高精度化に向けた努力を引き続き進めて参ります。

第 23 回 JTF 翻訳祭セッション開催報告

機械翻訳課題調査委員会

1. はじめに

AAMT 機械翻訳課題調査委員会は第 23 回 JTF 翻訳祭にて「機械翻訳による競争戦略—機械翻訳で差別化できるか—」と題し、パネリスト 3 名とモデレーターにより 90 分のパネルディスカッションを実施した。本稿では、当日のディスカッションの内容と、参加者からの反響などについて紹介する。

2. JTF 翻訳祭と AAMT

2013 年 11 月 27 日にアルカディア市ヶ谷において、日本翻訳連盟 (JTF) 主催の JTF 翻訳祭が執り行われた。毎年開催されるこの国内最大の翻訳業界のイベントの参加者は 800 名を超える。AAMT は 2 年前の第 21 回、そして第 22 回と「翻訳プラザ」と呼ばれる展示コーナーにブースを構え、課題調査委員会のメンバーにより、AAMT の活動を紹介するための各種資料の配布、UTX の紹介、機械翻訳に関するアンケートを実施してきた。

JTF と AAMT のさらなる連携の強化、そして翻訳業界および翻訳者の機械翻訳に対する興味を高めることを目的とし、第 23 回目となる本年の JTF 翻訳祭では、AAMT 企画の講演セッションを新たに開催することになった。

今年の翻訳祭では、6 つのトラックに分かれて行われた 24 の講演セッションの内の 1 つ (11:30~13:00) を AAMT が担当した。

3. 機械翻訳への高い関心

120 名を収容するセッション会場は満席であり、立ち見が出るほどの盛況ぶりだった。翻訳者、翻訳会社、そして翻訳を発注するクライアント企業の機械翻訳に対する関心はかなり高いようである。

時間的な理由からセッション内では多くの質問を受けることができなかったが、質問のために翻訳プラザの AAMT ブースを訪れるセッション参加者もあり、ここでも機械翻訳に対する関心の高さが伺えた。

4. 機械翻訳による差別化とは

パネリストは課題調査委員会のメンバーで構成された。秋元圭委員 (株式会社クロスランゲージアルファンドディ) は機械翻訳ソフト開発ベンダー、目次由美子委員 (株式会社 シュタール ジャパン) は翻訳メモリーツール開発ベンダーおよび翻訳会社、山本ゆうじ委員 (秋桜舎) は翻訳者および辞書の標準化推進というそれぞれの立場から意見を述べ、長瀬友樹委員長 (株式会社富士通研究所) がモデレーターとしてパネルディスカッションを取りまとめた。

Open Office XML 翻訳

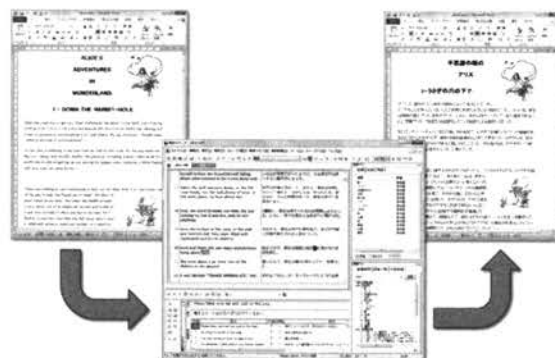


図 A : 機械翻訳ソフト PC-Transer の機能紹介

パネルディスカッションでは、ルールベース翻訳と統計翻訳との違いの紹介から始まり、MT（Machine Translation：機械翻訳）とTM（Translation Memory：翻訳メモリーツール）の連携を含む各ツールの最新動向などが紹介された。

翻訳支援ツールTransit の機械翻訳機能

2. エディターMT



図 B：翻訳メモリーツール Transit の機能紹介

さらに、翻訳プロセスを最適にするためには用語集の作成が重要である旨を解説した上で、AAMT 策定の UTX 用語集形式について紹介した。

UTX用語集の図解

さまざまな環境ですぐに共有・再利用できるように、データ形式を整える

用語集情報 #UTX1.20, en-US/ja-JP, 2013-04-01T19:00:00Z+09:00; copyright: AAMT (2013); license: CC-BY 3.0			
#src	tgt	src-pos	term status
Asia-Pacific Association for Machine Translation	アジア太平洋機械翻訳協会	proper noun	approved
dictionary administrator	辞書管理者	noun	approved
contributor	用語提供者	noun	provisional
domain	分野	noun	
glossary	用語集	noun	
bidirectional	双方向	adjective	approved
merge	統合する	verb	approved
原語	訳語	品詞	用語ステータス (省略可)

最適な訳語を1つ

図 C：UTX 用語集形式の紹介

また、機械翻訳を活用するポイントとして前処理が重要であること、つまり、人間にも機械にも読みやすい文章にすることで、より良い翻訳結果が得られることを指摘し、そのためにはクライアントの協力が不可欠であることを説明した。

現在、機械翻訳は翻訳メモリーツールと連携して活用されている一方で、品質は重視しない、早く安く提供される翻訳という新しいニーズも発生しつつあり、機械翻訳のさらなる活躍が期待されることも紹介した。

新しい翻訳ニーズに向けて

今後の機械翻訳活用法(早い、安い、品質そこそこの新マーケット)



現在の機械翻訳活用(翻訳メモリーとの連携)



図 D：機械翻訳の活用例

次世代の翻訳ツールとして、コンテンツマネージメントシステムのような企業単位で利用する大規模な翻訳ニーズへの対応、既存の翻訳資産を自動的に学習するツール、ルールベース・用例ベース・統計翻訳の融合による翻訳品質の向上などにも言及した。

5. 来年に向けて

翻訳プラザに設けた AAMT ブースには、セッション参加者の訪問があった。「聞きたいことが聞けなかった」(テーマに沿っていなかった)という厳しいご意見も受けたが、各社のツールについての質問も受け、最新動向および機械翻訳とは何かが分かったという声も聞かれた。翻訳業界における機械翻訳に対する関心の高さを実感できたことは、AAMT にとって大きな収穫だったと言える。来年の翻訳祭でも JTF の協力を得て AAMT セッションを開設し、翻訳業界における機械翻訳の活用を推進していきたい。

今回のパネルディスカッションは翻訳会社に魅力的なテーマを設定して集客に成功した一方で、その内容は必ずしも聴衆を満足させるものではなかったと認識している。参加者から受けたコメントやアンケート結果を検討材料とし、時間をかけて、来年の翻訳祭へ向けてじっくりと準備したい。たとえば、翻訳依頼側と受託側の双方が満足するような図式を、実例と共に示すなど工夫したい。

JTF との連携をさらに強化させて幅広く情報を共有することで、翻訳業界における機械翻訳に対する疑問や要望をより良く理解できると考えている。機械翻訳が翻訳業界で活躍する場を広げるためにも、機械翻訳課題調査委員会として JTF と協同して調査・広報活動を積極的に進めていきたい。

新製品紹介

Android スマートフォン向け文字入力システムと連携したクラウド型翻訳アプリ

翻訳キーボード

株式会社 高電社

株式会社高電社(本社：大阪市阿倍野区、代表取締役：岩城 陽子)は、オムロンソフトウェア株式会社(本社：京都市下京区、代表取締役社長：宮地 功)の文字入力システム「Wnn Keyboard Lab」と連携した、Android 向け翻訳アプリ『翻訳キーボード』を2013年11月より配信開始いたしました。

『翻訳キーボード』は、Android スマートフォン向け文字入力システムとして数多くのユーザーから支持される「Wnn Keyboard Lab」から直接起動することができるクラウド型翻訳アプリです。

メールやチャットはもちろん、文字入力を行う様々なシーンで利用することができます。

■本製品の主な特徴

- 文字を入力しながら簡単に翻訳が可能

「Wnn Keyboard Lab」の記号キーを長押しするだけで翻訳できるので、入力中の文章を保存して別の翻訳アプリを起動するなど、面倒な手順は一切不要です。入力した文字を簡単にスムーズに翻訳をすることができます。

●一括翻訳・部分翻訳が選択可能

入力した文章が、複数の句点を含む長い文章でも、一度に翻訳することができます。また、必要な部分を範囲選択して記号キーを長押しすれば、部分翻訳も可能です。

●英語・中国語・韓国語に同時翻訳可能

日本語の文章を入力して、複数の言語を訳文に指定すれば、複数言語の同時翻訳も可能です。SNS などのコミュニケーションツールなどを利用して、日本語の文章を複数の言語に翻訳してメッセージを送りたい



場合にも便利な機能です。また、編集エリアにある外国語も日本語に翻訳することができます。

■配信ページ（Google Play で配信）

<https://play.google.com/store/apps/details?id=jp.kodensha.wnnconnector>

■製品紹介ページ

http://www.kodensha.jp/mobile/android/honyaku_keyboard/

■注意事項

※「翻訳キーボード」のご利用には、Wnn Keyboard Lab 本体が必要です。

（Wnn Keyboard Lab 本体は、Google Play より無料でインストールできます。）

※「翻訳キーボード」の利用には、携帯電話回線(3G、4G)や Wi-Fi 環境などのインターネット接続環境が必要です。

■本体価格

無料

■サービス利用料金

1 言語のみ：月額 105 円（税込）

2 言語以上：月額 210 円（税込）

※ダウンロード後、30 回まで無料で翻訳をお試し頂けます。

■商標について

「Wnn」（うんぬ）は、オムロン株式会社の登録商標です。

■本製品に関するお問い合わせ先

株式会社 高電社

・大阪本社

大阪市阿倍野区昭和町 3-7-1

ソフトウェア事業部 開発 3 課

TEL：06-6628-2195

E-Mail：asupport@kodensha.jp

特許用語・専門用語を標準搭載した韓国語翻訳ソフト

J ソウル 9

株式会社 高電社

株式会社高電社は、日韓・韓日翻訳に対応した翻訳ソフト「J ソウル 9」（ジェイ・ソウル・ナイン）のダウンロード版を 2014 年 1 月 17 日より、パッケージ版を 2014 年 2 月 7 日より発売開始しました。

「J ソウル 9」は、特許用語・専門用語辞書を標準搭載することで、幅広い分野において高い翻訳精度を実現しました。また、作業効率を高める機能を多数搭載しており、翻訳業務の効率アップやコスト削減を図ることができます。

1. 新機能と強化ポイント

「J ソウル 9」で翻訳精度向上および翻訳支援環境を充実させるために追加された強化ポイントを示します。

(1) 特許用語・専門用語辞書

機械工学・化学・医学など 25 分野 140 万語以上の特許用語・専門用語辞書を標準搭載しています。収録されている用語は、膨大なデータの中から厳選したもののみを使用しており、単に収録語数を増やすだけでなく、全体の訳質を向上させるように工夫されています。

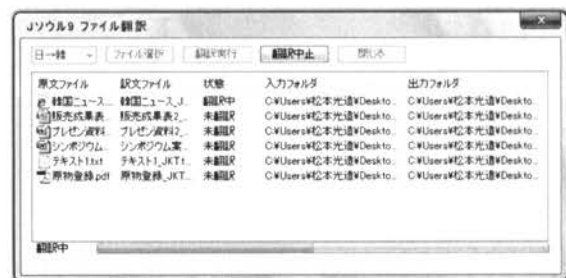
また、韓国を含めた非英語圏の特許出願数が急増しており、知的財産戦略や特許調査の重要性が高まっています。膨大な特許情報をスクリーニングしたり、韓国特許庁からの拒絶理由通知書を素早く把握したりする場面において、「J ソウル 9」はコスト的・時間的に絶大な効果を発揮します。



