

自動翻訳技術の概要と 情報通信研究機構の取組

情報通信研究機構（NICT）

内山将夫

日経エレクトロニクスセミナー

五感 × 脳（AI、xR、BMI） = 人間拡張が拓く新世界

2019年12月3日（火） 15:10～16:05

本講演の目的

- 自動翻訳技術の概要をお伝えする
- 自動翻訳技術の普及に関しての情報通信研究機構の取り組みの一部をご紹介します
- 自動翻訳技術の適切な使い方をお伝えする。

自動翻訳の歴史

- 計算機が発明されてすぐに自動翻訳の研究が開始(1950年代)
- 自動翻訳の考え自体は 1949 年にWarren Weaverが提案 (cf. Wikipedia)
- 半世紀以上の研究開発を経て、自動翻訳が一般に普及
 - みんなの自動翻訳@TexTra等のW e b 翻訳サービス
 - VoiceTra等のスマホアプリ
 - ポケトーク等の音声翻訳専用機

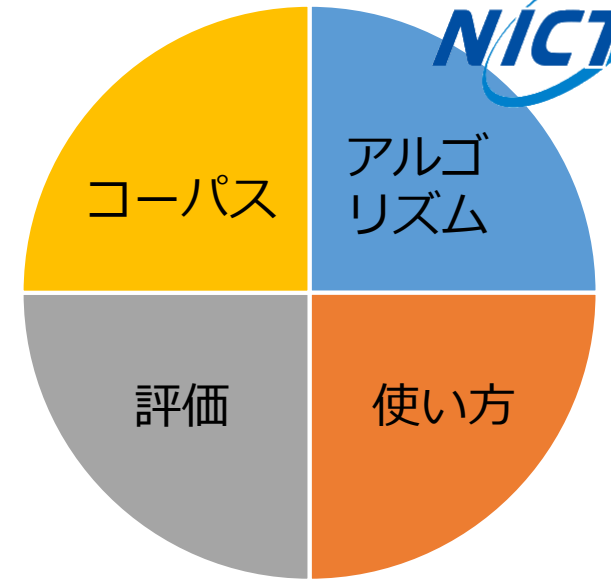


自動翻訳技術のタイプ

- 規則ベース自動翻訳
文法規則や辞書を人間が記述
上記に基づき自動翻訳を実施
- コーパスベース自動翻訳 (MT)
対訳コーパスから自動翻訳エンジンを自動学習
任意言語対に対して適用可能
ニューラル機械翻訳 (NMT) はこちら

コーパスベースMTの構成要素

- 自動翻訳アルゴリズム
1980年代からの発展⇒NMT
- 学習データとしての対訳コーパス
異なる言語で同一意味の文章の組からなるデータベース
- MTの評価手法
人間がどのように感じるかの観点からのMTの良さ
- MTの使い方
MTの普及により、MTの適切な使い方の周知が必要



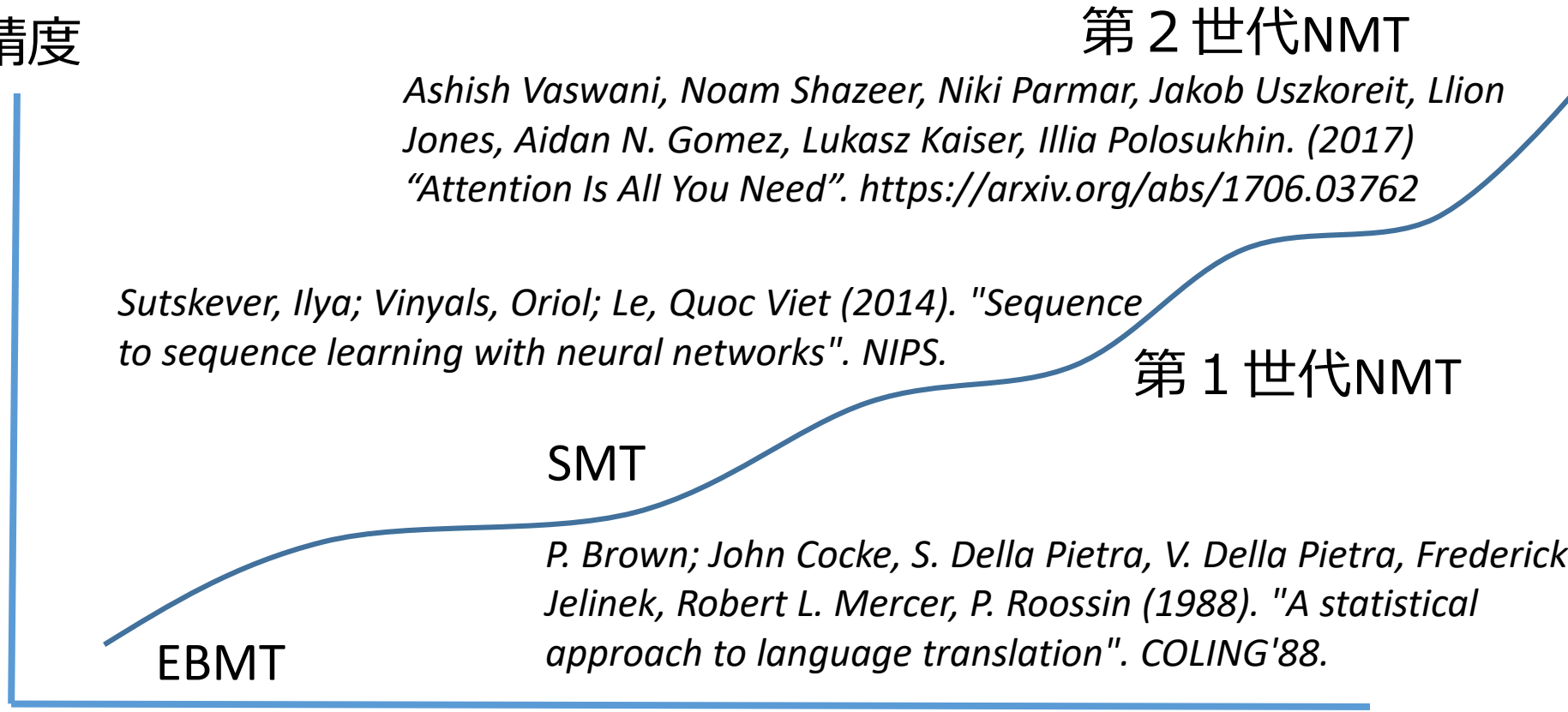
対訳コーパスの一例



- Italy have defeated Portugal 31-5 in Pool C of the 2007 Rugby World Cup at Parc des Princes, Paris, France.
- Italia berhasil mengalahkan Portugal 31-5 di grup C dalam Piala Dunia Rugby 2007 di Parc des Princes, Paris, Perancis.
- フランスのパリ、パルク・デ・プランスで行われた2007年ラグビーワールドカップのプールCで、イタリアは31対5でポルトガルを下した。
- អ៊ីតាលីបានឈ្នះលើព័រទុយហ្គាល់ 31-5 ក្នុងប្លុក C នៃពិធីប្រកួតពានរង្វាន់ពិភពលោកនៃកីឡាបាល់ទាត់ឆ្នាំ2007ដែលប្រព្រឹត្តទៅប៉ារីសខេត្តប្រ៊ីន ក្រុងប៉ារីស បារាំង។
- Itali telah mengalahkan Portugal 31-5 dalam Pool C pada Piala Dunia Ragbi 2007 di Parc des Princes, Paris, Perancis.
- ပြင်သစ်နိုင်ငံ ပါရီမြို့ပါဒက်စ် ပရင်စက် ၌ ၂၀၀၇ခုနှစ် ရပ်ဘို ကမ္ဘာ့ ဖလား တွင် အီတလီ သည် ပေါ်တူဂီ ကို ၃၁-၅ ဂိုး ဖြင့် ရေကူးကန် စီ တွင် ရှုံးနိမ့်သွားပါသည်။ ။
- Ý đã đánh bại Bồ Đào Nha với tỉ số 31-5 ở Bảng C Giải vô địch Rugby thế giới 2007 tại Parc des Princes, Pari, Pháp.
- អ៊ីតាលី បាន ឈ្នះ ព័រទុយហ្គាល់ ដោយ គុណភាព ៣១ ទល់ ៥ ក្នុង ក្រុម C នៃ ការ ប្រកួត ប្រជែង ពិភពលោក ឆ្នាំ ២០០៧ លើ កីឡា បាល់ ទាត់ ពិភពលោក ឆ្នាំ ២០០៧ ដែល ប្រព្រឹត្ត ទៅ ប៉ារីស ខេត្ត ប្រ៊ីន ក្រុង ប៉ារីស បារាំង ។
- Natalo ng Italya ang Portugal sa puntos na 31-5 sa Grupong C noong 2007 sa Pandaigdigang laro ng Ragbi sa Parc des Princes, Paris, France.

コーパスベースMTアルゴリズムの進展

翻訳精度



年代

第2世代NMT（汎用NT） @みんなの自動翻訳



The screenshot shows a web browser window with the URL <https://mt-auto-minhon-mlt.ucrj.jgn-x.jp/content/demo/>. The page title is "みんなの自動翻訳@TexTra". The interface includes a navigation menu with "翻訳データ", "自動翻訳", and "ツール". A "ヘルプ" button is visible in the top right.

