

機械翻訳について

— その急激な進化 —

NICT

隅田 英一郎

目次

1. 自己紹介 (**NICT** って何?)
2. 機械翻訳 -過去, **現在**, 未来-
3. 機械翻訳への**対応例**

- 情報通信研究機構は、情報の電磁的流通及び電波の利用に関する技術の研究及び開発、高度通信・放送研究開発を行う者に対する支援、通信・放送事業分野に属する事業の振興等を総合的に行うことにより、情報の電磁的方式による適正かつ円滑な流通の確保及び増進並びに電波の公平かつ能率的な利用の確保及び増進に資することを目的とする。（国立研究開発法人情報通信研究機構法第4条）
- 総務省所管。
- 情報通信技術の研究開発や、情報通信分野の事業支援等を総合的に行うことを目的とし、全国8か所の研究拠点、2か所の電波送信所をもつ。本部は東京都小金井市。

ALL for ONE

日本企業の研究者を糾合し、先進的音声翻訳研究開発推進センター（ASTREC）をNICT内に設立（2014年9月16日）

出向元：

Panasonic、NEC、富士通、日立、東芝、SONY、
NTT、KDDI、NHK、FEAT、ATR-TREK 等



ONE for ALL



商用事例① MT+PEで廉価/高速



凸版のプレスリリースから引用

<http://www.toppan.co.jp/news/2017/07/newsrelease170703.html>

音声翻訳の専用ハードウェア

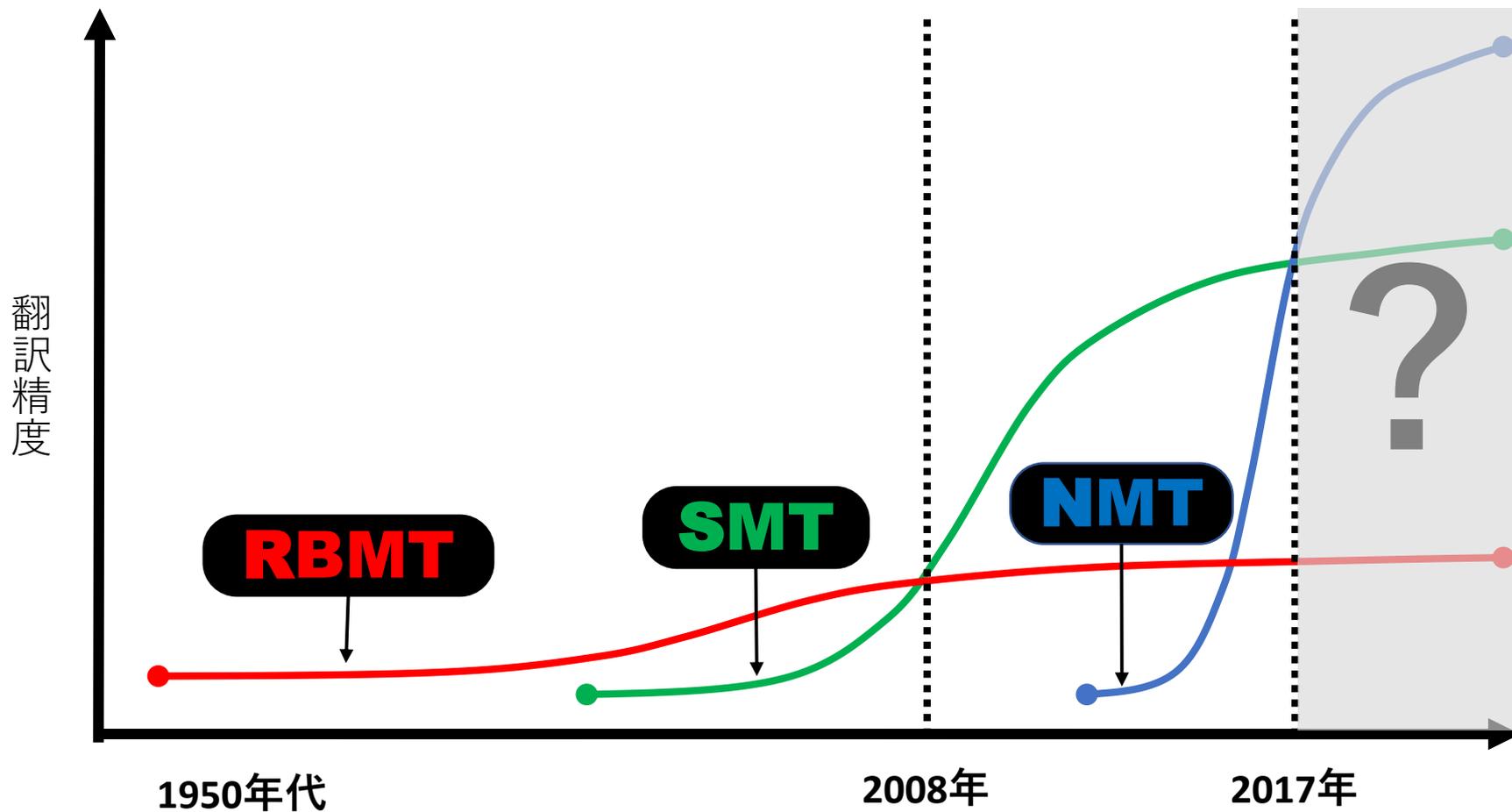