対訳エディタ画面

(2) ファイル翻訳

ファイルエクスプローラーで複数のファイルを選択したあとに、右クリックメニューで翻訳方向を指定するだけで、ファイルを開かずダイレクトに翻訳できる機能です。「J ソウル 9」では、Office ファイルや PDF ファイルに加え、特許情報データの標準形式である XML などの形式や動画の字幕などで多く採用されている SMI 形式にも新たに対応し利便性が高まっています。

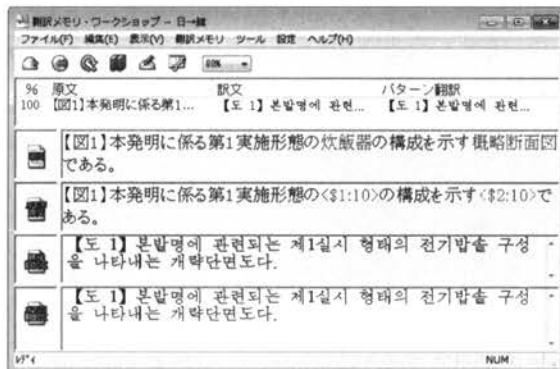


ファイル翻訳中の画面

(3) 翻訳メモリ

原文と対応する訳文を翻訳メモリ辞書にあらかじめ登録・蓄積することで、以後の翻訳作業を効率化する翻訳支援ツールです。「J ソウル 9」では、文字列や数字の場所をパターン登録することで、類似の文章に対しても適切な単語や数字などに置き換えて訳文が出力されます。

また、過去の翻訳で蓄積した翻訳資産を活用できるよう、Unicode 形式のタブ区切りテキストファイルやTrados 等で作成したTMX ファイルをインポートすることもできます。



翻訳メモリ ワークショップ画面

3. 翻訳精度

旧バージョンおよび Google 翻訳サービスとの翻訳結果（日韓、韓日）の例を示します。

（1）日韓翻訳

■原文

電力増幅器、方向性結合器、Duplexer などを一括して高周波モジュールを製作した場合、高調波レベル制御を容易かつ効率的に行うことができ、電力増幅器に用いられる電力増幅用半導体素子の特性を最大限に引き出すことができる高周波モジュールを提供する。

■訳文（J ソウル 9）

전력 증폭기, 방향성 결합기, Duplexer 등을 일체화해서 고주파 모듈을 제작했을 경우, 고조파 레벨 제어를 용이하고 효율적으로 실시할 수 있어, 전력 증폭기에 이용할 수 있는 전력 증폭용 반도체 소자의 특성을 최대한으로 끌어낼 수 있는 고주파 모듈을 제공한다.

■訳文（j・Seoul V8）

전력 증폭기, 방향성 결합기, Duplexer 등을 일체화해서 고주파 모듈을 제작했을 경우, 고주파 수준 제어를 용이 동시에 효율적으로 할 수 있고, 전력 증폭기에 이용할 수 있는 전력 증폭용 반도체소자의 특성을 최대한에 끌어낼 수 있는 고주파 모듈을 제공한다.

■訳文（Google 翻訳）

전력 증폭기, 방향성 결합기, Duplexer 등을 일체화하여 고주파 모듈을 제작하는 경우 고조파 레벨 제어를 쉽고 효율적으로 할 수 있는 전력 증폭기에 사용되는 전력 증폭 용 반도체 소자의 특성을 최대한으로 끌어낼 수 있는 고주파 모듈을 제공한다.

※Google 翻訳では、時制誤り、連結部分の不自然さ、分ち書き誤りが見られます。

（2）韓日翻訳

■原文

도 1 은 일반적인 주문형 단전문서서비스 시스템의 구성도로서, 코드분할다중접속(CDMA)신이동통신 시스템의 구성을 나타낸다.

■訳文（J ソウル 9）

図 1 は一般的なオンデマンド ショートメッセージ サービス システムの構成図として、符号分割多元接続(CDMA)移動通信システムの構成を現わす。

■訳文（j・Seoul V8）

図 1 は一般的な注文型停電文書サービス システムの構成図として、コード分割多重接続(CDMA)移動通信システムの構成を現わす。

■訳文（Google 翻訳）

図 1 は、一般的なオンデマンドただし専門的なサービスシステムの構成図であり、符号分割多元接続 (CDMA) 移動通信システムの構成を示す。

6. 翻訳業務を支援する便利な機能

翻訳業務に大変便利な機能やツールをいくつかご紹介いたします。

(1) テキスト文字比較ツール

「くらべてきすと！」は、2つのウインドウ上のテキストを比較して、相違点を色分け表示できるツールです。比較するテキストの改行位置が異なっている場合や、相違点が少なく内容が似たテキストでも、色分け表示で相違点を明確にすることが可能です。日本語は Shift-JIS コード・Unicode、韓国語は KS コード・Unicode に対応しています。

(4) 多言語 OCR ツール

186 言語に対応した多言語 OCR（文字認識）ツール「FineReader」を搭載しました。紙文書をスキャンしたデータや PDF・JPG などの画像ファイルを読み込み、OCR 処理することで、レイアウトを保持したままテキスト抽出可能なファイルに変換することが可能です。単一言語はもちろん、多言語が混在した文書でもテキスト抽出可能です。OCR ツールによるテキスト抽出機能と「J ソウル 9」の翻訳機能を利用することで、画像化されたファイル（PDF・JPG 等）もスムーズに翻訳することができ大変便利です。

(5) 韓国語入力支援ツール

「J ソウル 9 プレミアム」製品には、韓国で標準的に使用されている韓国語の文字コード「KSコード」での入力に加え、シフトJISコードにハングルや韓国語漢字を対応させた弊社オリジナル文字コード「KWコード」での入力が可能です。

また、見やすくデザイン性に優れた韓国語フォント、「簡単ハングル入力モード」や「日韓翻訳入力」をはじめとした入力支援ツールも多数搭載しています。

4. 製品と仕様

■製品ラインナップと価格

J ソウル 9 エキスパート	38,000 円＋税
J ソウル 9 プレミアム	53,000 円＋税

■動作環境

OS	Windows 8.1/8/7 日本語版など
HDD	エキスパート：4.9GB 以上 プレミアム：7.2GB 以上

■製品紹介ページ

<http://www.kodensha.jp/soft/js/>

5. おわりに

高電社における最初の韓国語ソフトは 1991 年に発売の MS-DOS 対応の日韓翻訳ソフト「j・ソウル」にさかのぼります。国内唯一、自社開発の韓国語翻訳エンジンを搭載した本ソフトは、23 年にわたり機械翻訳技術の研究を積み重ねるとともに、多くのユーザー様から寄せられたご要望に応えつつ進化して参りました。これからも、高電社は「言語の壁を超え、世界中の人々の心と心をつなぐ」という理念のもとに、さらに優れた翻訳ソフトやサービスの開発・提供に努めて参ります。

これまでの AAMT Forum メールマガジン

機械翻訳課題調査委員会 WG1・2

はじめに

これまで AAMT 会員の皆様に機械翻訳に関する様々な情報を提供させていただくため、AAMT Forum が活用されていました。例えば 2008 年 1 月から 3 月の 3 ヶ月間で 22 件のメールが配信されていました。その配信がめっきり少なくなったことが AAMT 機械翻訳課題調査委員会の席で話題となり、AAMT Forum の刷新の検討が始まりました（2012 年中期頃）。

委員会で検討した結果、定期的な配信を継続させるため、メルマガとして知られるメールマガジン形式への転換を試みることとなりました。その結果、2013 年 7 月 19 日配信の AAMT Forum メールマガジン第 1 号が皆様のもとへ配信されました。その後、隔週の金曜日に配信し、2014 年 1 月 9 日配信分で 12 号となりました。

最近、AAMT 会員の皆様を始め AAMT 関係者からご好評を耳にする機会がちらほらとあり、大変嬉しく思っております。そこで、皆様と一緒にこれまでのメルマガを振り返り、さらに良いメルマガへと発展させることができればという欲張りな気持ちが湧いてきたため、本誌において一つの節目としてこれまでの配信分をバックナンバーとして掲載させていただくこととなりました。ご一読いただき、お気づきの点などがありましたら AAMT 事務局までご意見をお寄せいただけましたら幸いです。

では、メールマガジン第 1 号から第 12 号までを再びお届けします。

AAMT Forum メールマガジン 1 号（2013/7/19 配信）

こんにちは。AAMT Forum メールマガ担当です。今回、メールマガジンは装いも新たに再スタートを切ります。製品情報からこぼれ話まで、機械翻訳に関する様々な情報をお届けいたしますので、ご愛読いただけますようお願いいたします。

■ パーソナル／ビジネス向け／医学向け翻訳ソフト新発売！

株式会社クロスランゲージから低価格英日・日英翻訳ソフトシリーズが 2013 年 7 月 19 日（金）より販売されます。

<http://prtimes.jp/main/html/rd/p/000000058.000002804.html>

■ Twitter、Bing 翻訳機能で 43 カ国語の自動翻訳が可能にデスクトップの Web ブラウザからアクセスする公式 Twitter サイトに自動翻訳機能が追加されました。

<http://www.itmedia.co.jp/news/articles/1307/01/news120.html>

■ 翻訳 企業に依頼で納期 1 ヶ月 450 万円 ネットで 7 日 7.5 万円

ネットの発達により薄利多売がいろいろな業種で可能になりました。翻訳もその例に漏れないようです。納期が 1/4、費用が 1/60 で翻訳精度も満足？！

http://www.news-postseven.com/archives/20130607_192062.html

■ プロ翻訳者のツール利用状況レポート

プロの翻訳者はどんなツールを使っているのか？自動集計によるランキングがほぼリアルタイムで公開されています。

<http://www.translator.jp/soft/>

■Google 翻訳で「do you」を日本語翻訳するとあのキャラ降臨

文章をほかの言語に自動翻訳してくれるサービス「Google 翻訳」。このたび「do you」と打ち込んで日本語訳すると、なぜか「やるお」と翻訳されることがネット上で話題になっています。

<http://nlab.itmedia.co.jp/nl/articles/1306/19/news125.html>

■スマホアプリ紹介：TranScope：スマホをかざせば即翻訳！英語・中国語・韓国語に対応

昔、香港に旅行した時、英語がまったく通じないレストランに入り、注文で苦労しました。こんなアプリがあったら、何が運ばれてくるかハラハラしないですみますね。

<http://octoba.net/archives/201300329-android-app-transcope-188142.html>

AAMT Forum メールマガジン 2号（2013/8/2 配信）

こんにちは。AAMT Forum メールマガ担当です。このところ、不安定な天候が続いていますがいかがお過ごしでしょうか？8月は夏祭りのシーズンですね。どうぞ天候に恵まれますように。では、再スタート後の第二号をお送りします。