The main content area is titled "自動翻訳". It features a language selection interface with "英語" and "日本語" buttons, a bidirectional arrow, and a dropdown menu currently set to "汎用NT 【英語 - 日本語】 1". A green "翻訳" button is to the right.

The source text (English) is: "The dominant sequence transduction models are based on complex recurrent or convolutional neural networks in an encoder-decoder configuration. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train."

The translated text (Japanese) is: "支配的なシーケンス変換モデルは、エンコーダ/デコーダ構成における複雑なリカレントまたは畳み込みニューラルネットワークに基づいている。また、最適なモデルは、アテンション機構を介してエンコーダとデコーダを接続する。本稿では、注意機構のみに基づき、リカレントと畳み込みを完全に省く、新しい簡単なネットワークアーキテクチャTransformerを提案する。2つの機械翻訳タスクに関する実験では、これらのモデルは、より並列化可能であり、トレーニングに要する時間が大幅に短縮される一方で、品質が優れていることが示されている。"

At the bottom, there are checkboxes for "文章を結合する" and "文章を区切る", and buttons for "← 訳文コピー" and "× リセット".

汎用NT @みんなの自動翻訳

原文英語	汎用NT (第2世代)
<p>The dominant sequence transduction models are based on complex recurrent or convolutional neural networks in an encoder-decoder configuration.</p> <p>The best performing models also connect the encoder and decoder through an attention mechanism.</p> <p>We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely.</p> <p>Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train.</p>	<p>支配的なシーケンス変換モデルは、エンコーダ/デコーダ構成における複雑なリカレントまたは畳み込みニューラルネットワークに基づいている。</p> <p>また、最適なモデルは、アテンション機構を介してエンコーダとデコーダを接続する。</p> <p>本稿では、注意機構のみに基づき、リカレントと畳み込みを完全に省く、新しい簡単なネットワークアーキテクチャTransformerを提案する。</p> <p>2つの機械翻訳タスクに関する実験では、これらのモデルは、より並列化可能であり、トレーニングに要する時間が大幅に短縮される一方で、品質が優れていることが示されている。</p>

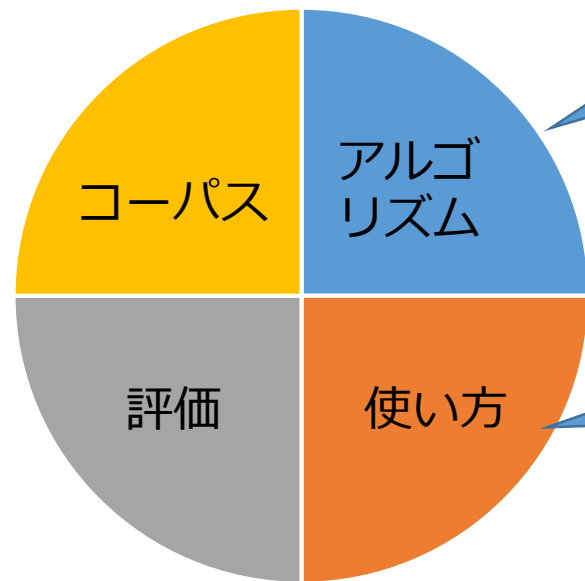
English sentences are part of Abstract from Ashish Vaswani, et al. (2017) "Attention Is All You Need".
<https://arxiv.org/abs/1706.03762>

汎用NT @みんなの自動翻訳

原文日本語	汎用NT（第2世代）
<p>2016年のニューラル機械翻訳（NMT）の実用化は、翻訳業界に衝撃を与え、ポケトークのような自動翻訳端末の市場拡大につながるなど、社会に大きなインパクトを与えた。</p> <p>ただし、翻訳技術や自然言語処理技術（NLP）分野では、その後も革命級のブレークスルーが相次いでいる。</p> <p>翻訳を含む言語系の人工知能（AI）が従来の常識を次々と塗り替え、ありえないペースで発展している。</p>	<p>The practical application of Neural Machine Translation (NMT) in 2016 gave a shock to the translation industry and had a large impact on society, such as the expansion of the market for automatic translation terminals such as PokeTalk.</p> <p>However, breakthroughs in translation and natural language processing (NLP) have continued.</p> <p>Artificial intelligence (AI), a linguistic system that includes translation, is evolving at an incredible pace, breaking conventional wisdom.</p>

原文は次の一部：野澤 哲生「AI翻訳が人間超え、言葉の壁崩壊へ」日経エレクトロニクス、2019/08/20
<https://tech.nikkeibp.co.jp/atcl/nxt/mag/ne/18/00046/00002/>

NMTにより新しくなった点



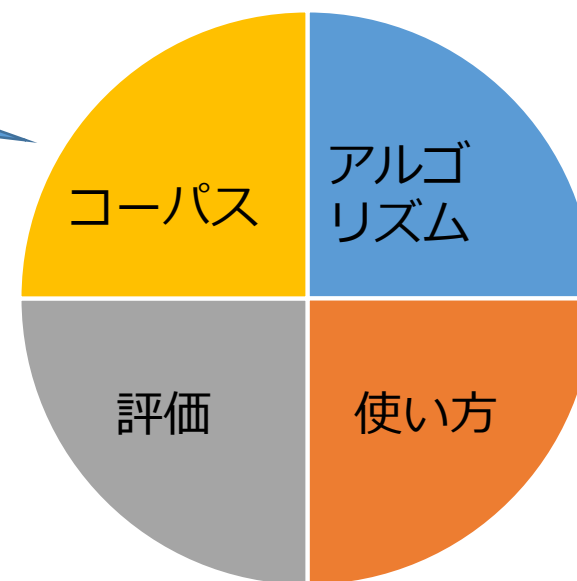
ニューラルネットワークによる高精度アルゴリズム

高精度自動翻訳の社会的普及

NMTでも以前と同じ点

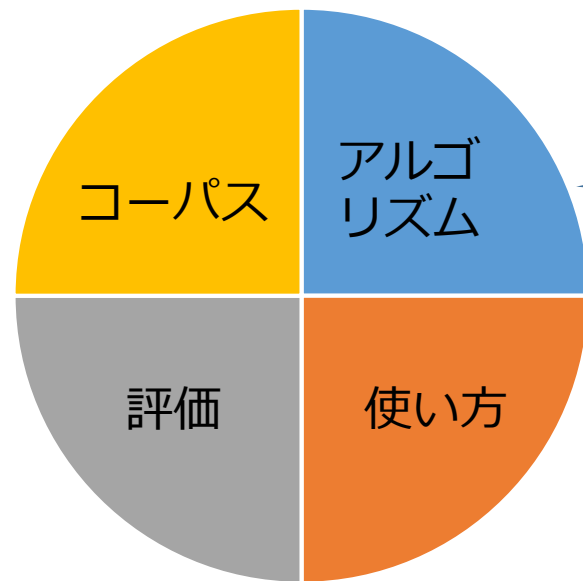
大規模対訳
コーパスが必要

人間による評価
が必要



NMTアルゴリズムの概要

- 基本訓練
- アダプテーション



ニューラルネットワークによる高精度
アルゴリズム

NMTの訓練とアダプテーション概要

【訓練】 1文ずつNMTモデルのパラメタを調整する

- 大雑把には：
 - 入力文を翻訳
 - 参照訳文と比較
 - 翻訳文と参照訳文の違いに応じてNMTのパラメタを更新
 - 以上を大規模に繰り返す（数億回になることもある）

【アダプテーション】

- 訓練済みNMTモデルに、上記訓練を特定分野データで追加
- 訓練済みモデルをベースにするので、比較的少量データで高精度

NMTの基本構造 (第1・2世代ほぼ共通)

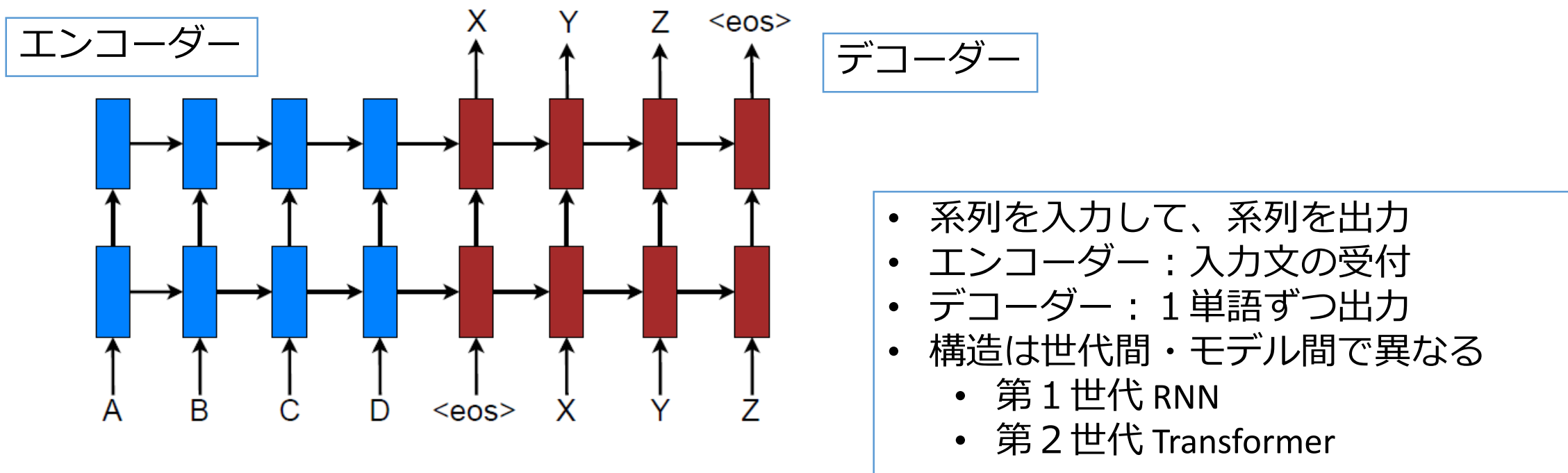


Figure 1: **Neural machine translation** – a stacking recurrent architecture for translating a source sequence A B C D into a target sequence X Y Z. Here, <eos> marks the end of a sentence.