<https://iamili.com/ja/>

機械翻譯

過去, **現在**, 未來

自動翻訳の仕組みの変遷



RBMT(Rule-Based MT)

SOVから
SVOに

日本語文法規則

日英変換規則 +
日英対訳辞書

英語文法規則

SMT(Statistical MT) のモデル

① 対訳データ

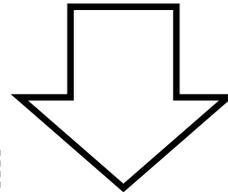
I will go to school. 私は通学します。

I will go to Kyoto. 京都に行きます。

I will go to the park. 私は公園に行きます。

I will go to shopping. 買い物に行きます。

I will go to walking. 私は散歩します。



② 確率付き対訳辞書

I will go します $2/5=40\%$

I will go 行きます $3/5=60\%$

AI時代の自動翻訳の原理

NMT(Neural MT、ニューラル機械翻訳)

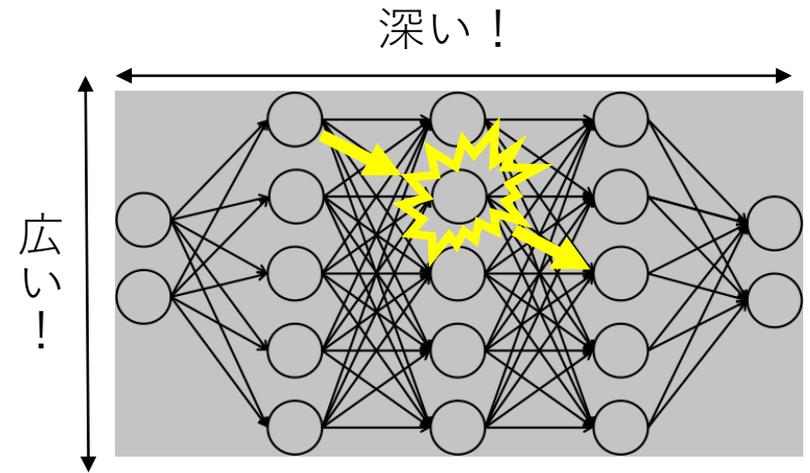
人間の神経回路網 (NN)



引用: <http://www.geeky-gadgets.com/wp-content/uploads/2011/07/Brain-epicness1.jpg>

単純な機能しか持たないニューロンの組み合わせで知的な処理を実現

人工の神経回路網 (NN)



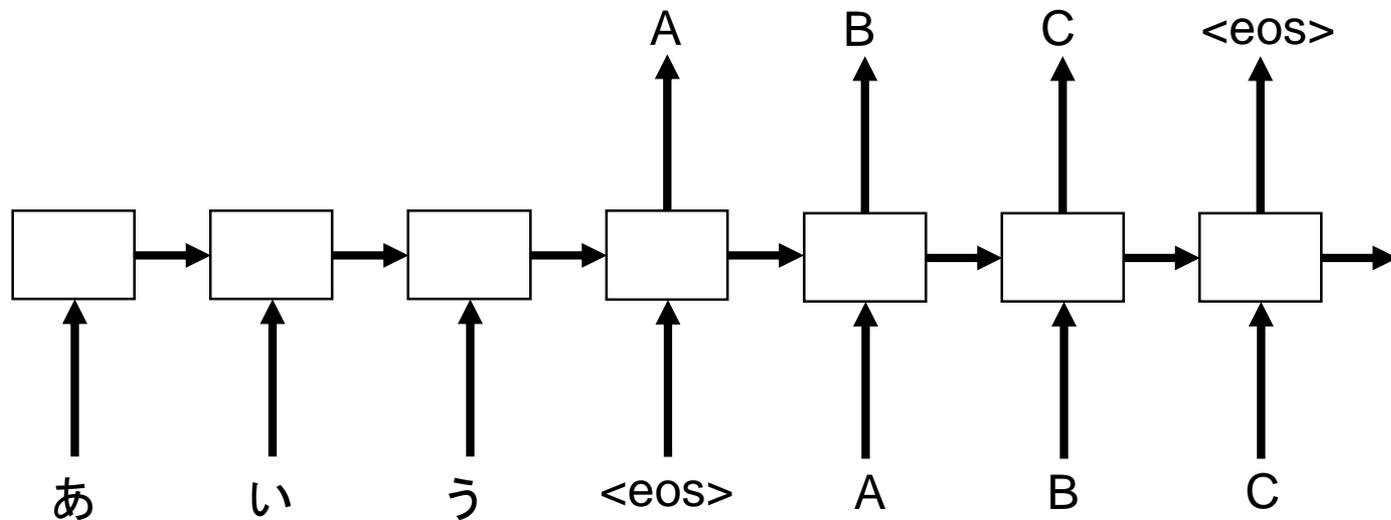
広い!