本メールマガジンでは、製品情報からこぼれ話まで、機械翻訳に関する様々な情報をお届けいたします。

■Google 翻訳のデスクトップ向け Web サイトで手書き入力をサポート
米 Google は、Google 翻訳サービスにて手書き入

力機能を設け、漢字の入力にも対応しました。

<http://news.mynavi.jp/news/2013/07/25/227/>

■日本語を英語に自動変換する独自システムで英語サイトと完全連携を開始

株式会社 HEARTS は、独自システムで自動翻訳を行うことにより自動車関連パーツを販売する自社運営サイトと英語圏向け販売サイトのリアルタイム連携を可能にしました。

<http://japan.cnet.com/release/30048911/>

■「1秒間に約4語」世界9000人が翻訳 ゲンゴの挑戦

ゲンゴ社の翻訳サービス (<http://gengo.com/ja/>) は世界に散らばる約9000人の翻訳家のネットワークを使い、安くてスピーディーで高品質な翻訳を提供。

<http://www.nikkei.com/article/DGXBZO57739780V20C13A7000000/>

■Kinect 活用、手話をリアルタイムで翻訳

「Kinect」等を使って、手話とテキストをその場で翻訳できるシステムが各種開発されています。

<http://wired.jp/2013/07/23/sign-language-translation-kinect/>

■スマホアプリ紹介：TransApp：iPhone アプリ /Android アプリに特化した、日本初の翻訳サービス

日本初、iPhone アプリ/Android アプリ専門の翻訳サービス『TransApp』（トランスアップ）サービス提供開始。

<http://octoba.net/archives/201300329-android-app-transcope-188142.html>

AAMT Forum メールマガジン 3号（2013/8/23 配信）

こんにちは。AAMT Forum メルマガ担当です。連日、厳しい暑さが続いていますね。皆さまはいかがお過ごしですか？私はコンビニでアイス買って帰るのがすっかり日課になってしまいました。ちなみに最近のお気に入り、は、「贅沢最中」シリーズの「ゴルゴンゾーラチーズ&いちじく」です。皆さまもアイスで暑さ対策はいかがですか？では、メールマガジン第3号をお届けします。

■Android版「Google 翻訳」v2.8 にアップデート

Android版「Google 翻訳」がv2.8にアップデートされ、手書き・カメラ入力の手書き言語が拡大しました。

<http://getnews.jp/archives/399690>

■『モンスターズ』続編、興収 70 億突破！ヒットを生み出す吹替え台詞の妙

日本語吹替え版翻訳を手掛けてきた映像翻訳家・杉田朋子氏インタビュー。性別、世代、国境をも越えて愛される作品を、日本人の心に届ける翻訳の妙とは？

http://www.47news.jp/topics/entertainment/oricon/movie_anime/131980.html

■米フェイスブック、翻訳アプリ「ジビゴ」開発会社の買収検討

フェイスブックは、音声認識・言語翻訳アプリメーカー、米モバイル・テクノロジーズの買収を検討。

<http://jp.reuters.com/article/topNews/idJPTYE97C00A20130813>

■スマホ＆ウェブアプリ向け翻訳に特化した新サービス

八楽株式会社は、スマホ＆ウェブアプリ向け翻訳に特化した新サービス「WorldJumper for アプリ」の提供を開始しました。

<http://prw.kyodonews.jp/opn/release/201308163908/>

■語学学習ツールとしての Google 翻訳

Google 翻訳や相互添削型 SNS を使って試行錯誤しながら学ぶ語学学習法が紹介されています。

<http://mongolia.seesaa.net/article/372254581.html>

AAMT Forum メールマガジン 4号 (2013/9/6 配信)

こんにちは。AAMT Forum メルマガ担当です。台風の影響により西日本を中心に大雨が続きましたが、いかがお過ごしでしたでしょうか？雨が止めばまた暑い日が戻ってくるこのごろ、体調にはお気を付け下さい。では、メールマガジン第4号をお届けします。

■NAIST、文末を待たずに翻訳を開始する同時自動音声通訳技術を開発

精度を維持したまま 20%のスピード向上を実現できることが確認され、また、通訳経験初級者(1年以上経験者)と同等の同時通訳性能と速度も達成されました。

<http://news.mynavi.jp/news/2013/08/28/181/>

■Google 翻訳がアフリカのローカル言語に対応

それぞれ約 1,000 万人から 2,000 万人が話している言語である、ハウサ語、イボ語、ソマリ語、ヨルバ語、ズールー語の 5 言語に対応しました。

http://www.icr.co.jp/newsletter/global_perspective/2013/Gpre201398.html

■平成 24 年度中国特許文献の機械翻訳のための中日辞書整備及び

機械翻訳性能向上に関する調査報告書について
特許庁が辞書データ及び対訳コーパスの作成を行い、辞書データ、中日対訳コーパスを機械翻訳ソフトウェアに追加した場合の翻訳品質の向上について調査を行いました。

http://www.jpo.go.jp/shiryoutoushin/chousa/h24_chunchi_honyaku.htm

■音声翻訳アプリ「NariTra」が産学官連携功労者表彰 総務大臣賞を受賞！

「NariTra」の技術開発から実用化、無償公開までの一連の普及促進・社会還元活動が評価されました。

<http://www.nict.go.jp/press/2013/08/29-2.html>

AAMT Forum メールマガジン 5号 (2013/9/20 配信)

こんにちは。AAMT Forum メールマガジン担当です。秋の気配が近づき、暑さも少しずつ和らぎ過ぎやすくなってまいりました。とはいえ、先の大型台風など、天候が不安定になりやすい季節の変わり目でもあります。急な天気の変化にご注意ください。では、メールマガジン第5号をお届けします。

■クラウド翻訳 Conyac のエニドア、サイト翻訳管理ツールの八楽と東京オリンピックに向けて業務提携

提携の第一弾としてワールドジャンパーの翻訳修正機能に Conyac の翻訳サービスが搭載され、低価格な修正翻訳がクラウドソースで出来るようになります。

<http://sankei.jp.msn.com/economy/news/130910/prl13091012170031-n1.htm>

■未来技術遺産:九大の自動翻訳実験用計算機、選定

1950年代、国内で初めて開発された日英独3カ国語の自動翻訳機で、その後の自然言語処理研究の原点になった計算機です。

<http://mainichi.jp/feature/news/20130911ddlk40040329000c.html>

■第2回「未来の翻訳研究に関するワークショップ」開催

本ワークショップは「クラウドソース翻訳」と「翻訳支援技術」に注目しており、大阪駅直近の「グランフロント大阪」で開催されます。

<http://www.mastar.jp/wfdtr/index.html>

■PUBMED の日本版がついに登場！・・・かと思いきや、間違えようのない、高精度翻訳機の超巨大なお試し訳サイトを公開

株式会社ロゼッタは、同社が開発している翻訳機のお試し訳サイトとして PUBMED（世界中の医学・薬学・ライセンサイエンス関連の論文や文献をリアルタイムで収集し公開しているサイト）の日本語版のように一見みえるサイトを公開しました。

http://www.asahi.com/and_M/information/pressrelease/PRT201309170001.html

AAMT Forum メールマガジン 6号 (2013/10/4 配信)

こんにちは。AAMT Forum メールマガジン担当です。日が落ちるのが早まり、夜長の頃となってきました。過ごしやすいつまみが増えるこの季節、皆様、有意義に使われていますでしょうか。では、メールマガジン第6号をお届けします。

■手のひらサイズで25か語以上の音声翻訳が可能になる「SIGMO」

本体サイズは4センチ×4センチと、とてもコンパクト。既存のオンライン翻訳システムを使用しており、専用アプリをインストールした iPhone や Android スマートフォンと Bluetooth 接続して動作します。

<http://gigazine.net/news/20130924-sigmo-voice-translating-device/>

■NICT と KI、IT 英文マニュアルに特化した「自動

翻訳ソフトウェア」を開発

情報通信研究機構(NICT)と川村インターナショナル(KI)は人手翻訳のコストを 30%削減した英日の「自動翻訳ソフトウェア」を開発し、10月1日より同ソフトウェアを活用したサービスの提供を開始しました。

<http://news.mynavi.jp/news/2013/10/01/247/>

■iPad 対応多言語メニューアプリ 「MOMONGA 翻訳・外国語メニューBOOK」

エクスウェア株式会社は、飲食店向けに、日本語で作成したメニューを様々な国の言語に自動翻訳する iPad 対応多言語メニューアプリ『MOMONGA 翻訳・外国語メニュー』を、2013年10月1日より無料公開しました。

<http://www.atpress.ne.jp/view/39218>

AAMT Forum メールマガジン 7号 (2013/10/18 配信)

こんにちは。AAMT Forum メルマガ担当です。吹きつける風に涼しさよりも肌寒さを感じやすいこの頃、皆様はもう衣替えを済ませたでしょうか。季節に合わせた服装で快適に過ごしましょう。では、メールマガジン第7号をお届けします。

■自社 Web サイトを安価かつ高精度に多言語対応するには

Web サイトを多言語化する必要性と Web サイトの翻訳サービス「WorldJumper」の仕組みを八楽代表取締役の坂西優氏が述べています。

<http://it.impressbm.co.jp/e/2013/10/03/5143>

■さらなる翻訳業務効率の向上を実現するサーバ型翻訳ソフトウェア「The 翻訳エンタープライズ V16」を発売

東芝ソリューション株式会社は、中日/日中・韓日/日韓の翻訳精度を向上させるとともに、ホーム

ページ翻訳などの業務効率を高める機能を充実させたサーバ型翻訳ソフトウェア「The 翻訳エンタープライズ V16」を10月7日より発売しました。

http://www.toshiba-sol.co.jp/news/detail/131007.htm?_1ink=top

■コピーした英単語を通知バナーで即翻訳！単語帳にもなる辞書アプリ

辞書アプリ「Biscuit--辞書&単語帳」は英単語を調べる手間をできるだけ減らしたアプリです。英単語をコピーした瞬間、通知バー上に翻訳します。

<http://www.appbank.net/2013/10/14/iphone-application/682888.php>

AAMT Forum メールマガジン 8号 (2013/11/1 配信)

こんにちは。AAMT Forum メルマガ担当です。伊豆大島をはじめ、各地に甚大な被害を及ぼした季節はずれの台風 26・27 号ですが、皆様は大事無かつたでしょうか。自然の猛威を改めて感じさせられました。では、メールマガジン第8号をお届けします。

■第23回 JTF 翻訳祭

翻訳者、翻訳会社、クライアント、翻訳支援ツールメーカーなど業界関係者が一堂に会する国内最大規模の翻訳フェスタが2013年11月27日(水)に東京都アルカディア市ヶ谷で開催されます。

<http://expotoday.com/event/3620/>

■ジゾンと Translations.com が戦略パートナー提携

株式会社ジゾンは、同社が開発・販売・保守を手掛けるウェブコンテンツ管理システム (CMS) 「HeartCore」と、Translations.com が提供するローカリゼーション管理ツール「GlobalLink Localization Suite」を連携することで合意したと発表しました。

<http://www.zaikai.co.jp/releases/131175/>

■業務向け翻訳ソフト「PC-Transer 翻訳スタジオ V21 for Windows」

株式会社クロスランゲージは、ビジネス・科学技術文書に適した業務用の翻訳ソフトの新バージョン「PC-Transer 翻訳スタジオ V21 for Windows」を、2013 年 11 月 29 日（金）より販売を開始します。
<http://sankei.jp.msn.com/economy/news/131028/prl13102815300073-n1.htm>