- 系列を入力して、系列を出力
- エンコーダー：入力文の受付
- デコーダー：1単語ずつ出力
- 構造は世代間・モデル間で異なる
 - 第1世代 RNN
 - 第2世代 Transformer

Thang Luong; Hieu Pham; Christopher D. Manning. (2015)
Effective Approaches to Attention-based Neural Machine Translation. EMNLP

第1世代NMT：注意機構が重要

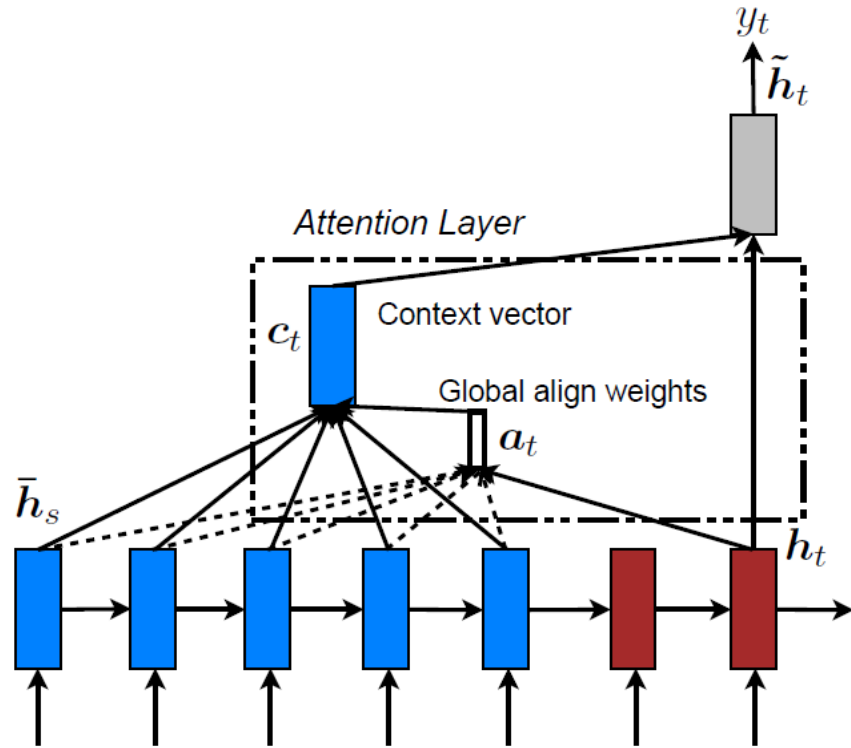


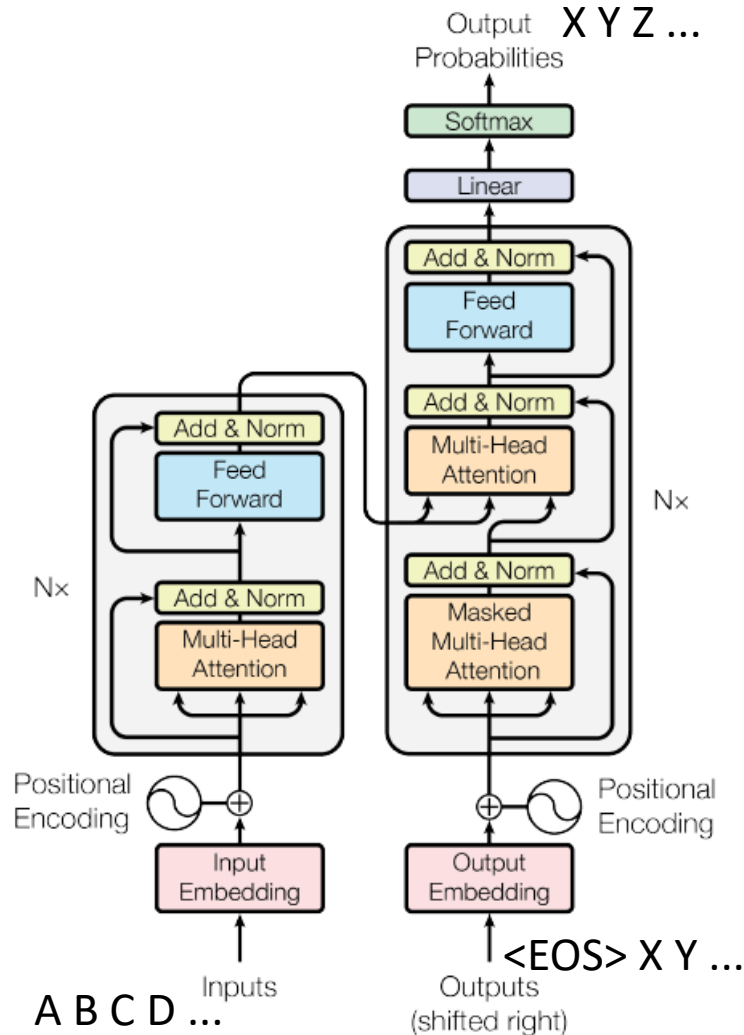
Figure 2: **Global attentional model** – at each time step t , the model infers a *variable-length* alignment weight vector \mathbf{a}_t based on the current target state \mathbf{h}_t and all source states $\bar{\mathbf{h}}_s$. A global context vector \mathbf{c}_t is then computed as the weighted average, according to \mathbf{a}_t , over all the source states.

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}_c[\mathbf{c}_t; \mathbf{h}_t])$$

$$\begin{aligned} \mathbf{a}_t(s) &= \text{align}(\mathbf{h}_t, \bar{\mathbf{h}}_s) \\ &= \frac{\exp(\text{score}(\mathbf{h}_t, \bar{\mathbf{h}}_s))}{\sum_{s'} \exp(\text{score}(\mathbf{h}_t, \bar{\mathbf{h}}_{s'}))} \end{aligned}$$

Thang Luong; Hieu Pham; Christopher D. Manning. (2015) *Effective Approaches to Attention-based Neural Machine Translation*. EMNLP

第2世代NMT : Transformer



- デコーダーからエンコーダーへのアテンション
RNNと似ている
- エンコーダーでセルフアテンション
各層で一つ下の階層にアテンション
- デコーダーでセルフアテンション
各層で一つ下の階層にアテンション (未来除く)
- 訓練が高速

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017) "Attention Is All You Need". <https://arxiv.org/abs/1706.03762>

Figure 1: The Transformer - model architecture.

第1世代と第2世代の比較

Table 2: The Transformer achieves better BLEU scores than previous state-of-the-art models on the English-to-German and English-to-French newstest2014 tests at a fraction of the training cost.

Model	BLEU		Training Cost (FLOPs)	
	EN-DE	EN-FR	EN-DE	EN-FR
ByteNet [18]	23.75			
Deep-Att + PosUnk [39]		39.2		$1.0 \cdot 10^{20}$
GNMT + RL [38]	24.6	39.92	$2.3 \cdot 10^{19}$	$1.4 \cdot 10^{20}$
ConvS2S [9]	25.16	40.46	$9.6 \cdot 10^{18}$	$1.5 \cdot 10^{20}$
MoE [32]	26.03	40.56	$2.0 \cdot 10^{19}$	$1.2 \cdot 10^{20}$
Deep-Att + PosUnk Ensemble [39]		40.4		$8.0 \cdot 10^{20}$
GNMT + RL Ensemble [38]	26.30	41.16	$1.8 \cdot 10^{20}$	$1.1 \cdot 10^{21}$
ConvS2S Ensemble [9]	26.36	41.29	$7.7 \cdot 10^{19}$	$1.2 \cdot 10^{21}$
Transformer (base model)	27.3	38.1	$3.3 \cdot 10^{18}$	
Transformer (big)	28.4	41.8	$2.3 \cdot 10^{19}$	

翻訳の比較的難しい英独で BLEU 4ポイント程度の大きな改善

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017) "Attention Is All You Need". <https://arxiv.org/abs/1706.03762>

アダプテーションの有効性の事例

自動車法規文の自動翻訳をニューラル技術で高精度化

～トヨタとの共同研究を通じ、英日・中日翻訳の実用度が向上～

- 自動車業界からトヨタが翻訳バンクに協力、翻訳データを提供
- 自動車法規を対象とした翻訳をニューラル英日翻訳システムで
24%実用度向上

法規原文	人による翻訳	ニューラル翻訳
This Australian Design Rule prescribes requirements for the number and mode of installation of lighting and light signalling devices on motor vehicle other than L-group vehicles.	本オーストラリア設計規則はLグループ車両以外の自動車への灯火および灯火信号装置の数と取り付け方法に関する要件を規定する。	本オーストラリア設計規則は、Lグループ車両以外の自動車への灯火および灯火信号装置の数および取り付け方法に関する要件を規定する。