$$y_j = f\left(\sum_i w_{ji} x_i\right)$$

ニューロンの結合にデータを与えて重み w_{ji} を学習して知的な処理を実現

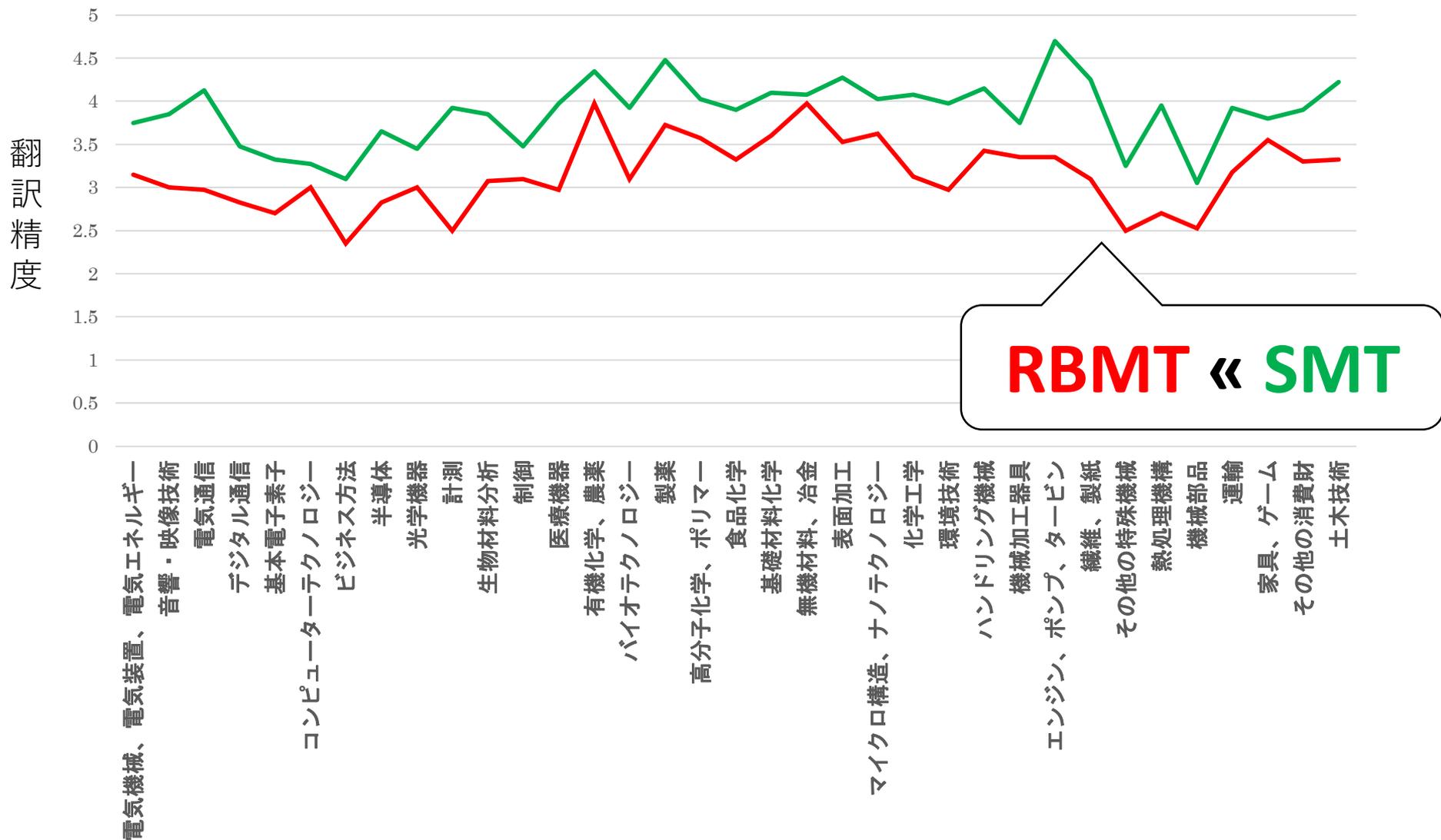
NMT(Neural MT、ニューラル機械翻訳) 系列変換のNNの深層学習で実現