**AAMT Forum メールマガジン
9 号（2013/11/15 配信）**

こんにちは。AAMT Forum メルマガ担当です。寒さが厳しくなりはじめ、地域によっては暖房器具を利用されているでしょうか。初雪も観測されはじめ、冬の訪れを感じるこの頃、寒さへの備えをしっかりしましょう。では、メールマガジン第 9 号をお届けします。

■五輪見据え、企業利用に耐える英・中・韓へ翻訳無料

自動翻訳ソフトウェアの開発などを手がけるベンチャー企業の八楽は 2013 年 11 月 6 日、企業の Web サイトの翻訳サービス「ワールドジャンパー」を無料で提供するサービスを開始しました。
http://www.nikkei.com/article/DGXNASFK06032_W3A101C1000000/

■特許公報全文、日本語検索 OK 日本特許情報機構が来春新サービス

日本特許情報機構（J a p i o）は発明通信社と共同で、日本と米国、欧州、中国、特許協力条約（P C T）の特許公報の全文を日本語で横断的に検索できる新サービスを、来春をめどに始めます。
<http://www.sankeibiz.jp/compliance/news/131105/cpd1311050600007-n1.htm>

■フィリップ・コーン博士インタビュー

10 月に開催された NICT 主催「未来の翻訳研究に関するワークショップ」会場におけるエディンバラ大学のフィリップ・コーン博士へのインタビューになります。

<http://jtfjournal.homepage.com/special/id=254>

**AAMT Forum メールマガジン
10 号（2013/12/6 配信）**

こんにちは。AAMT Forum メルマガ担当です。師走となり、今年も残すところあとわずかとなりました。何かと気忙しくなる年の瀬ですが、小さな見落としなどがないようにしたいものです。では、メールマガジン第 10 号をお届けします。

■中国特許に特化したサーバー組み込み型自動翻訳エンジンが販売開始！

株式会社高電社は、中国特許に特化したサーバー組み込み型自動翻訳エンジン『J-SERVER Advance 中国語特許エディション』を 2013 年 12 月 2 日（月）より販売開始しました。
http://www.asahi.com/and_M/information/pressrelease/ATP201312020019.html

■50 以上の言語の翻訳が行える「Just Translate」が Mavericks に対応

インフィニシスは、50 以上の言語の翻訳がワンクリックで行えるソフト「Just Translate」(Mac 版)のバージョン 2.9 を、同社運営のインターネット通信販売サイトを通じて販売開始しました。
<http://news.mynavi.jp/news/2013/11/12/430/>

■医療関係者に専門性の高い日英・英日翻訳を提供
翻訳ソフトメーカーの株式会社クロスランゲージが、医学・医療文書の翻訳に特化して開発した英日・日英翻訳ソフトの新バージョン「MED-Transer V13 for Windows」を、12 月 13 日より販売開始と発表しました。

<http://www.qlifepro.com/news/20131129/122013-to-launch-translation-software-specialized-in-the-publication-of-the-medical-and-health-care.html>

AAMT Forum メールマガジン 11号 (2013/12/18 配信)

こんにちは。AAMT Forum メールマガ担当です。年の瀬もいよいよ押し詰まってきたこのごろ、忙しくお過ごしかと思います。最後まで気を抜かず、気持ちのよい新年を迎えられますように。では、メールマガジン第11号をお届けします。

■特定の意味を含む文書を高速に検出するテキスト含意認識技術を開発

NEC は、大量のテキストデータから特定の意味を含む文書を漏れなく検出する「テキスト含意認識」を、従来に比べて約 24,000 倍高速化することに成功しました。

http://jpn.nec.com/press/201311/20131114_03.html

■ウェブ翻訳ツール「WorldJumper」、個人・小規模店舗向けサービス開始

業界最低水準となる格安のプランを開発したことで、これまで翻訳サービスの利用機会がなかった方でも、精度の高い外国語サイトを手間なく運用できる環境の創出を目指しています。

<http://prw.kyodonews.jp/opn/release/201312126989/>

■韓国語翻訳アプリ「ジーニートーク」、英語・日本語に続き中国語機能を追加

ソウルで開催された創造経済博覧会で韓国電子通信研究所が開発した自動翻訳アプリ「ジーニートーク (Genie Talk)」の韓国語・中国語翻訳機能が紹介されました。

<http://news.livedoor.com/article/detail/8350971/>

AAMT Forum メールマガジン 12号 (2014/01/10 配信)

こんにちは。AAMT Forum メールマガ担当です。新年あけましておめでとうございます。新しい年を迎え、気持ちも新たに自分の目標に向かって頑張りをしましょう。では、メールマガジン第12号をお届けします。

■ほんやくコンニャクまでもう少し!? 次期 AQUOS の翻訳機能がすごい

ソフトバンクモバイル向け「AQUOS PHONE Xx 302SH」および「AQUOS PHONEXx mini 303SH」に搭載の、英語の文章をカメラのファインダーを通すと日本語にすぐに翻訳してくれる「翻訳ファインダー」機能の紹介です。

<http://news.livedoor.com/article/detail/8246944/>

■ウェブで調べた外国語の意味を復習できる Chrome 拡張機能

単語や語句、文章を範囲選択すると翻訳がポップアップされ、意味や発音も調べられるブラウザー拡張機能の紹介です。

<http://weekly.ascii.jp/elem/000/000/190/190853/>

実務翻訳における機械翻訳の利用に関するアンケート結果報告

機械翻訳課題調査委員会 WG2

1. はじめに

AAMT 機械翻訳課題調査委員会 WG2（調査・広報・啓蒙）では、産業翻訳業界における機械翻訳システム活用の実態を把握することを目的として、2012 年と 2013 年に開催された JTF 翻訳祭においてアンケートを実施した。JTF 翻訳祭とは、翻訳者、翻訳会社、翻訳発注企業などの翻訳業界団体である日本翻訳連盟（JTF）の主催により、講演会、展示会で構成される 1 年に 1 度のイベントである。

本稿はそのアンケートの報告である。まず 2013 年度に実施したアンケート結果を説明し、その後で 2012 年度の結果と対比させて経年変化について論じる。

2. アンケートの概要

● 調査実施日

第 22 回 JTF 翻訳祭当日（2012 年 11 月 28 日）

第 23 回 JTF 翻訳祭当日（2013 年 11 月 27 日）

● 調査方法

アンケート用紙（A4 表裏）配布による無記名の調査とした。

翻訳祭の展示コーナーに AAMT がブースを確保し、ブースの前を通りかかる人に AAMT の委員から直接協力を依頼した。AAMT ブースにアンケート用紙と筆記用具を用意しておき、その場でアンケートに回答いただき、回収した。なお、回答者に抽選で当たる景品として、500 円分クオカード 10 枚と、翻訳ソフト「PC-Transer」（クロスランゲージ株式会社様のご好意による供出）を用意した。

● 回答状況

回答者数は、2012 年度が 60 名、2013 年度が 64 名であり、翻訳祭参加者（800 名と仮定）全体に対する回答率はどちらの年もおおよそ 8%であった。

3. 2013 年度アンケートの調査結果

本節では、回答者のうち翻訳作業そのものを生業とする 35 人のみを対象とした集計結果を紹介する。なお、内訳の合計が 35 人に満たない項目があるが、これは回収したアンケートに未回答の項目が含まれるためである。

3.1. 回答者

● 回答者の職種

翻訳業務の従事者といっても様々な仕事がある。以下は回答者 64 名の職種の内訳である。

[1] フリーランスの翻訳者 17 人

[2] 翻訳会社の社内翻訳者 7 人

[3] 企業（翻訳会社以外）・官公庁等の翻訳者

11 人

[4] 翻訳会社の翻訳業務管理者、コーディネーター

15 人

[5] 企業（翻訳会社以外）・官公庁等の翻訳業務管理者、

コーディネーター 6 人

[6] その他・無回答 8 人

回答者 64 名のうち 35 名（55%）が翻訳を生業としている人であり、21 名（33%）が翻訳の管理を仕事にしている人であった。

● 性別・年齢

年齢は 20 代から 70 代まで広く分布しているが、40 代が 29 人と突出して多かった。以下、30 代（17 人）、20 代台と 50 代（ともに 6 人）と続く。

性別は女性 42 人に対して男性 22 人であり、女性の回答者の割合が多かった。

3.2. 機械翻訳の利用頻度

翻訳者が普段どのくらいの頻度で機械翻訳を使って

いるかについて、翻訳者のタイプ別に集計した結果を表 1 に示す。

表 1 機械翻訳の利用頻度

	フリーランス	翻訳会社	企業・官公庁
[1] ほぼ毎日	7	2	5
[2] 週 1～3 回	3	2	1
[3] 月に数回	2	3	1
[4] 年に数回	1	0	1
[5] 使っていない	4	0	2

質問：翻訳ソフト/サイトの利用頻度は？

翻訳者の 4 割以上（14 人）が機械翻訳を「ほぼ毎日」使っていると回答している。「使っていない」と答えた人は 17%（6 人）にすぎず、機械翻訳は翻訳者に幅広く浸透しているといえそうである。

3.3. 利用している機械翻訳の種類

機械翻訳は、インターネット上の無料サービスから、一本が 10 万円以上する高価なソフトまで、様々な形態で提供されている。翻訳者はどのようなソフトあるいはサービスを使っているのだろうか。表 2 は翻訳者が良く利用する機械翻訳の種類についてまとめたものである。

表 2 よく利用する機械翻訳の種類

	フリーランス	翻訳会社	企業・官公庁
[1] 無料の翻訳サイト	10	2	4
[2] 有料の翻訳サイト	1	0	1
[3] 自社または翻訳会社などで提供される翻訳システム(サービス)	0	1	1
[4] 市販の PC 翻訳ソフト	2	4	1

質問：現在最もよく使う翻訳ソフトや翻訳サイトは？

翻訳者がよく使っている機械翻訳のトップは、無料の翻訳サイトという結果となった。翻訳会社等から翻訳ソフトが提供される例は少数に留まり、翻訳者が個人的に利用するケースが大半を占めるようである。翻訳者をタイプ別に見ると、フリーランスの翻訳者が圧倒的に無料サイトの利用が多いのに対し、翻訳会社内の翻訳者は無

料サイトよりも市販の PC 翻訳ソフトを利用する割合が多くなっている。

3.4. 機械翻訳への期待

翻訳者は機械翻訳にどのくらいの品質を期待しているのだろうか。表 3 に「翻訳ソフト/サイトにどのような品質を求めますか？」という質問の回答をまとめる。

表 3 機械翻訳が期待される品質

	フリーランス	翻訳会社	企業・官公庁
[1] 辞書代わり	2	3	3
[2] おおよその意味が分かること	6	0	3
[3] 翻訳の下訳として使えること	7	3	5

質問：翻訳ソフト/サイトにどのような品質を求めますか？

機械翻訳を翻訳の下訳として使いたいという翻訳者の期待が読みとれる一方で、機械翻訳が辞書引きレベルで使えれば十分と考えている翻訳者もかなりの割合でいることがわかる。

3.5. 機械翻訳の満足度

表 4 は機械翻訳の満足度を 5 段階で評価した結果をまとめたものである。「とても満足」と回答した人は皆無だったものの、18%（6 人）が「やや満足」と答えており、現在の翻訳システムでも使い方しだいで翻訳者の役に立つ場面があるようである。