アダプテーションで専門用語に適応

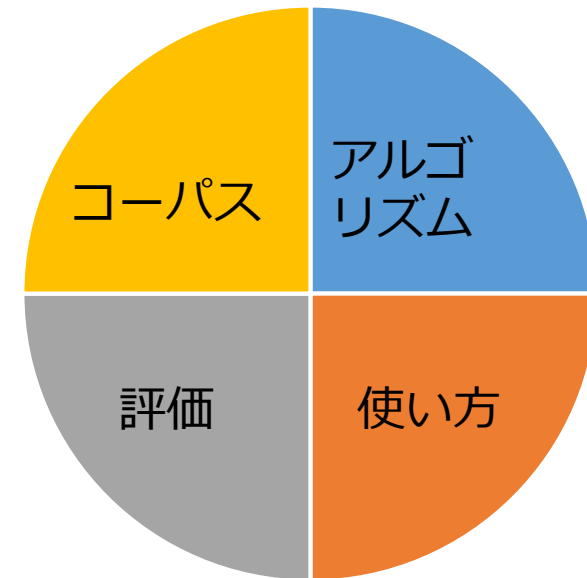


原文	第1世代汎用NMT	第2世代汎用NMT	第1世代金融NMT	第2世代金融NMT
本取引は、相互売買による取引実績が豊富な本投資法人ならではの取引であると考えており、本投資法人は今後も戦略的かつ継続的に資産入替を実施していく方針です。	This transaction is considered to be a transaction unique to this investment corporation, which has a rich record of transactions through mutual trade. This investment corporation will continue to change assets in a strategic and ongoing manner.	We believe that this transaction is unique to the Investment Corporation, which has a rich track record of transactions through mutual transactions. The Investment Corporation will continue to implement this transaction strategically and continuously.	The Transaction is considered as a transaction that is unique to the Investment Corporation, which has a wealth of transaction experience through mutual transaction. The Investment Corporation will continue to implement strategic and continuous asset replacement in the future.	The Investment Corporation believes that the Transaction is unique to the Investment Corporation, which has a wealth of transaction records through mutual transactions. Therefore, the Investment Corporation will continue to implement strategic and continuous asset replacement.

M T の評価の概要

- M T の評価はM T を改善する原動力
- 誤訳報告⇒誤訳の改善
- 適切な評価⇒適切なM T

人間による評価
が必要



2 タイプの評価

- 内的基準による評価
 - アプリケーションと関係ないMTの評価
 - 自動評価
 - 人間評価
- 外的基準による評価
 - アプリケーションの達成度によるMTの評価

自動評価

- 簡便
- M T 訳と参照訳を比較
- 評価値を自動計算
- 開発にはとても役に立つ！
 - システム改修後、自動評価値に大きな変化がないことを確認
 - 大きな数値差（BLEUなら5程度）があれば品質差もほぼある

自動評価尺度 BLEU

- M T 訳と参照訳の類似度の尺度
- 1 0 0 0 文～5 0 0 0 文程度の文章で計算
- 一般的に BLEU 3 0 ～4 0 以上であれば高精度

Papineni, Kishore and Roukos, Salim and Ward, Todd and Zhu, Wei-Jing. (2002)
BLEU: a Method for Automatic Evaluation of Machine Translation. ACL

みんなの自動翻訳でのBLEU計算法

- 自動翻訳 > 自動翻訳評価 > 「+新規登録」



The screenshot shows a web browser window with the URL <https://mt-auto-minhon-mlt.ucrjgn-x.jp/content/eval/edit/index.html>. The page title is "みんなの自動翻訳@TexTra". The main heading is "自動翻訳評価 登録". The form includes a name input field with the placeholder "名前を入力してください。", a language direction selector set to "日本語 ↔ 英語", and an automatic translation type selector with "標準" and "共有" options. Below these are several checkboxes for different translation methods, all currently unchecked:

- 汎用NT【日本語 - 英語】 ⓘ
- 汎用NMT【日本語 - 英語】 ⓘ
- 特許NT【日本語 - 英語】 ⓘ
- 特許NMT【日本語 - 英語】 ⓘ
- 特許請求項NMT【日本語 - 英語】 ⓘ
- 対話NMT【日本語 - 英語】 ⓘ
- ミニ対話NMT【日本語 - 英語】 ⓘ

自動評価尺度BLEUでの比較

	第1世代	第2世代
汎用	25.1	29.0
金融	34.6	38.4

第1世代 << 第2世代
汎用 << 金融

AAMT 2019, Tokyo: パネルディスカッション「機械翻訳もある総合的な翻訳サービスの模索～金融・IR分野を例に～」の内山のスライドの1枚より引用
日英翻訳の例について、みんなの自動翻訳でBLEUを計算

人間評価

- 原文とMT訳を人間が比較
- 人間が誤訳かを分析
- MTの研究開発に必須
- 第2世代NMTでは、おおむね正しい翻訳ができるようになり、以前は見逃されていた誤訳が目立ち始めた
 - 「NICT」⇒「N I C T」（なぜか空白が挿入）
- ピンポイントな誤訳を検出する自動評価方法は知られていない
- ピンポイントな誤訳は、ピンポイントなルールなどで改善

アプリケーションに基づく評価

A/B テスト

原文

MT 訳

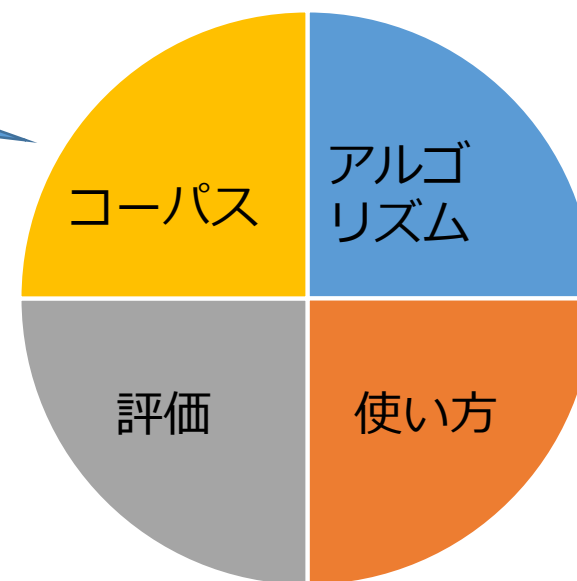
ポスト
エディット

全部
人間訳

どの文章が一番利益になるか？
(訳質は第 1 の目標ではない)

NMTでの対訳コーパスの重要性

大規模対訳
コーパスが必要

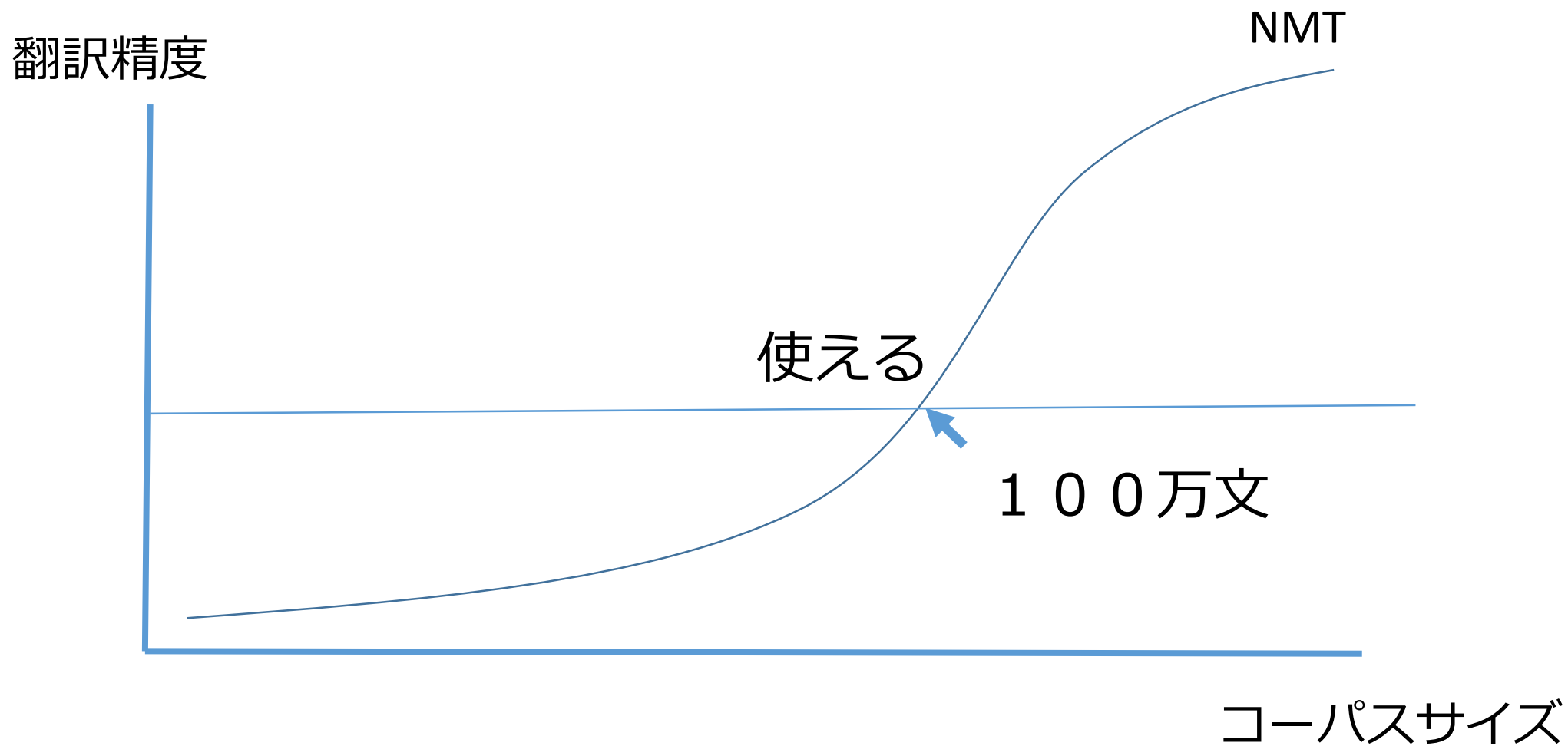


対訳コーパスの一例 (再掲)