- 翻訳データ (原文と訳文) を与えて翻訳関係を学習させる。



特許の全分野でSMTが強い

～ 平成27年度中国特許文献の機械翻訳の品質評価及び辞書整備に関する調査（特許庁）～



RBMT:YAHOO翻訳のサービス終了

YAHOO! 翻訳
JAPAN

Yahoo!翻訳 サービス終了のお知らせ

Yahoo!翻訳をご利用いただきありがとうございます。
誠に勝手ながら、Yahoo!翻訳は2017年6月29日に終了いたします。
日頃よりご利用いただいている皆様にはご迷惑をおかけすることとなり、
誠に申し訳ございません。
長年のご愛顧に、Yahoo!翻訳スタッフ一同、心よりお礼申し上げます。

サービス終了の理由

Yahoo! JAPANでは、機械翻訳利用のすそ野を広げることで言語間の壁を低くすべく、2005年より「Yahoo!翻訳」の提供を行ってまいりました。

パソコンから携帯、スマートフォンと、この10年余りの間に、インターネット利用環境は、大きく変化を遂げています。

当初は翻訳された言葉も不自然であり、またサービスの選択肢も少なかった翻訳サービスですが、どんどん進化を遂げ、より自然な機械翻訳サービスや、通信を利用した対面での翻訳サービスなどを選べる時代となりました。これは翻訳サービスの理想的な進化です。