表 4 機械翻訳全般の満足度

	フリーランス	翻訳会社	企業・官公庁
[1] とても不満	1	0	1
[2] やや不満	4	2	2
[3] どちらでもない	10	3	5
[4] やや満足	2	2	2
[5] とても満足	0	0	0

質問：翻訳ソフト/サイトの満足度は？（5 段階で評価）

満足度の根拠を具体的に知るために、機械翻訳ソフト/サイトの「品詞」と「機能」について別々に満足度を

調査した。以下にその結果を紹介する。

● 品質

表5を見ると、翻訳品質に満足していると答えた人（5人）の割合は不満であると答えた人（12人）の半分以下であり、翻訳者を満足させるためには、さらなる翻訳品質の向上が不可欠といえる。

表5 翻訳品質の満足度

	フリーランス	翻訳会社	企業・官公庁
[1]とても不満	4	0	1
[2]やや不満	3	1	3
[3]どちらでもない	8	4	5
[4]やや満足	1	2	1
[5]とても満足	1	0	0

質問: 翻訳ソフト/サイトの品質の満足度は? (5段階で評価)

● 機能（使い勝手）

表6から、機械翻訳は機能としては「満足」と答えた人（11人）と「不満」と答えた人（10人）の割合が拮抗していることがわかる。翻訳ソフト/サービスの使い勝手については品質ほど不満に感じている翻訳者はいないようである。

表6 翻訳機能の満足度

	フリーランス	翻訳会社	企業・官公庁
[1]とても不満	1	0	0
[2]やや不満	3	3	3
[3]どちらでもない	9	1	3
[4]やや満足	4	3	4
[5]とても満足	0	0	0

質問: 翻訳ソフト/サイトの機能の満足度は? (5段階で評価)

3.6. 翻訳メモリの利用状況

翻訳者を支援するツールとして、過去の訳例を蓄えておき、次回以降の翻訳で類似訳例を参照することができ「翻訳メモリ」がある。翻訳の現場では、機械翻訳よりも先に翻訳メモリの活用が進んでいると考えられる。表7は翻訳者による翻訳メモリの利用状況を調査した結果である。

表7 翻訳メモリの利用状況

	フリーランス	翻訳会社	企業・官公庁
[1] 現在使っている	6	6	2
[2] 以前使っていたことがあるが現在は使っていない	2	1	2
[3] 使ったことがない、翻訳メモリを知らない	9	0	7

質問: 翻訳メモリを使っていますか?

翻訳メモリを使ったことがない人が16人（46%）であり、現在使っている人の割合（14人（40%））よりも多くなっている。ただし、翻訳メモリの利用は、翻訳者のタイプによって大きく傾向が異なる。翻訳会社の中の翻訳者が1人を除いて翻訳メモリを使っていると回答しているのとは対照的に、フリーランスと企業・官公庁の翻訳者で翻訳メモリを使っている人は全体の3分の1にも満たない。このことから、翻訳会社の中において、翻訳メモリの活用が通常の業務プロセスに取り込まれていることが考えられる。

表8は、機械翻訳の種類と翻訳メモリの利用状況をクロス集計した結果である。これを見ると、二つの翻訳ツールの連携状況がある程度推定できる。

表8 翻訳メモリ利用状況と利用中機械翻訳の種類

	現在使用	過去に使用	未使用
[1] 無料の翻訳サイト	9	3	19
[2] 有料の翻訳サイト	1	1	2
[3] 自社または翻訳会社などで提供される翻訳システム(サービス)	2	0	1
[4] 市販のPC翻訳ソフト	10	2	0
[5] 業務では使っていない	5	3	5

翻訳メモリを使っていない翻訳者の多く（70%）は無料の翻訳サイトを利用する傾向にある。また、市販の翻訳ソフトを使っている翻訳者の大半（83%）は翻訳メモリのユーザでもあることがわかる。翻訳メモリまたは機械翻訳をどちらも使っていない翻訳者は8人（13%）の

みで、22 人（35%）は両方のツールを使っていると回答している。

3.7. 機械翻訳を利用するうえでの翻訳者の工夫

アンケートの中に、「翻訳業務の効率化のために工夫していることはありますか？」という設問を設け、回答は自由に文章を記述してもらった。27 人（42%）から回答があった。

27 人中 11 人が用語集の整備の重要性を指摘しており、9 人が翻訳メモリの活用を挙げていた。この 2 つが群を抜いて数の多い回答だった。これら以外では、

「原文の編集をしている」

「マクロやワイルドカード、正規表現などを活用して
原文を書き換えている」

というように原文修正の工夫を指摘するものや、

「自分で対訳表示をつくり原文と訳文を並べている」

「大きなモニタ 2 つを使用して、複数の翻訳関係のサ
イトを開いておく。」

といった、自分用の翻訳環境を作っているという回答があった。

4. 2012 年度アンケート結果との比較

4 節では 2013 年度アンケートの結果について説明した。冒頭で述べたとおり 2012 年度にも同じ内容のアンケートを実施しているが、調査結果は 2013 年度の結果とほぼ同じ傾向を示している。2 回のアンケートの間隔が 1 年と短いこともあり、下記の機械翻訳への期待に関する項目以外で特筆すべき経年変化を発見することはできなかった。

「機械翻訳にどのような品質を求めていますか？」という問いについては、やや異なった傾向を示した。2012 年と 2013 年を比較すると、「辞書代わり」と回答した人が 17%から 28%に増えた一方で、「下訳として使えること」と回答した人は 40%から 34%に減少した。現実的な利用をしようとする意識の広がりが見られるのか、機械翻訳への期待感が下がっているのか、この結果だけでは明らかにならない。

5. まとめ

2013 年度の第 32 回 JTF 翻訳祭で実施したアンケート結果を中心に、実務翻訳における機械翻訳の利用実態についてまとめた。

これまで実務翻訳の現場で機械翻訳の活用は難しいと思われてきたが、今回の調査で、翻訳者の 4 割以上が「毎日」機械翻訳を使っていることが判明した。しかしながら、毎日使っている機械翻訳の多くはインターネットで無償提供されている翻訳サイトであり、翻訳ベンダーが長年取り組んできた翻訳パッケージの利用はまだまだ浸透していないのが現実である。無料翻訳サイトの具体的な活用法については、今回の調査では踏み込むことができなかったが、次回アンケートなどを利用して調査を行う予定である。

用語集の整備と翻訳メモリの活用については、すでに実践して効果をあげている翻訳者もいるようである。機械翻訳を用語集や翻訳メモリと連携させて効果を引き出すことができるかどうか、実務翻訳における機械翻訳普及の鍵になると思われる。

課題調査委員会では、今後も日本翻訳連盟との連携を深めてイベント共催などを進めながら、実務翻訳者の実態調査を継続するとともに、機械翻訳の効率的な活用方について翻訳者に向けた情報発信を行っていきたいと考えている。

シンプルな用語集形式 UTX とその活用 ービッグ データと「グッド データ」の関連性ー

機械翻訳課題調査委員会 WG3

山本ゆうじ[†]

[†] 秋桜舎・代表（住所略）／アジア太平洋機械翻訳協会（AAMT）事務局（住所略）

あらまし 用語集形式 UTX の概要と、UTX がどのように活用されうるかを紹介する。特に、自然言語処理のビッグ データ（コーパス）から少数の良質な「グッド データ」を抽出し、正規化し、継続して管理・共有・再利用する際に、UTX がどのように用語集の正確性、的確性、および信頼性を向上させるかを示す。

キーワード UTX、TBX、用語集、辞書、ビッグデータ、機械翻訳、RBMT、コーパス

UTX, a Simple Glossary Format, and Its Applications ーHow "Good Data" Interact with Big Dataー

Yamamoto Yuji[†]

[†] CosmosHouse, Representative (address omitted) / Asia-Pacific Association for Machine Translation
(address omitted)

Abstract This presentation introduces the glossary format UTX and how it can be used. Especially it demonstrates how UTX can improve accuracy, appropriateness, and reliability of a glossary when extracting, normalizing, continuously managing, sharing, and reusing selected, fine-tuned "good data" out of linguistic big data (corpus).

Keyword UTX, TBX, glossary, dictionary, big data, machine translation, RBMT, corpus

1. はじめに

本発表では、用語集形式 UTX と、それがどのように活用されうるかを紹介する。自然言語処理のビッグ データ（コーパス）から少数の良質な「グッド データ」を抽出し、正規化し、継続して管理・共有・再利用する際に、UTX がどのように用語集の正確性、的確性、および信頼性を向上させるかを示す。

2. 体系的翻訳のアプローチ

UTX 自体と、UTX が必要とされる環境を説明する前に、体系的翻訳について概説する。

日本語のビジネス文書では、表記や用語の面で多くの乱れがある。このような乱れは、ビジネス文書の理解度、検索性、発信力、ウェブプレゼンス（ウェブ上での検索などを通じた存在感）を弱め、企業の信頼性を下げることもある。また、企業／組織向けに商業ベースで行われる

規模の翻訳では、翻訳そのものの正確性が求められることがある。このような場合は、「体系的翻訳」が必要となる。体系的翻訳では、出版翻訳や研究者が個人で行う、その場限りの翻訳などと異なり、以下の特徴がある。

- 翻訳文書量が多い
- 複数の翻訳者・チェッカーが協働する
- （依頼元企業などから）用語と表記の統一が求められる
- 上記の理由から体系的翻訳用の専門的ツール（翻訳メモリー ツール、翻訳ソフト、用語ツール、校正ツールなど）を使用する
- 専門的ツールを使用して、可能な限りの言語資源を蓄積し、再利用する

「翻訳メモリー ツール」とは、主に人間翻訳者の翻訳を蓄積・共有・再利用するツールである。「用語ツール」とは、用語データベース（用語ベース）を、作成・検索・管理するツールで

ある。「校正ツール」とは、単独、または他のツールに組み込まれた形で、文書内の表記や文法の問題をチェックするツールである。体系的翻訳のアプローチは、ICT 業界では定着しつつあるが、それ以外の業界では、方法論やツールが十分に普及していないことも多い。

体系的翻訳では、スタイルガイドと用語集が使われるが、これらは相互に関連している。スタイルガイドとは、「企業や組織で文書を作成するときの表記基準、表記の規則」^[1]である。一例として、JTF（日本翻訳連盟）が作成した翻訳用スタイルガイド^[2]が無償で公開されている。用語集とは、一般には、特定分野での用語を列挙したものである。用語の定義や解説を含む場合もあるが、対訳用語集では、定義や解説はなく、原語を見出しとして、対応する訳語を記載することもある。

体系的翻訳では、スタイルガイドと用語集を正しく使うことで、的確な用語および表記を選び、統一できる。このような統制は、「企業に入ってくる文書」よりも、「企業内文書」と「企業が発信する文書」で重要である。

3. 用語集形式 UTX

UTX (Universal Terminology eXchange) は、アジア太平洋機械翻訳協会 (AAMT: Asia-Pacific Association for Machine Translation) が策定した、シンプルかつオープンで汎用的な用語集形式である。

UTX 形式の用語集を翻訳ソフトのユーザー辞書とすることで、わずかな語数の UTX 辞書でも複数の翻訳ソフトの翻訳精度を改善できることが示されている (Bond et al. 2009)^[3]、(Okura et al. 2011)^[4]。