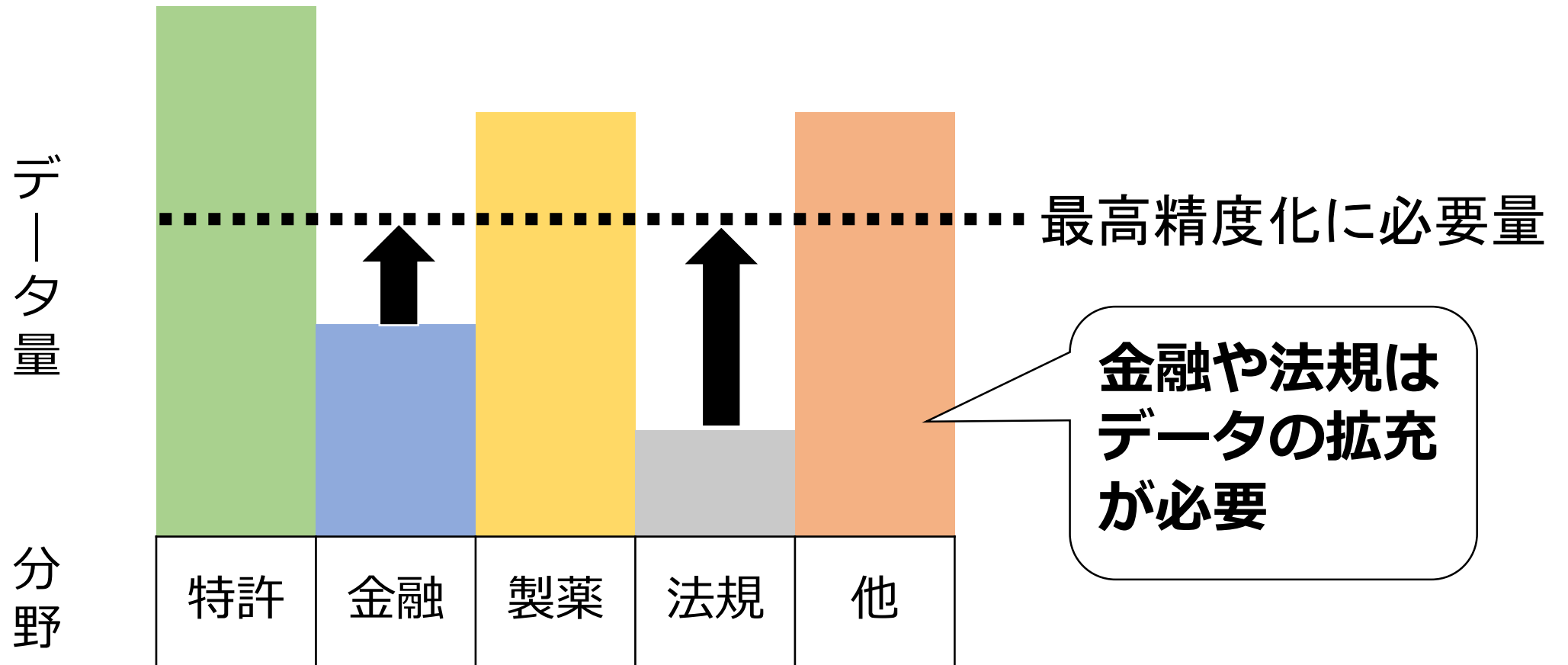


- Italy have defeated Portugal 31-5 in Pool C of the 2007 Rugby World Cup at Parc des Princes, Paris, France.
- Italia berhasil mengalahkan Portugal 31-5 di grup C dalam Piala Dunia Rugby 2007 di Parc des Princes, Paris, Perancis.
- フランスのパリ、パルク・デ・ランスで行われた2007年ラグビーワールドカップのプールCで、イタリアは31対5でポルトガルを下した。
- អ៊ីតាលីបានឈ្នះលើព័រទុយហ្គាល់ 31-5 ក្នុងប្លុក C នៃពិធីប្រកួតពានរង្វាន់ពិភពលោកនៃកីឡាបាល់ឱបឆ្នាំ2007ដែលប្រព្រឹត្តទៅប៉ារីសខេត្តប្រ៊ីន ក្រុងប៉ារីស បារាំង។
- Itali telah mengalahkan Portugal 31-5 dalam Pool C pada Piala Dunia Ragbi 2007 di Parc des Princes, Paris, Perancis.
- ပြင်သစ်နိုင်ငံ ပါရီမြို့ပါဒက်စ် ပရင့်စက် ဌ ၂၀၀၇ခုနှစ် ရုပ်ဘီ ကမ္ဘာ့ ဖလား တွင် အီတလီ သည် ပေါ်တူဂီ ကို ၃၁-၅ ဂိုး ဖြင့် ရေကူးကန် စ တွင် ရှုံးနိမ့်သွားပါသည်။ ။
- Ý đã đánh bại Bồ Đào Nha với tỉ số 31-5 ở Bảng C Giải vô địch Rugby thế giới 2007 tại Parc des Princes, Pari, Pháp.
- อิตาลีได้เอาชนะโปรตุเกสด้วยคะแนน31ต่อ5 ในกลุ่มC ของการแข่งขันรักบี้เวิลด์คัพปี2007 ที่สนามปาร์กเดแพรงส์ ที่กรุงปารีส ประเทศฝรั่งเศส
- Natalo ng Italya ang Portugal sa puntos na 31-5 sa Grupong C noong 2007 sa Pandaigdigang laro ng Ragbi sa Parc des Princes, Paris, France.

大規模対訳コーパスでNMTの性能向上



高精度 N M T には対訳データが重要



分野にマッチした対訳が重要 特許対訳は普通の文には不向き

This is **the** desk.



特許対訳

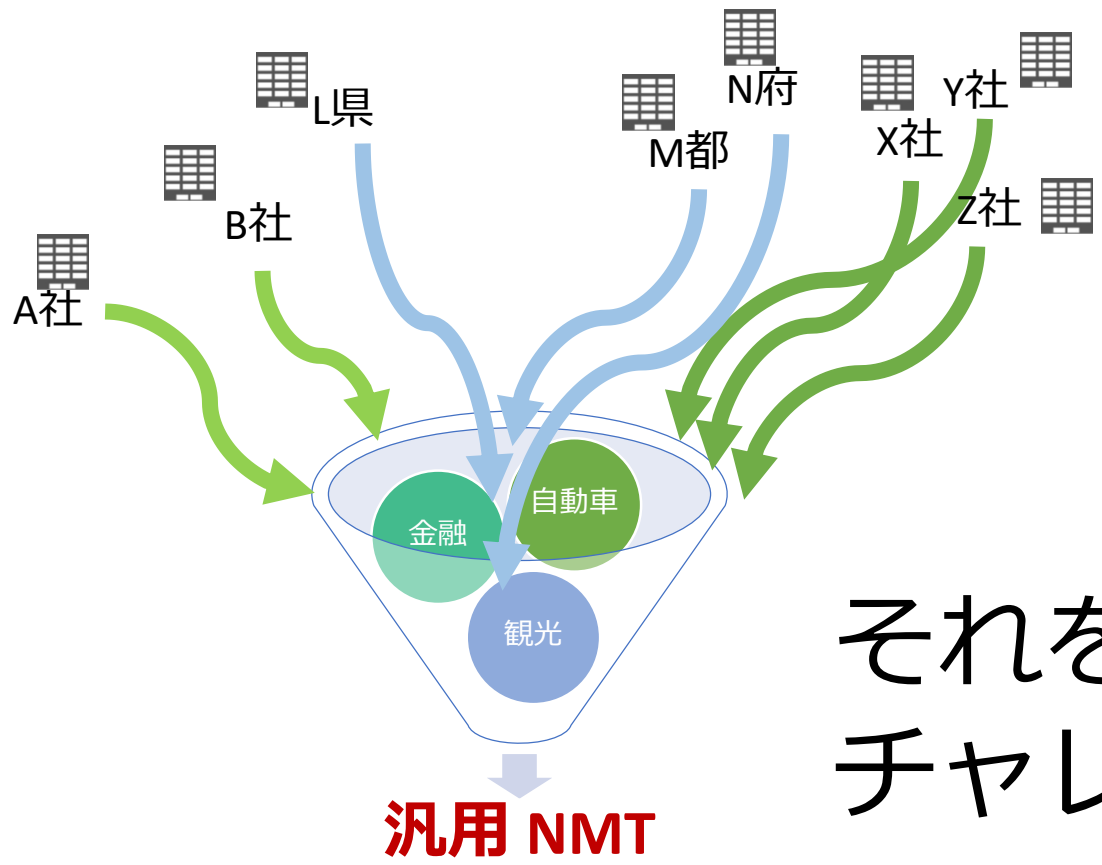


特許NMT



これは**前記**机である。 ???