そのため、機械翻訳利用のすそ野を広げることを目指してきた「Yahoo!翻訳」は、その役目を果たしたと考え、サービスを終了させることにいたしました。

ご意見は「[Yahoo!翻訳へのご意見](#)」までお願いいたします。

Q&A

Q.Yahoo!翻訳はいつ終了しますか？

A. 2017年6月29日に終了いたします。

Q. Yahoo!翻訳に代わる

A. 本ページ内で、ご利用可
い。

他社が提供する翻訳サービス
検討ください。

[Google 翻訳](#) テキ

[Weblio翻訳](#) テキ
単語

[エキサイト翻訳](#) テキ
対応。

[infoseekマルチ翻
訳](#) テキ

[Bing 翻訳](#) テキ
ト翻訳

[So-net 翻訳](#) テキ

<http://honyaku.yahoo.co.jp/info>

飛躍的改善(↗)を確認！

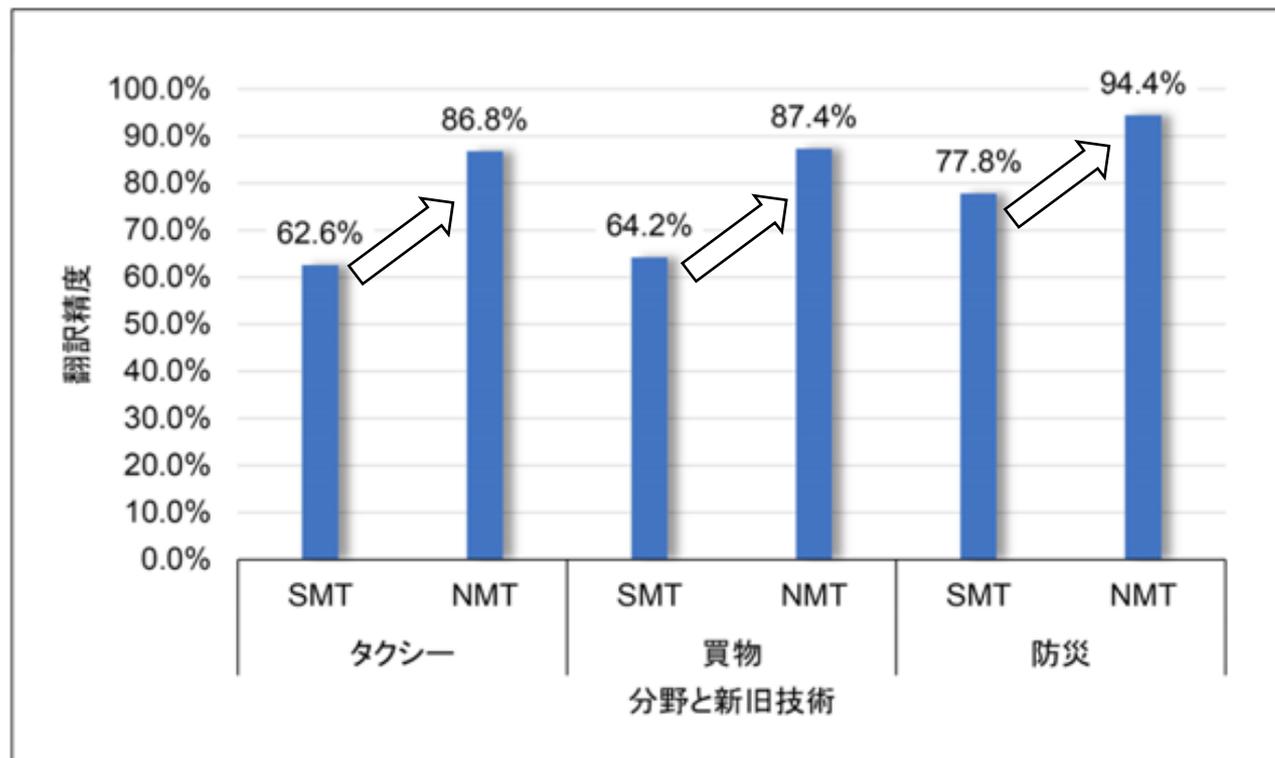
VoiceTraの日・英にニューラル機械翻訳を先行的に導入

ニューラル機械翻訳で音声翻訳アプリVoiceTraが 更なる高精度化を実現

～話し言葉の翻訳精度が大幅アップ・洗練された表現へ

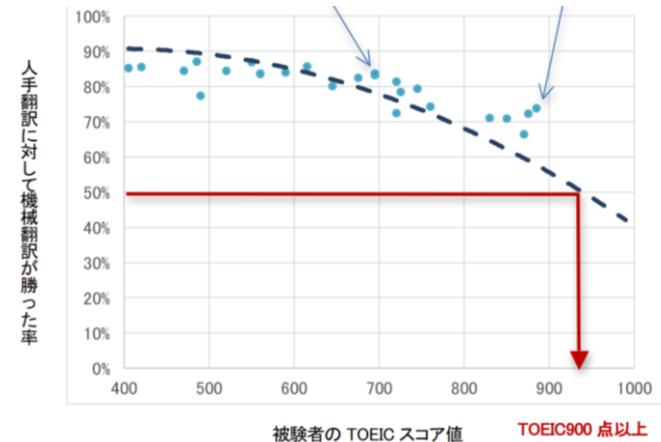
ポイント

- 日・英、双方向の話し言葉の翻訳を対象として、ニューラル
- 従来技術（統計翻訳）に比べ大幅に高い9割前後の翻訳精度
- 多言語音声翻訳アプリVoiceTraにニューラル機械翻訳技術を



※ 従来技術(統計翻訳(SMT)) AI時代の技術(ニューラル翻訳(NMT))