UTX の公式ページ^[5]より、詳細な UTX 1.11 の仕様、サンプル用語集、クイックガイドなどが日本語と英語でダウンロードできる。

UTX には、以下の利点がある。

- シンプルな表形式
 - 作りやすく管理しやすい
- 相互変換して共有・再利用
 - 翻訳ソフトや用語ベース用の形式と相互変換できる
- 双方向用語集
 - たとえば英日と日英を1つの用語集で管理できる

UTX は、タブ区切りテキストであり、シンプルな構造であるため、特殊なツールを使用せず

図 1: UTX の例 (UTX 1.20 での双方向用語ステータス)

AAMT用語辞書 (UTX 1.20)英日双方向.utx.xlsx - Excel					
ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 開発 アドイン ATOK拡張ツール ACROBAT					
G12 : X ✓ fx 5					
A	B	C	D	E	F
1	#UTX 1.20beta; en-US/ja-JP; 2013-11-13T11:00:00Z+09:00; copyright: AAMT (2013); license: CC-BY 3.0				
2	#description: This is a sample dictionary for AAMT-related terminology. It is not an official dictionary.				
3	/ この辞書はサンプル用のAAMT関連の用語辞書です。AAMTの公式の辞書ではありません。				
4	#src	tgt	src:pos	term status:src	term status:tgt
5	entry	エントリー	noun	approved	forbidden
6	entry	項目	noun	approved	approved
7	dictionary	ディクショナリー	noun	approved	forbidden
8	dictionary	辞書	noun	approved	approved
9	merge	マージする	verb	approved	forbidden
10	merge	統合する	verb	approved	approved
11	optional	オプション	adjective	approved	forbidden
12	optional	省略可能	adjective	approved	approved
13	unique	ユニークな	adjective	approved	forbidden
14	unique	一意の	adjective	approved	approved
15	blank	ブランク	adjective	approved	forbidden
16	blank	空白	adjective	approved	approved
17	glossary	グロッサリー	noun	approved	forbidden
18	glossary	用語集	noun	approved	approved
19	provisional term	暫定語	noun	approved	approved
20	provisional word	暫定語	noun	non-standard	
21	approved term	承認語	noun	approved	approved
22	approved word	承認語	noun	non-standard	
23	non-standard term	非標準語	noun	approved	approved
53	license	使用許諾	noun		provisional
54	line comment	行コメント	noun		

に、表計算ソフトで編集・管理できる(図 1)。しかし、編集が容易な表計算ソフトで扱えるというだけで、データの構造が標準化・規格化されていないと、用語データとして活用することは困難である。標準化された用語集仕様としては、ISO 規格 TBX^[6]があるが、仕様が複雑で、特に日本では広く普及していないのが現状である。UTX は、柔軟な拡張性を保ちつつ、それぞれのツールで実際に活用できる情報を利用できる構造である。

また、UTX 用語集は、前述の TBX や各種翻訳ソフトのユーザー辞書形式と、相互に変換できる。

UTX は、日本発の国際標準規格となることを目指している。UTX プロジェクトでは、公開中のすべての文書が日本語と英語の両方で用意されている。UTX チーム メンバーが、言語処理学会(2008 および 2011 年、東京)、IWIC2009(ハワイ)、MT Summit 2011(アモイ)、LISA China Focus(2009 年、北京)、LISA Open Standards Summit(2011 年、ボストン)など、世界各国で UTX の成果を発表している。さらに、英文翻訳業界誌 *MultiLingual*^[7]、"GALaxy Newsletter"^[8]や、国内では日本特許情報機構(Japio)の『Japio Year Book 2012』^[9]、『Japio Year Book 2013』^[10]などでも UTX を紹介している。

このような普及活動の結果、現在、各国の開発者により、UTX 形式が採用されつつある。一例として、翻訳業界で広く使われている用語チェック ツール Xbench^[11]のバージョン 3.0 で UTX がサポートされている。UTX 仕様は無償で公開されており、このようなツールへの組み込みは自由に行うことができる。AAMT でも、オープン ソースの UTX 変換ツールを 2013 年 11 月現在、開発中である。

また、サンプル用語集として、2010 年、法律と医学分野で UTX 形式の英日・日英対訳辞書(それぞれ 5451 語、27126 語)を作成して、無償で一般公開した。これらは UTX 公式ページからダウンロード可能^[12]である。

4. 用語ステータス

UTX では、各項目に以下の用語ステータスを付与することで、用語を管理できる。

- 暫定 (provisional) 「未承認の用語」

- 承認 (approved) 「使うべき用語」
- 非標準 (non-standard) 「使われることがある用語」
- 禁止 (forbidden) 「使うべきでない用語」

用語ステータスについては、大倉らが以下のように示している(以下の引用では「辞書」となっているが、これは「用語集」と読み替えられる)。

- 「
- 暫定(provisional): その項目がまだ辞書管理者(後述)によって承認されていないことを意味する。辞書管理者は、可能な限り、暫定語を以下で説明する「禁止」「承認」「非標準」のいずれかに決定するか、削除することが望ましい。
 - 禁止(forbidden): 用語管理の観点から、項目の中に使ってはならない訳語が含まれていることを意味する。翻訳ツールが複数の辞書間の優先順位を適切に扱わない場合は、異なる分野別辞書との競合を避けるため、目標言語の用語を抑制する必要があることもある。例えば ICT の文脈では、英語の用語である "window" は日本語の単語である「窓」と翻訳されることは極めて少ない。この「窓」という訳語を機械翻訳システムが適切に扱えないのであれば、この訳語は明示的に抑制する必要があるだろう。禁止用語は、翻訳ツールの外で用語チェックに使用するために抽出することができる。
 - 承認(accepted): 項目が辞書管理者により承認されていることを意味する。その用語が必ず使われなければならないことを示す。何か明らかな理由がある場合、起点言語の 1 つの用語に対して複数の目標言語の用語を割り当てることは可能であるが、用語ステータスが承認となるのはそのうちの 1 項目だけである。承認語は常に双方向性をもつ。つまり、辞書に定義されている翻訳方向と反対方向にして使える。同一概念(後述)に属する複数の項目が存在する場合、翻訳方向を逆にしたときに、唯一有効な項目である。
 - 非標準(non-standard): 1 つの項目に対し、1 つ以上の非標準の起点言語の用語があること

を意味する。非標準用語は、単に起点言語の用語の異形に対応するためにのみ許可される。非標準用語は目標言語の用語としては使われてはならない。UTX 辞書が翻訳のためではなく、文書のオーサリングのための用語集として使用される場合、非標準用語は用語として使ってはならない。非標準用語が辞書に登録されているのは、起点言語の文書の筆者が承認されていない不適切な語を使用した場合でも、自動翻訳が翻訳できるようにするためにすぎないからである。」(大倉他 2011) ^[13]

表 1 は、UTX 1.11 での用語ステータスの例である。

表 1：用語ステータス (UTX 1.11)

原語	訳語	用語ステータス
outlet	コンセント	approved
outlet	アウトレット	forbidden
power point	コンセント	non-standard
PowerPoint	PowerPoint	approved
plugin	プラグイン	approved
plug-in	プラグイン	non-standard

なお、用語ステータスのより詳細な情報は、UTX 仕様^[14]に記載されている。

2013 年 11 月現在、UTX 仕様の現行バージョンは 1.11 である（前述の引用は UTX 1.10 についてであるが、同等）。UTX 1.11 では、用語ステータスは、原語と訳語の 1 つのペアに対して 1 つ付与される。また、この用語ステータスにより、UTX 1.11 は、二言語双方向用語集をサポートする。二言語双方向とは、翻訳方向を反転させて使用できることを示す。たとえば、基本的には英日である対訳用語集を、日英として使用する場合である。しかし、UTX 1.11 では、正方向で使用するが主で、逆方向で使用することは二次的という位置付けであった。これは、実際の翻訳業務では翻訳方向を反転して使用する

状況は比較的少なく、正方向のみで使用する場合に特化してシンプルにしたほうがよいと推測されたからである。

現在検討中のバージョン 1.20 では、多言語双方向用語集の仕様を策定中である。UTX 1.20 では、用語ステータスを原語と訳語の両方に付けることで、双方向とする予定である。さらに、起点言語（原語の言語）と、複数の目標言語（訳語の言語）間での翻訳方向の反転が可能になる。やや複雑にはなるものの、この「双方向用語ステータス」（図 1 を参照）により、原語と訳語の翻訳方向を反転させたとき、原語と訳語をより等しく扱えるようになる。双方向用語ステータスについては、今後公開予定の UTX 1.20 仕様に記載される。

UTX では、実際の翻訳業務での必要性から、この用語ステータスを使用した「用語の品質管理」という観点があることが特徴である。コーパスから生成されただけの「用語集」は品質が低く、多くの問題点が含まれることがよって指摘されている (Melby 2012) ^[15]。UTX 形式では、各種専門分野の現場の専門知識に基づいて用語データを収集し、専門家（用語集管理者）が用語の品質管理を行うことで、高品質な用語集をすばやく作成できる。さらに用語ステータスにより、用語の品質を保証し、信頼性を持たせることにより、既存の用語データを再利用・共有しやすくできる。だいたいの意味が分かればいい「概訳」と異なり、企業用途翻訳では、専門用語が確実に使用されることが重要である。

統計機械翻訳は、機械翻訳の一種だが、文法解析はせず、統計的パターンに基づいて翻訳する方式である。統計機械翻訳では、ビッグデータが活用されている。しかし、英仏、日韓など類似言語間ではともかく、日英のような異種言語間では期待するほどの相関関係が得られず、「ビッグなだけのデータ」になりかねない。特に、用語が的確に訳されているかという点では、十分な信頼性が得られない。これに対して、「使うべき用語」、「使うべきでない用語」などが明確に区別され、品質管理された用語集は、「グッドデータ」といえる。

5. まとめ

用語集は、自然言語ビッグデータから「グッドデータ」を用語抽出して作成されることがあ

る。その一方で、用語集が既存のデータを統一し規則性をもたらすことで、データの品質を高められる。つまり、作成した用語集は、(スタイルガイドも併用して)言語資源を評価するためにも使える。その意味で、用語集は、自然言語ビッグデータを制御する「手綱」だといえる。UTXは、実用的な用語管理の枠組みであり、各種専門分野の、現場の専門知識を活用してビッグデータの品質を高めることができる。

UTXは、たとえば、ルールベース機械翻訳でのユーザー辞書として使用することに加えて、統計機械翻訳での用語チェックにも活用できる。ルールベース機械翻訳(RBMT: rule-based machine translation)は、機械翻訳の一種で、文法解析に基づく方式である。信頼できる、しっかりした用語集があれば、統計機械翻訳の用語の不確かさという欠点を補うことができる。

UTXは、一例として特許庁機械翻訳システムでも使われている。日中機械翻訳の調査では、約80万語の対訳辞書データが作成された。特に複合名詞の翻訳精度が大きく向上したと報告書に記載されている。

「複合名詞ほどRBMT1で誤訳となる割合が高く、ほぼ単名詞とみなせる中国語2文字以下の正訳率が71%なのに対し、複合名詞とみなせる中国語3文字以上の用語の正訳率は49%であった。RBMT2ではこれがそれぞれ81%、73%に改善しており、特に複合名詞の翻訳精度に大きな向上が見られた。」^[16]