汎用対訳コーパスが汎用NMTには必要

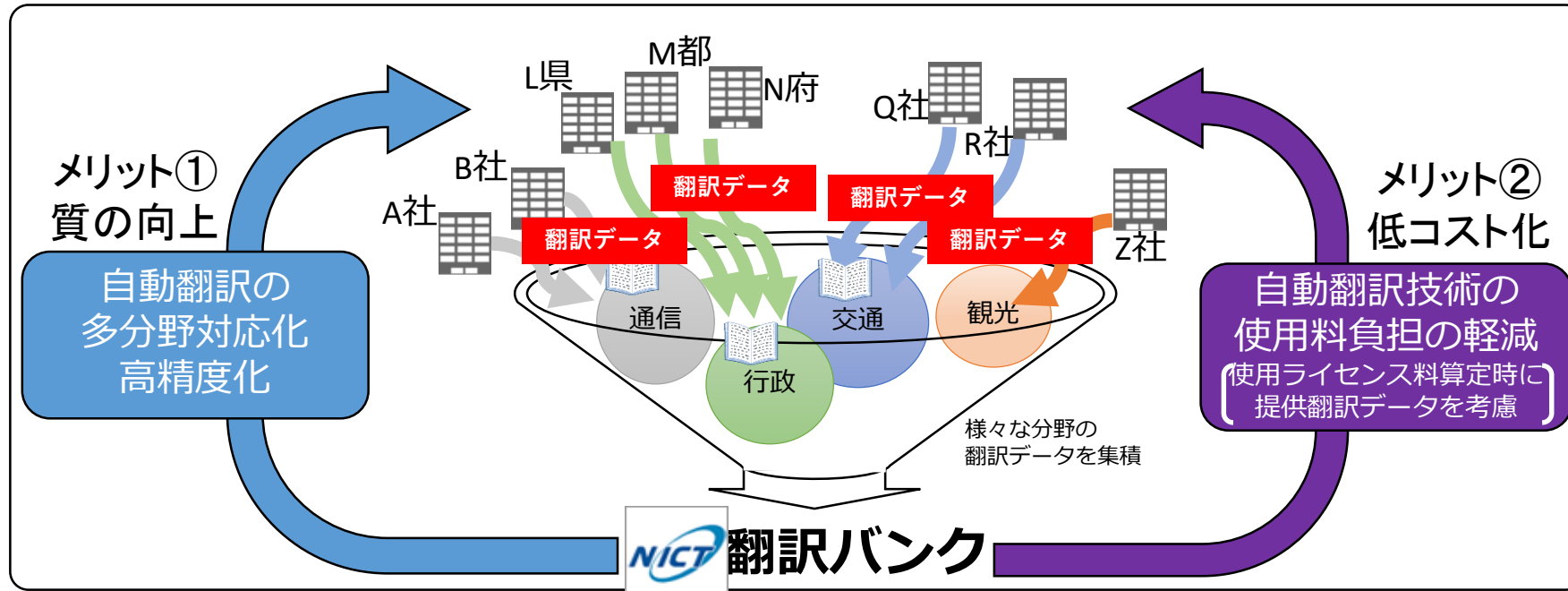


対訳コーパスは
散らばっている

それを集めるのが
チャレンジ

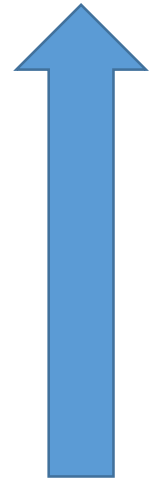
翻訳バンクで高品質対訳データ

NICTの自動翻訳技術の使用ライセンス料の算定の際に、提供が見込まれる翻訳データを勘案して負担を軽減する仕組みを導入

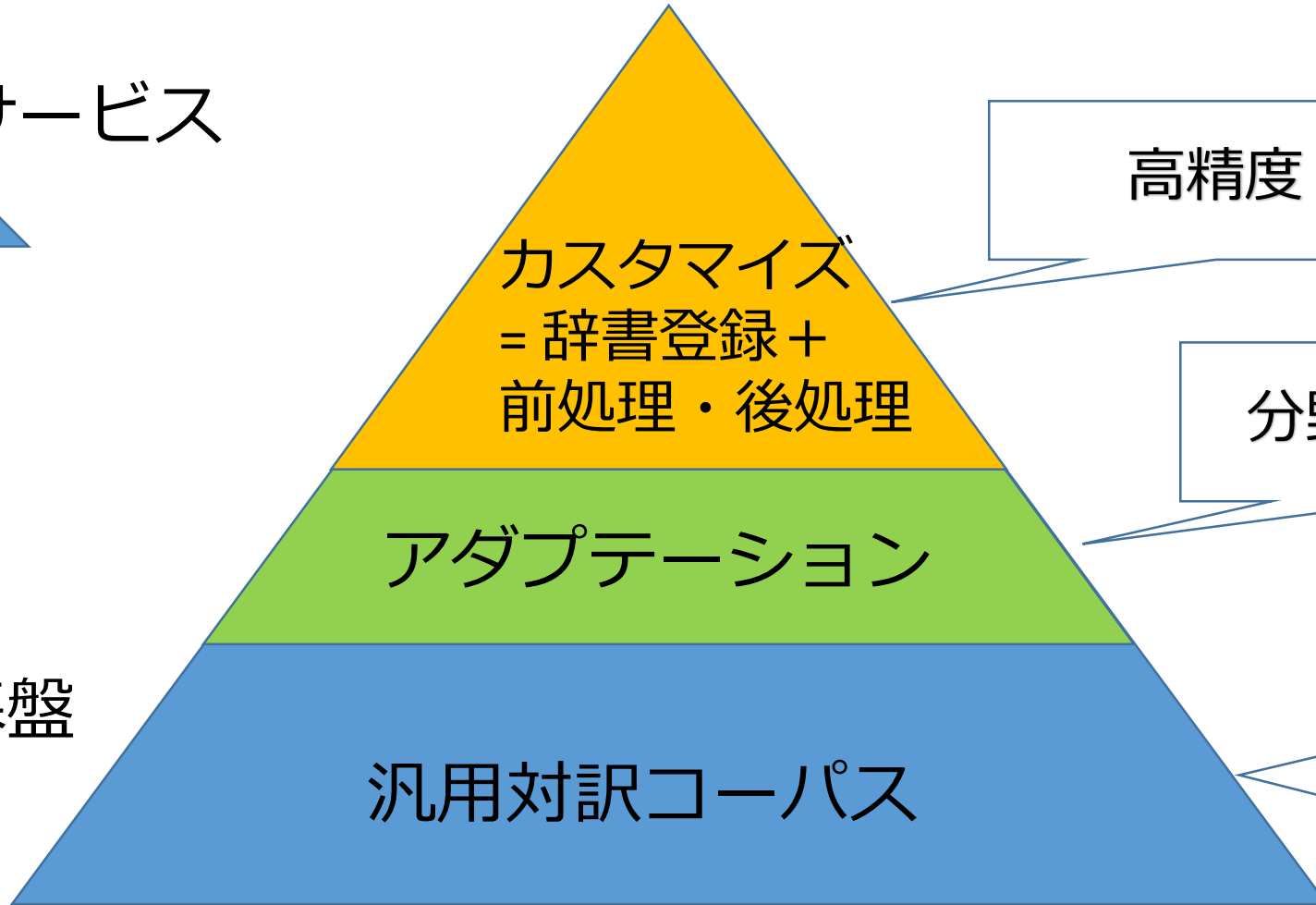


高精度MT = 汎用NMT + アダプテーション + カスタマイズ

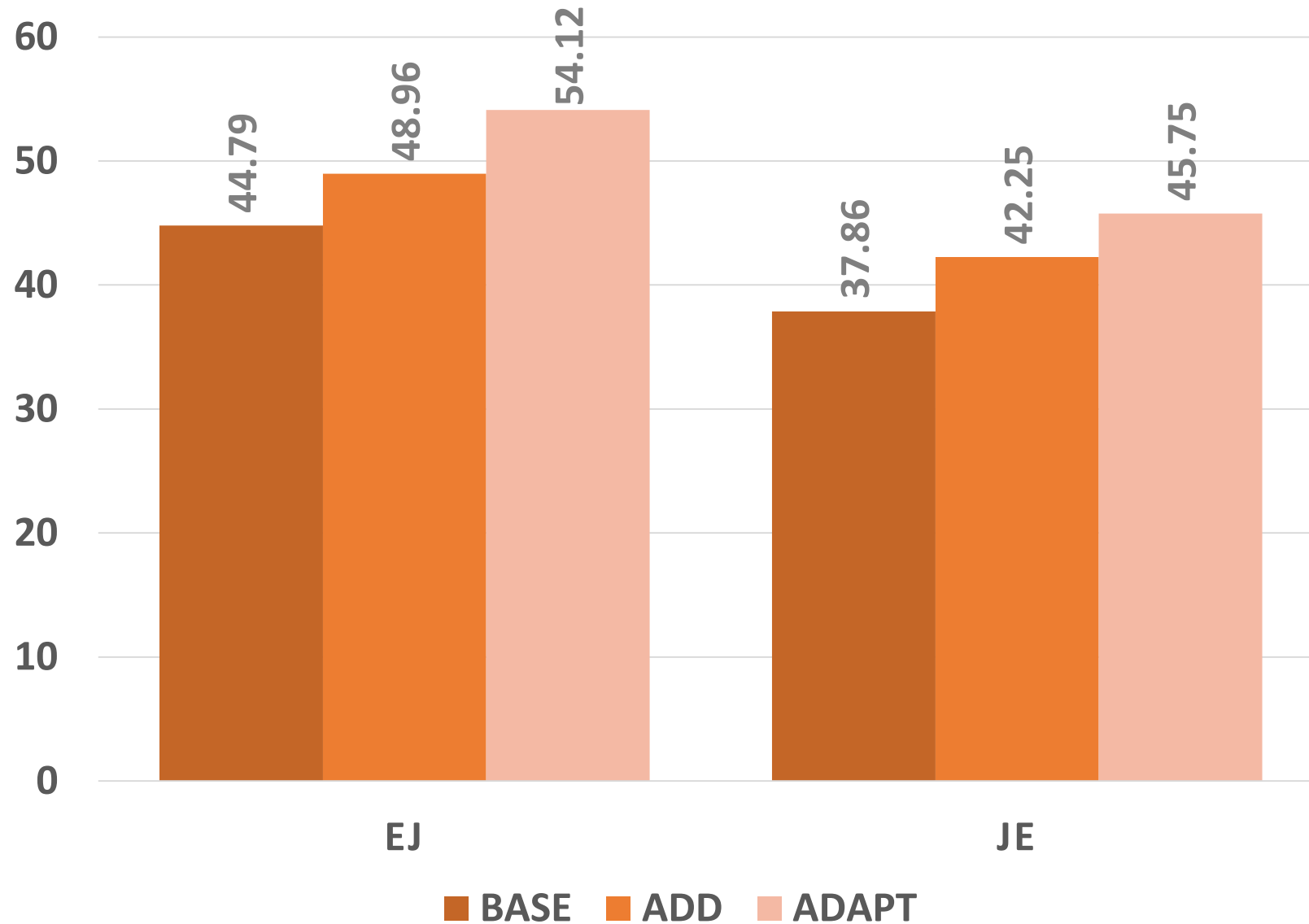
個別サービス



共通基盤
NMT

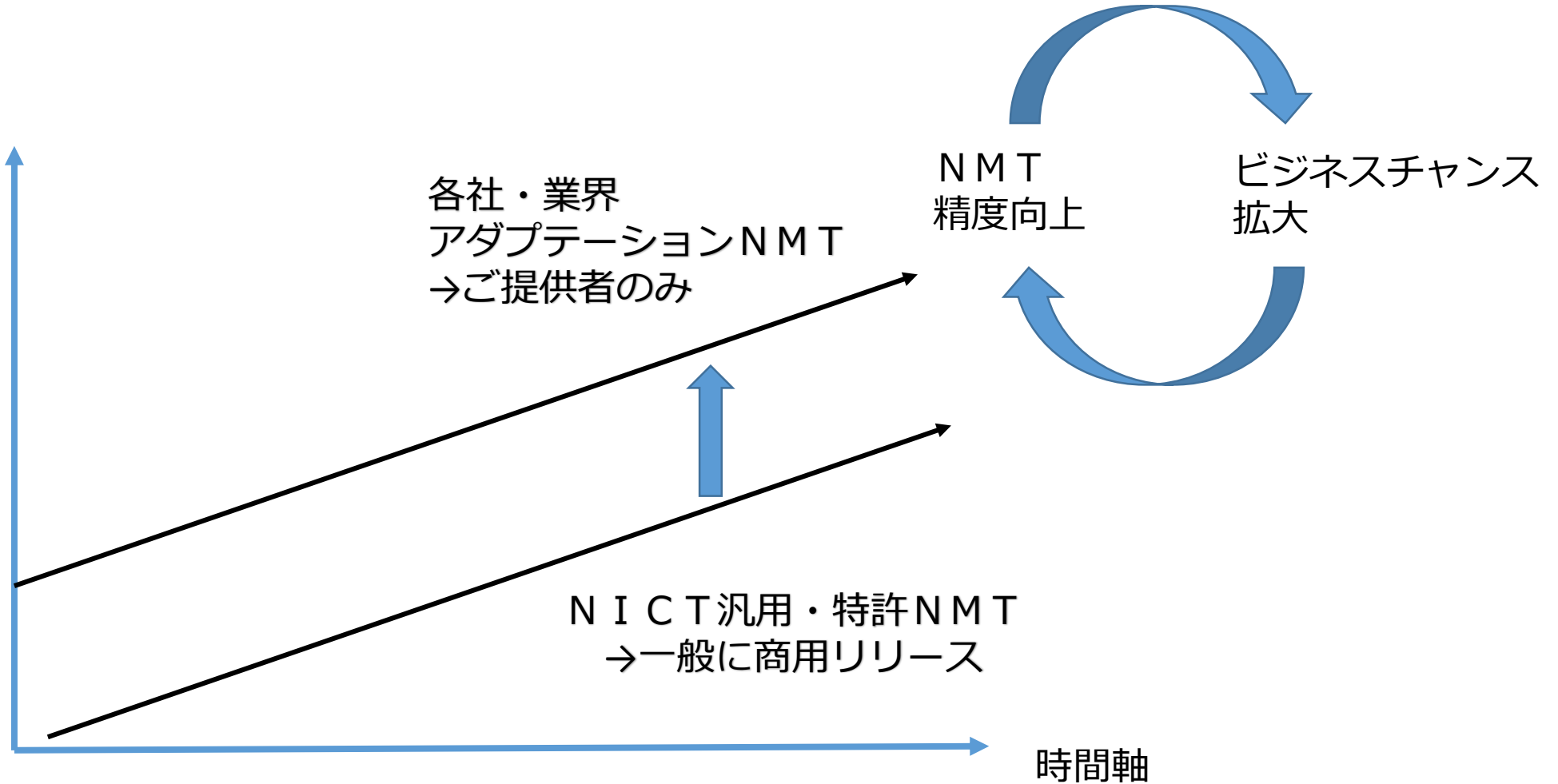


翻訳精度BLEUの2段階向上



対訳は2度使う

翻訳精度



翻訳バンクへのご協力方法

1 『みんなの自動翻訳@TexTra®』 の利用



2 NICTとの翻訳データの 提供に関する契約の締結



3 ライセンス契約に伴う 翻訳データの提供



翻訳データ量でライセンス料の負担軽減

NMTの社会的な普及と安全な使い方



2019年はNMTがビジネスで普及

	トピック	リリース主体