日英翻訳は（翻訳者
以外の）大半の日本
人を凌ぐ能力！



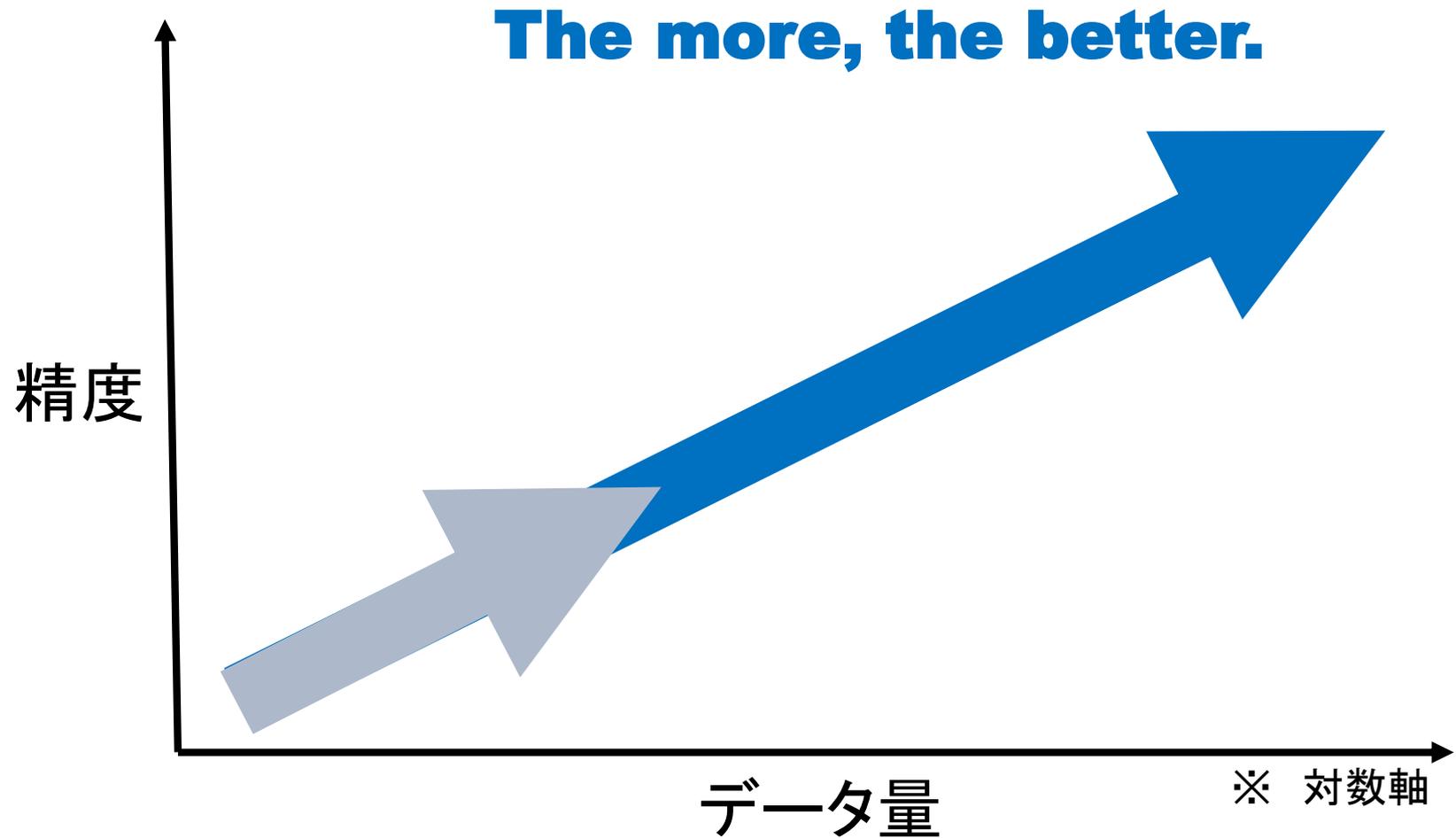
TOEIC900 点以上の英作文能力を持つ
深層学習による機械翻訳エンジンをリリース

【ポイント】

- 英作文能力は TOEIC900 点相当以上となり、従来型の機械翻訳に比べ翻訳精度が大幅に向上
- 企業向けの機械翻訳アプリケーションサービスの申込みを受け開始
- 書き言葉系のニューラル機械翻訳エンジンのデモを、みらい翻訳ホームページにて提供開始

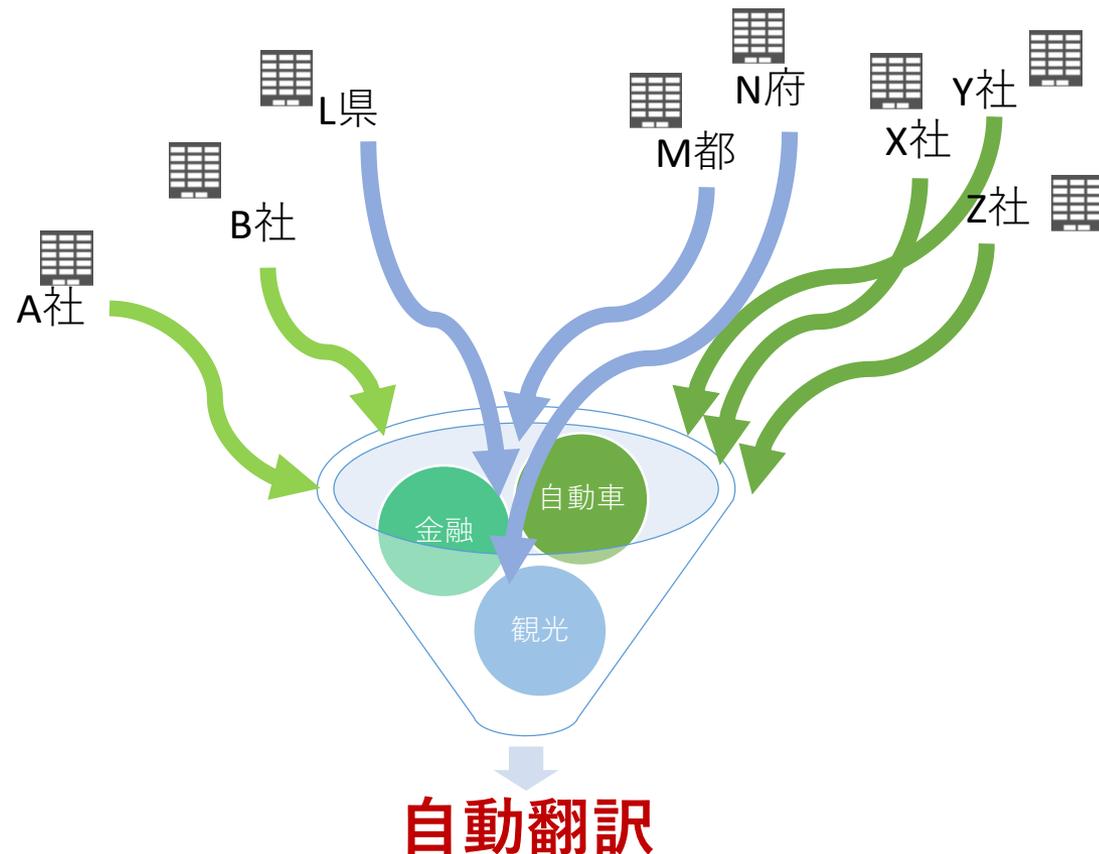
出典 <https://miraitranslate.com/uploads/2017/06/2d5778dcdee47e4197468bc922352179.pdf>