上記のRBMT1は辞書データを機械翻訳辞書に追加する前の環境、RBMT2は追加後の環境を示す。

今後、UTXに関する研究課題としては、より効果的な用語抽出の支援が挙げられる。ビッグデータから頻度情報のみに基づいて用語を抽出するのではなく、専門家の承認を経た用語ステータスを付与することで、用語集の信頼性を高められる。このような仕組みがあれば、商業翻訳での用語をよりの確に翻訳することができる。

また、学会論文や特許の出願で、UTX形式でのシンプルな対訳用語データの提出を求めるようにすれば、高精度な用語データベースを容易に構築できよう。

シンプルでありながら、実用的で一貫性のある用語品質管理の情報を記述できるUTXにより、高品質な用語集を共有・再利用する環境が

今後整うことが期待される。

本文章は、2013年12月4日にグリー株式会社で行われた、電子情報通信学会(IEICE)言語理解とコミュニケーション研究会(NLC)主宰による、第五回集合知シンポジウム
<<http://www.ieice.org/~nlc/cs2013.html>>提出論文である。『電子情報通信学会技術研究報告』信学技報 Vol. 113 No. 338 より一部を修正して転載した。

Copyright ©2013 by IEICE

- [1] 山本ゆうじ, 『IT 時代の実務日本語スタイルブック——書きやすく、読みやすい電子文書の作文技法』, p. 53, ベレ出版, 2012.
- [2] JTF, 『JTF 日本語標準スタイルガイド (翻訳用)』, 2013.
<http://www.jtf.jp/jp/style_guide/styleguide_top.html>
- [3] Francis Bond et al., "Sharing User Dictionaries Across Multiple Systems with UTX-S," *Proceedings of the Second International Workshop on Intercultural Collaboration (IWIC2009)*, Stanford, 2009.
<<http://web.mysites.ntu.edu.sg/fcbond/open/pubs/2009-iwic-utx.pdf>>
- [4] Seiji Okura et al., "UTX 1.11, a Simple and Open User Dictionary/Terminology Standard, and its Effectiveness with Multiple MT Systems," pp. 587-94, MT Summit XIII: the Thirteenth Machine Translation Summit, September 2011, Xiamen, , 2011.
<<http://www.mt-archive.info/MTS-2011-Okura.pdf>>
- [5] AAMT, 「UTX 公式ページ」
<<http://www.aamt.info/japanese/utx/>>
- [6] LISA, "Term Base eXchange (TBX)"
<<http://www.ttt.org/oscarStandards/tbx/>>
- [7] Yamamoto Yuji, "How to create glossaries in UTX," *MultiLingual*, June 2011.
- [8] Yamamoto Yuji, "UTX (universal terminology eXchange): a simple terminology format," "GALaxy Newsletter," Q2 2013.
<<http://www.gala-global.org/node/82208>>
- [9] 山本ゆうじ, 「実務日本語と UTX から特許明細書の改善を考える」, 『Japio Year Book 2012』, 2012.
<http://www.japio.or.jp/00yearbook/files/2012book/12_5_04.pdf>
- [10] 山本ゆうじ, 「日本語文章の課題と前編集手法—用語集形式 UTX と実務日本語の観点から」, 『Japio Year Book 2013』, 2013.
- [11] Xbench.net, <<http://www.xbench.net/>>
- [12] AAMT, 「UTX 公式ページ」
<<http://www.aamt.info/japanese/utx/>>
- [13] 大倉清司他, 「翻訳支援のためのシンプルでオープンな辞書仕様 UTX-Simple 1.10」, 「言語処理学会 第 17 回年次大会 発表論文集」, 2011.
- [14] AAMT, 『UTX 1.11 仕様書』, 2011.
<<http://www.aamt.info/japanese/utx/#download>>
- [15] Alan K. Melby, "Terminology in the age of multilingual corpora," *The Journal of Specialised Translation*, p. 12, 2012.
<http://www.ttt.org/tbx/art_melby-JoSTransNum18.pdf>
- [16] 特許庁, 『平成 24 年度 中国特許文献の機械翻訳のための中日辞書整備及び機械翻訳性能向上に関する調査 調査報告書 概要版』, p. 10, 2013.
<<http://www.jpo.go.jp/shiryoutoushin/chous>>

AAMT会員のひろば

AAMT 会員の新たな交流の場を AAMT Journal 誌面上で提供するべくスタートいたしました「AAMT 会員のひろば」、会員の皆さまのご助力をいただきまして、第一回の No.41 のスタートから第七回を迎えることができました。今号では、法人会員一社、個人会員二名の皆さまからのご寄稿をいただいております。

独自のお取り組みのご紹介、機械翻訳研究への提言、AAMT の活動へのご要望など、今回も貴重なご意見をお寄せいただきました。

AAMT Journal では今後も引き続き、会員の皆さまからのご寄稿を心よりお待ちしております。

ご寄稿・お問い合わせは AAMT 事務局(E-mail: AAMT-info@AAMT.info)まで宜しくお願いいたします。

個人会員（敬称略・50 音順）

会員名

Alun Simpson

Freelance Technical Translator, Japanese to English

simpson_translations@me.com

自己紹介

I am from Cardiff in Wales. I graduated from Cambridge University, England, in 1999 with a First Class degree in history, including Japanese history from the Meiji Restoration to the present day. A couple of months after graduating I came to Japan on the JET Programme, teaching at five junior high schools in Gifu City, where I fell in love with Japan (and snowboarding!).

After JET, I worked as an in-house translator at a company in Nagoya before becoming freelance in 2006. I specialize in technical translation, especially software, computers, consumer electronics and automobiles.

I came to Tokyo in 2008, when I shocked all my Nagoya friends by becoming a passionate Yakult Swallows fan! I also like soccer (Cardiff City), travelling, art, restaurants and movies, as well as making sure I know everything about the latest gadgets and technologies.

MT/翻訳とのかかわり

MTおよび翻訳業界に期待すること

Like many other technical translators, I use translation memory software (CAT tools) in most of my work. When I started translating in 2003, Trados was already widely used across the industry and one of my first tasks was to learn how to use it. CAT tools have always been an integral part of my workflow, and I find it hard to imagine doing technical translation without them. For this reason, I have always had an open mind towards new translation technologies.

Despite the understandable feelings of anxiety towards new technologies felt by some translators and LSPs, I think that this is a very exciting time to be involved in our industry. The translation industry continues to grow and grow, as one of the industries favored most by globalization. Translation was one of the few industries to continue growing all through the Great Recession, and trends show it will continue to grow in the years ahead. Add this to the internet revolution, crowd and cloud computing and the recent breakthroughs in machine translation (MT) quality, and I think we are working at the frontier of many of the world's most exciting new trends. I am certain that over the next 5 to 10 years we will see many innovative business solutions that meld different combinations of the crowd, social media and MT. Having said that, from the perspective of my day job as a freelance translator, I do not think that the MT quality in my language combination is good enough yet to be useful. Japanese to English is one of the hardest language directions for MT. Even so, it will not take much more improvement before MT becomes a useful tool for professional translations, as it already is in other language combinations.

I think a key issue is how to integrate CAT tools and MT in an efficient way that actually helps translators. For an integrated tool to be useful to translators, it is important to intelligently rank and sort fuzzy match segments in the translation memory and those created by the MT. To do this, I think it is important for developers of MT to involve working translators from an early stage to get feedback about what is useful and what is not.

AAMTへの要望

AAMT is great place for organizations involved in MT and other interested individuals (like myself) to communicate and exchange ideas. Too often, communication between engineers and translators is poor, with translators seeing all new technology as a threat, and engineers not fully understanding the needs of translators. AAMT helps to break down these barriers.

I would like AAMT to perhaps offer more educational seminars or courses, to give members the opportunity to expand their knowledge of MT and other current technologies.

協会活動報告

(2013 年 9 月～2014 年 3 月)

機械翻訳課題調査委員会

【委員会】

2013 年 9 月 6 日 (2013 年度 第 5 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
 - ・メルマガの手順確認
 - ・AAMT/Japio 特許翻訳研究会との連携について状況報告
 - ・中日基本文テストセットの構築について
 - ・特許用テストセット構築について
 - ・ジャーナル連載・講演会開催など
 - ・UTX 変換ツール開発について
 - ・UTXVer1.20 仕様
- ③ 活動内容の報告 (各 WG から)
- ④ 活動内容についての議論
 - ・第 23 回 JTF 翻訳祭(2013/11/27)パネルディスカッションに関する検討
 - ・UTX 変換ツール開発に関する費用の検討
 - ・その他
- ⑤ まとめと次回委員会について

2013 年 10 月 18 日 (2013 年度 第 6 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
 - ・中日基本文テストセットの構築について
 - ・特許用テストセット構築について
 - ・ジャーナル連載・講演会開催など
 - ・第 23 回 JTF 翻訳祭について
 - ・UTX 変換ツール要求仕様書検討
 - ・第 5 回集合知シンポジウムについて
- ③ 活動内容の報告 (各 WG から)
 - ・AAMT/Japio 特許翻訳研究会(10/11)の参加報告
 - ・活動内容についての議論
 - ・第 23 回 JTF 翻訳祭について
 - ・その他
- ④ まとめと次回委員会について

2013 年 11 月 15 日 (2013 年度 第 7 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
 - ・ 中日基本文テストセットの構築について
 - ・ 特許用テストセット構築について
 - ・ 第 23 回 JTF 翻訳祭 (AAMT 翻訳プラザ出展、アンケート) について
 - ・ 第 5 回集合知シンポジウムについて
 - ・ UTX1.20 仕様検討
- ③ 活動内容の報告 (各 WG から)
- ④ 活動内容についての議論
 - ・ 第 23 回 JTF 翻訳祭について
 - ・ 言語処理学会第 20 回年次大会のテーマセッション/ワークショップへの参加について
- ⑤ まとめと次回委員会について
- ⑥ AAMT セッション講演について講演者打ち合わせ

2013 年 12 月 6 日 (2013 年度 第 8 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
 - ・ 中日基本文テストセットの構築について
 - ・ 特許用テストセット構築について
 - ・ 人工知能研究振興財団助成金について
 - ・ UTX 変換ツールについて
 - ・ UTX 仕様検討
- ③ 活動内容の報告 (各 WG から)
 - ・ 集合知シンポジウムの参加報告
- ④ 全体会議
 - ・ 各 WG 検討内容報告
 - ・ 第 23 回 JTF 翻訳祭について(報告)
 - ・ AAMT 企画講演会について
 - ・ その他
- ⑤ まとめと次回委員会について

2014 年 1 月 17 日 (2013 年度 第 9 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)

- ・中日基本文テストセットの構築について
- ・特許用テストセット構築について
- ・AAMT ジャーナル記事執筆について
- ・AdWords の利用の検討
- ・UTX 仕様検討
- ・今年度成果と来年度計画

③ 全体会議

- ・各 WG 検討内容報告
- ・JTF 翻訳祭について(振り返り)
- ・AAMT ジャーナル記事執筆について
- ・AAMT 総会併設企画について
- ・その他

④ まとめと次回委員会について

2014 年 2 月 14 日 (2013 年度 第 10 回)

① 前回委員会の議事録の確認

② 各 WG の活動について (各 WG に分かれて議論)

- ・中日基本文テストセットの構築について
- ・特許用テストセット構築について
- ・言語処理学会の発表について
- ・AdWords について
- ・UTX 仕様検討の方針
- ・今年度成果と来年度計画