3月	オール・ジャパンの枠組みを活用「みんなの自動翻訳@K I（カスタム版）」の提供を開始	川村インターナショナル
4月	機械翻訳サービスの和文英訳がプロ翻訳者レベルに、英文和訳はTOEIC960点レベルを達成	みらい翻訳
4月	自動車法規文の自動翻訳をニューラル技術で高精度化 ～トヨタとの共同研究を通じ、英日・中日翻訳の実用度が向上～	NICT
5月	特許庁から受注した機械翻訳システムの稼動開始について ～NICTが開発した最新のニューラル機械翻訳エンジンの採用により、正確で自然な翻訳を実現～	東芝デジタルソリューションズ
7月	AI翻訳サービス「T-tact AN-ZIN」を発表 「みんなの自動翻訳@TexTra®」が商用利用可	十印
9月	IR・金融分野向け自動翻訳エンジンの性能向上を確認 ～日本財務翻訳との共同研究により実用化へ～	NICT
10月	大規模翻訳データによる製薬業界向けAI自動翻訳システムの最適化 ～情報通信研究機構（NICT）とR&D Head Clubの共同開発～	NICT

N I C T の自動翻訳技術のご紹介

- VoiceTra
- みんなの自動翻訳@TexTra
- NICTの自動翻訳アプリは非商用利用にご利用いただけます
- 商用利用には技術移転先のサービスをご利用いただけます



多言語音声翻訳アプリ

< ボイストラ >

無料アプリ

VoiceTra[®]