経験則： 精度 \propto データ量



『翻訳バンク』

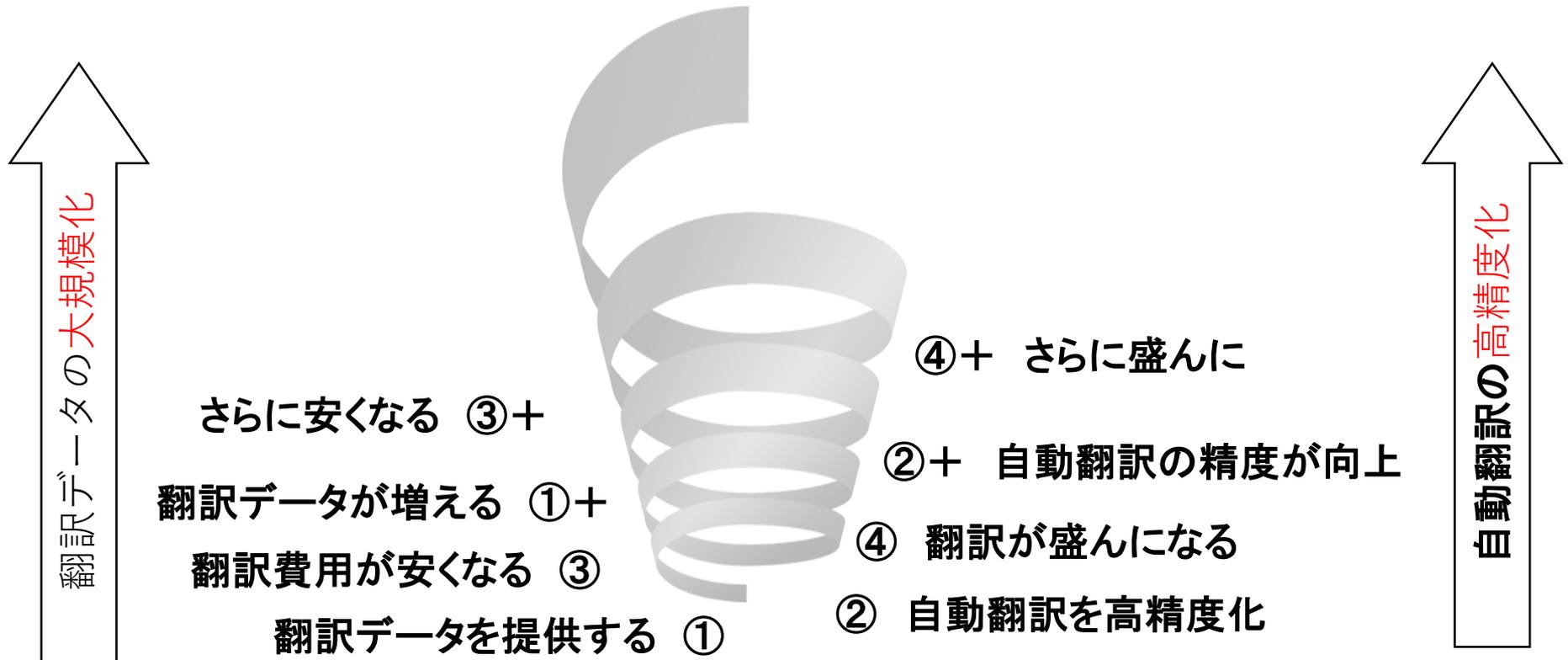
～5億文/年なので、たくさん集まりそう～



参考情報 http://www.soumu.go.jp/menu_news/s-news/01tsushin03_02000220.html

自動翻訳のエコシステム

自動翻訳を育てながら、高性能な自動翻訳を活用していくサイクル



同時通訳の自動化



NICT 情報通信研究機構
National Institute of Information and
Communications Technology

the sentence

The main verb is usually at the end of the sentence .