③ 全体会議

- ・各 WG 検討内容報告
- ・AAMT 総会併設企画について
- ・その他

④ まとめと次回委員会について

2014 年 3 月 7 日 (2013 年度 第 11 回)

① 前回委員会の議事録の確認

② 各 WG の活動について (各 WG に分かれて議論)

- ・中日基本文テストセットの構築について
- ・AAMT/Japio 特許翻訳研究会評価拡大部会活動報告
- ・言語処理学会の発表について
- ・AAMT ジャーナルの会員からの投稿の扱い
- ・AdWords について

- ・変換ツールについて
 - ・今年度成果と来年度計画
- 来年度の活動計画に必要な予算の見積もり

③ 全体会議

- ・AAMT 総会併設企画について
- ・新しい AAMT ページのレビュー
- ・その他仕様

④ まとめと次回委員会について

【対外活動】

① 第 23 回 J T F 翻訳祭 (2013/11/27) への参加

- ・翻訳プラザへの出展。
課題評価委員会の活動紹介およびMTに関するアンケートを行った。
- ・セッション参加：「機械翻訳による競争戦略--機械翻訳で差別化できるか---」
J T F 様よりパネルディスカッションの場をいただき、上記のテーマで講演および会場からの質問もまじえて討論を行った。(モデレーター：長瀬友樹(富士通研究所)、パネラー：秋元圭(クロスランゲージアールアンドディ)、目次由美子 (シュタールジャパン)、山本ゆうじ(秋桜舎))。

② 言語処理学会第 20 回年次大会(2014/3/19)への参加

- ・ポスターセッション：「実務翻訳における機械翻訳の利用に関する調査報告」
J T F 翻訳祭にて機械翻訳の利用状況に関するアンケートを行った結果について分析報告した。
(長瀬 友樹(富士通研究所)、小谷 克則(関西外語大学)、工藤 竜広(サンフレア)、佐久間 みゆき(富士通)、 秋葉 泰弘(NTT メディアインテリジェンス研究所))

編集委員会

2013 年 11 月 5 日（2013 年度 第 2 回）

- ① AAMT Journal No.55 の記事の執筆依頼計画について
- ② AAMT Journal No.55 のスケジュールについて

2014 年 2 月 26 日（2013 年度 第 3 回）

- ③ AAMT Journal No.55 の記事状況について
- ④ AAMT Journal No.56 の執筆依頼計画について

インターネット WG

- ⑤ 新 AAMT ホームページ公開に向けた更新体制の立ち上げ
- ⑥ 既存 AAMT ホームページの更新
- ⑦ AAMT Forum メールマガジンのコンテンツ作成・配信業務の、課題調査委員会への移管作業（完了）
- ⑧ 他委員会の AAMT ホームページ利用に関するサポート
- ⑨ その他、事務局ネットワークのインフラ管理全般

AAMT/Japio 特許翻訳研究会

2013 年 10 月 11 日(金) (2013 年度 第 4 回)

1. 委員メンバーについて (NHK 後藤様、NICT 隅田様)
2. 前回議事録の確認
3. MT Summit の出張報告
4. 機械的な評価手法についての活動方針
5. NTT の機械翻訳システムについて紹介(須藤様)
6. Coling 14 の Workshop
7. 次回の開催について
 - 開催の日時
 - 主な議題

2013 年 11 月 29 日(金) (2013 年度 第 5 回)

1. 前回議事録の確認
2. 平成 25 年度報告書について
3. 特許・情報フェア&コンファレンスおよび Japio YEAR BOOK 2013 Web 公開について
4. 次回の開催について
 - 開催の日時
 - 主な議題

2014 年 1 月 24 日(金) (2013 年度 第 6 回)

1. 前回議事録の確認
2. 「第 3 回拡大評価部会の報告」
3. 報告書内容概要
4. 来年度の研究費について (事務局)
5. 次回の開催について
 - 開催の日時
 - 主な議題

2014 年 3 月 7 日(金) (2013 年度 第 7 回)

1. 前回議事録の確認
2. NICT 渡辺様講演「SMT の最前線」
3. 来年度について
 - 開催の日時
 - 主な議題

AAMT/Japio 特許翻訳研究会 拡大評価部会

2013 年 10 月 11 日（金）（2013 年度 第 2 回）

1. 前回議事録の確認
2. 進捗報告
 - 2.1 自動評価
 - 2.2 人手評価
 - 2.3 テストセット
3. その他

2014 年 1 月 24 日（金）（2013 年度 第 3 回）

1. 前回議事録の確認
2. 報告書内容概要(各グループより)
3. その他

・ 5th Workshop on Patent Translation （研究会主催）

2013 年 9 月 2 日(月) Acropolis Conference Centre, Nice, France

（MT Summit XIV 併設 Workshop として開催）

AAMT ジャーナル編集委員会委員長
筑波大学 システム情報系 知能機能工学域
宇津呂 武仁

AAMT ジャーナル 55 号をお送りします。

今号の巻頭言は、東芝ソリューション株式会社 熊野明様より、御寄稿を頂きました。

AAMT の対外活動の一環として、機械翻訳課題調査委員会が第 23 回 JTF 翻訳祭にて行ったパネルディスカッション「機械翻訳による競争戦略－機械翻訳で差別化できるか－」の内容を掲載いたしました。

また、理事の坂本義行先生からは、機械翻訳の歴史に関する貴重な御寄稿の第二弾を頂きました。

新たなサービス・製品の紹介として、日本特許情報機構(Japio)様より、Japio 世界特許情報検索サービス (Japio-GPG) の紹介を、また、高電社様より、Android スマートフォン向け文字入力システムと連携したクラウド型翻訳アプリ「翻訳キーボード」、および、特許用語・専門用語を標準搭載した韓国語翻訳ソフト「J ソウル 9」の紹介を、それぞれ御寄稿いただきました。

AAMT 内の活動報告として、機械翻訳課題調査委員会から、JTF 翻訳祭において実施した「実務翻訳における機械翻訳の利用に関するアンケート」の結果報告、用語集形式 UTX の最新情報、および、「AAMT フォーラム メールマガジンバックナンバー」を掲載いたしました。

その他、「AAMT 会員のひろば」の企画におきましては、個人会員の紹介文 1 件を掲載しました。

AAMT

Asia-Pacific Association for Machine Translation

AAMT 入会のご案内

AAMT は、機械翻訳の発展を目的として、機械翻訳の研究者、開発者、製造者、利用者が集まった任意の組織です。委員会による定期的な調査研究をはじめ、機関誌の発行、シンポジウム、セミナー等各イベントの開催など幅広く活動を行っています。
機械翻訳にご関心のあるすべての方にご入会をお勧めします。

* * AAMT 会員の特典 * *

1.AAMT Journal の購読ができます。

会員には、機関誌である AAMT Journal（年 2～3 回発刊予定）が送付されます。購読料は年会費に含まれています。

2.機械翻訳関連の最新情報をメールでお届け

会員専用メーリングリストで、最新の機械翻訳関連の情報をお届けします。

MT 新製品、新サービスの紹介、国際会議、シンポジウムのお知らせ、WEB での MT 関連記事の紹介など盛りだくさんです。

3. AAMT が組織する委員会や調査活動に参加し、機械翻訳や翻訳に関心のある方との交流を深め、知見を広めることができます。

機械翻訳に関する言語資料の調査、広報、標準化活動に参加したり、AAMT Journal や会員専用メーリングリストで、自社製品、サービスの紹介を行うことができます。

4.関連機関の主催する国際会議に参加できます。

IAMT の主催で隔年開催される MT Summit をはじめ、AAMT、AMTA*、EAMT** の主催する会議やワークショップに参加できます。

AMTA* : Association for Machine Translation in the Americas

EAMT** : European Association for Machine Translation

年会費は以下の通りです。

法人会員：入会金	1 口	10,000 円	年会費	1 口	50,000 円
個人会員：入会金		1,000 円	年会費		5,000 円（学生は学生会費 1,000 円）

ご関心のある方は、事務局までお問い合わせください。

アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

電子メール：aamt-info@aamt.info

AAMT

Asia-Pacific Association for Machine Translation

入会申込書

以下の通り、アジア太平洋機械翻訳協会の会員申し込みを致します。

申込日 20 年 月 日

氏名（ローマ字）	
氏名（漢字）	
電話番号	
メールアドレス	
所属先	
所属先住所	〒
種別	☐ ユーザ ☐ 研究開発者 ☐ その他
機械翻訳に関するお知らせメールの配信	☐ 希望する ☐ 希望しない

コメント

--

アジア太平洋機械翻訳協会(AAMT)

総会のご案内

アジア太平洋機械翻訳協会

会 長 中岩 浩巳

アジア太平洋機械翻訳協会では、総会および報告会・懇親会などの関連行事を下記の通り開催いたします。

会員の皆様には、以下の日時を予定してぜひご参加いただければ幸いです。

また関連行事は会員・非会員の方々のいずれも参加できます。
ぜひ多数ご参加くださいますよう、ご案内申し上げます。
プログラムなど詳細につきましては、後日MLおよび書面にてご案内いたします。

1. 日 時 2014 年 6 月 17 日 (火) 午後

2. 会 場:ホテルアジュール竹芝

アクセス: <http://www.hotel-azur.com>

2.1)総会および関連行事

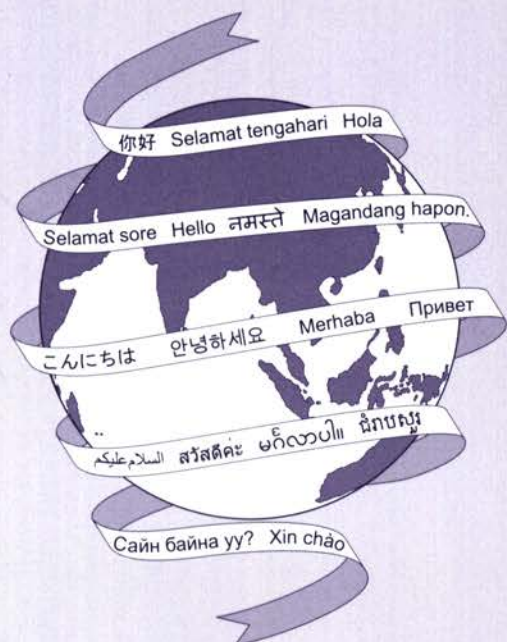
同ホテル内 16 階 曙

2.2)懇親会

同ホテル内 21 階フレンチレストラン『ベイサイド』

※懇親会会場は今年も東京湾の夜景が見える最上階です。

AAMT



AAMTジャーナル No.55

発行：アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

住所：〒171-0014 東京都豊島区池袋2-55-2鈴木ビル3階
(株)日本システムアプリケーション内

phone : 03-5951-3961 fax : 03-5951-3966

編集委員会：宇津呂 武仁 小谷 克則 大倉 清司

鈴木 博和 阿部 さつき

表紙(図部分)デザイン：阿部 さつき

事務局：神崎 享子 荻野 孝野

印刷所：株式会社ユリクリエイト

Asia-Pacific Association for Machine Translation (AAMT)

c/o Japan System Application Co., Ltd.

Suzuki Building 3F 2-55-2, Ikebukuro, Toshima-ku Tokyo 171-0014, JAPAN