本動詞は普通、文の最後にあります。

機械翻訳の特徴と近未来

平均的
日本人
を凌ぐ
語学力

多言語
を扱え
る

速い

電子辞書のように
日常的に使われる

進化は
まだ停
まりそ
うにな
い。

24時間
365日
稼働

安い

進化する機械翻訳への対応策

- A) 機械翻訳を使った成功体験を勉強の動機づけとする。**
- B) 機械翻訳の使い方を学ぶ。**
 - ① 上手く機械翻訳できる日本語を書く。
 - ② 機械翻訳を使って英語を読む。
 - ③ 機械翻訳を使って英語を書く（日本語から英語に翻訳する）。
- C) 語学は勉強せず他に時間を使う。**

A) 機械翻訳を使った成功体験を勉強の動機づけとする。



出典

https://www.youtube.com/watch?v=cUGz_q8ldmQ

B) 機械翻訳の使い方を学ぶ



出典 <https://kotobank.jp/word/%E7%82%8A%E9%A3%AF%E5%99%A8-540126>

B)① 上手く機械翻訳できる日本語を書く。



出典
<http://digitalnature.slis.tsukuba.ac.jp/category/people/>

意味が明確な日本語文章は、ほぼ完璧に自動翻訳してくれます。「自動翻訳は、ちゃんと翻訳してくれない」と文句をつける人の多くは、そもそも自分が日本語で明確な文や機械が翻訳しやすい文を書けていないのではないのでしょうか？

(落合陽一、「これからの世界をつくる仲間たちへ」小学館、2016年4月2日、『コミュニケーションは「ロジック」がすべて』(P.147~))

B)② 機械翻訳を使って英語を読む。

The article is far more significant than is apparent.

「三人寄れば文殊の知恵」手法

1. この記事は明らかに明らかである。
2. この記事では、見かけよりはるかに重要である。
3. 物品が見かけよりもはるかに大きい。
4. その記事は明らかに見た目よりはるかに大きい。
5. 物品が見かけよりもはるかに顕著である。
6. 論文は見かけよりはるかに重要である。
7. 本論文では見かけよりもはるかに著しい。

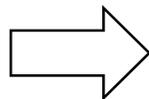
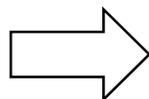
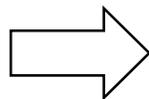
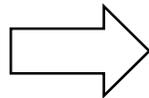
その論文は見掛けよりずっと重要な意味を持っている

B)③機械翻訳を使って英語を書く (日本語から英語に翻訳する)。

ソフトバンク、シンガポール配車アプリに出資 2017/7/15 10:13

http://www.nikkei.com/article/DGXLASDC15H1D_V10C17A7NNE000/

- 米紙ウォール・ストリート・ジャーナル（電子版）は14日、ソフトバンクグループがシンガポールの配車アプリ会社Grab・タクシー・ホールディングスに出資すると報じた。
- ソフトバンクは中国の配車アプリ大手、滴滴出行とともにGrabが実施する最大20億ドル（約2200億円）の増資を引き受ける。
- Grabは東南アジアを中心に7カ国、65都市で事業を展開。
- 米配車アプリ大手ウーバーテクノロジーズの東南アジアでの最大ライバルとされる。



- The Wall Street Journal reported on 14 that the Softbank Group will invest in the Singapore ride-hailing app company Grab Taxi Holdings.
- SoftBank will take up up to \$ 2000000000 (about 220000000000 yen) in capital to grab, with Didi Chuxing, a major Chinese application.
- The company operates in 65 cities in 7 countries, mainly Southeast Asia.
- It is said to be the biggest rival for the U.S. taxi application company, Uber Technologies in Southeast Asia.

1 文目

翻訳

米紙ウォール・ストリート・ジャーナル（電子版）は14日、ソフトバンクグループがシンガポールの配車アプリ会社Grab・タクシー・ホールディングスに出資すると報じた。

The Wall Street Journal reported on the 14 that the Softbank Group will invest in the Singapore application company Grab Taxi Holdings.

The Wall Street Journal (edition) reported on 14 Softbank group taking a stake in ride-hailing app company grab the Singapore Taxi Holdings.

2 文目

翻訳

ソフトバンクは中国の配車アプリ大手、滴滴出行とともにGrabが実施する最大20億ドル（約2200億円）の増資を引き受ける。

SoftBank will take up up to \$ 2000000000 (about 220000000000 yen) in capital to grab a stake in the company, along with drops from a major Chinese application.

SoftBank's application for taxi companies, offering up to \$2000000000 in (about 220000000000 yen) to grab and Didi's row.

ありがとうございました