話した内容を外国語に翻訳
31言語に対応！

他のサイズはありますか。

ショッピングで

次の列車は何時ですか。

駅で

穴子ください。

寿司屋で



旅行会話に最適！

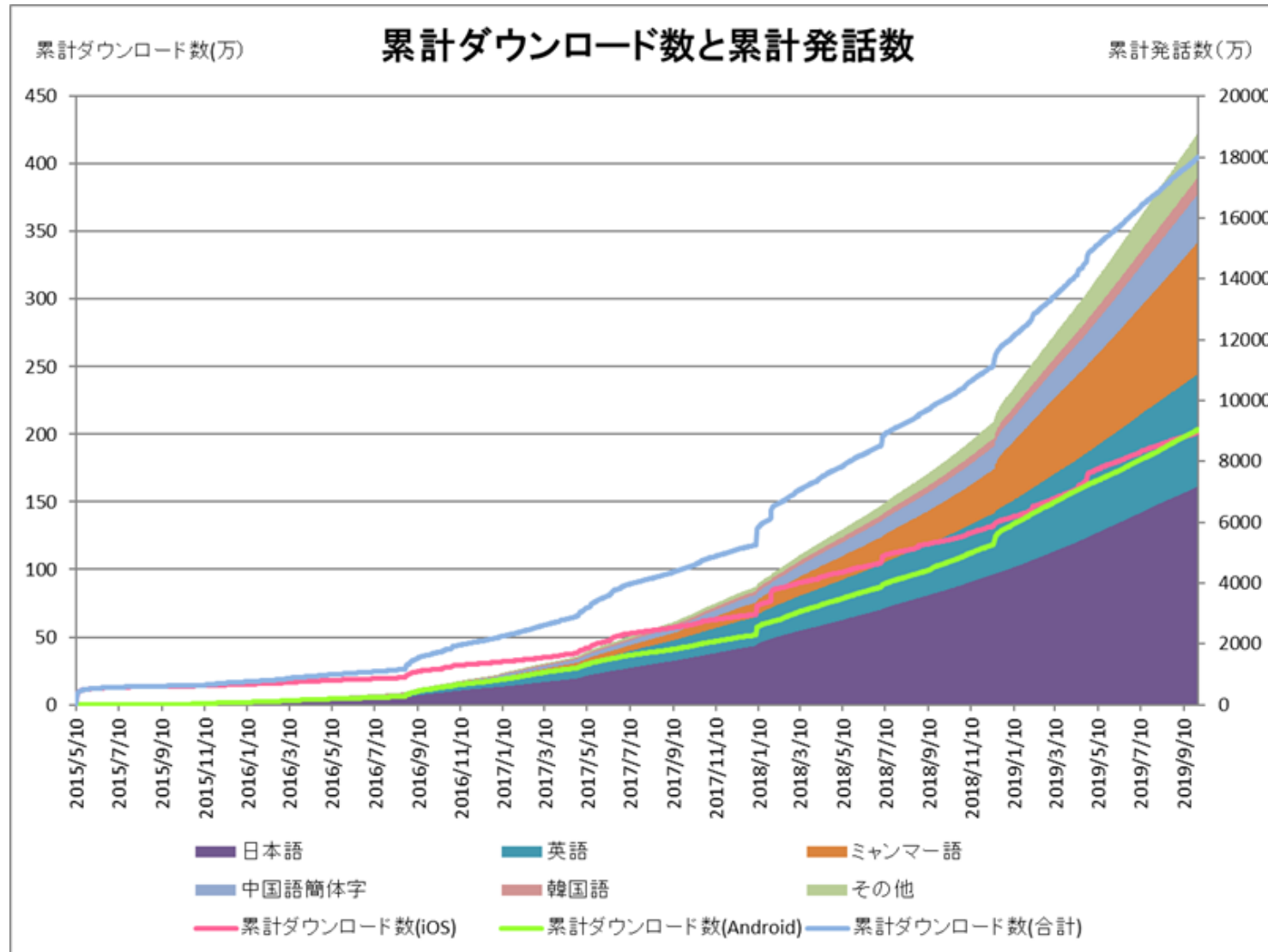
チェックインは何時ですか。

ホテルで

少し熱があります。

病院で

累計400万DL, 2億発話



2019年9月30日現在

NMTの使い方@みんなの自動翻訳

<https://mt-auto-minhon-mlt.ucri.jgn-x.jp/>

まずは
無料で
新規登録



The screenshot shows the homepage of the 'みんなの自動翻訳@TexTra' website. The browser address bar displays the URL <https://mt-auto-minhon-mlt.ucri.jgn-x.jp/>. The page features a green header with the site name and a 'Language' dropdown menu. On the left side, there is a login and registration form with fields for '名前' (Name) and 'パスワード' (Password), a 'ログイン' (Login) button, a checkbox for 'ログインしたままにする' (Stay logged in), and a link for 'パスワード再発行' (Reset password). Below this is a prominent red button labeled '新規登録' (New registration). To the right, a section titled 'みんなの自動翻訳@TexTra®とは' (About) includes a 'ヘルプ一覧' (Help list) link, a graphic with characters 'ア', 'ア', and 'あ', and text explaining that the site is a free automatic translation service developed by NICT. A footer section at the bottom contains the text '便利な「翻訳エンジン」を無料でご利用いただけます。' (You can use the convenient 'translation engine' for free).

テキストを貼り付けて翻訳

自動翻訳

[ヘルプ](#)

英語



日本語

汎用NT 【英語 - 日本語】 1

翻訳

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks in an encoder-decoder configuration.

The best performing models also connect the encoder and decoder through an attention mechanism.



支配的なシーケンス変換モデルは、エンコーダ/デコーダ構成における複雑なリカレントまたは畳み込みニューラルネットワークに基づいている。

また、最適なモデルは、アテンション機構を介してエンコーダとデコーダを接続する。

 文章を結合する

↓ 訳文コピー

× リセット

自動翻訳結果を逆翻訳で確認



The screenshot shows a web browser window with the URL <https://mt-auto-minhon-mlt.ucri.jgn-x.jp/content/demo/>. The page title is "みんなの自動翻訳@TexTra®". The interface is in Japanese, with a language dropdown set to "mutiyama3". The main heading is "自動翻訳". Below it, there are buttons for "日本語" and "英語", and a dropdown for "汎用NT【日本語 - 英語】". A green "翻訳" button is visible. A dialog box is open with the text "自動翻訳結果が正しいかを逆翻訳で確認します。". To the right, a detailed view of the reverse translation process is shown, including the original text, the translated text, a 70.2% score, and a "修正" button. A note at the bottom explains that corrected translations will be used for future translations and training.

自動翻訳結果が正しいかを逆翻訳で確認します。

Check the machine translation result by reverse translation.

原文
自動翻訳結果が正しいかを逆翻訳で確認します。

逆翻訳 70.2%

機械翻訳結果を逆翻訳で確認します。

修正する

Check the machine translation result by reverse translation.

修正

※修正した訳文は、今後の翻訳に再利用されます。具体的には、「リサイクル」に編集済みとして登録され、次に同じ原文があなたにより入力されたときには、訳文として再利用されます。また、「共有自動翻訳」の「汎用NMT+」の訓練に使われるなど、自動翻訳技術の改良の目的に利用される可能性があります。

- 文章を結合する
- 文章を区切る
- 文章を読点で区切る

利用可能な翻訳モデル

VoiceTra翻訳モデル	第1世代NMT	第2世代NMT
対話NMT ミニ対話NMT	汎用NMT 特許NMT 特許請求項NMT サイエンスNMT 金融NMT 汎用NMT+	汎用NT 特許NT 金融NT

第2世代NMT導入中

翻訳エディタで翻訳 (インストール不要)

Step.1 翻訳したいサイトを決めます。



Oldest Poodle Has Left Hospital
By Angelika Moshin

Sachiko, the world oldest known female poodle has been discharged from Ontago Dog Hospital following lung surgery yesterday. The surgery was quite successful and the 28-year-old poodle is in a good condition.

Step.2 翻訳エディタでサイトを読み込むと自動で日本語に翻訳。英語のサイトが見た目はそのまま日本語に。



最も古いプードルは退院
アンジェリカ・モシン

世界最古の女性プードルであるサチコは、昨日、肺手術でOntago犬病院から退院しました。手術はかなり成功しており、28歳のプードルは良好な状態にある。

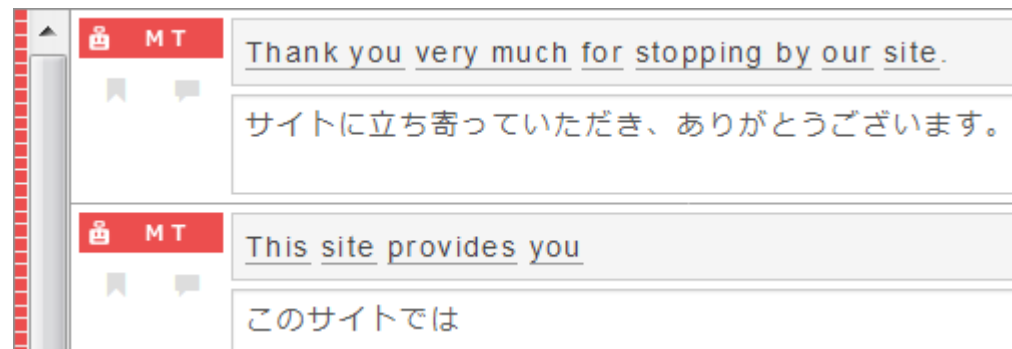
Step.3 自動翻訳結果を修正していきます。まずはタイトルから…



最高齢プードル、退院
アンジェリカ・モシン

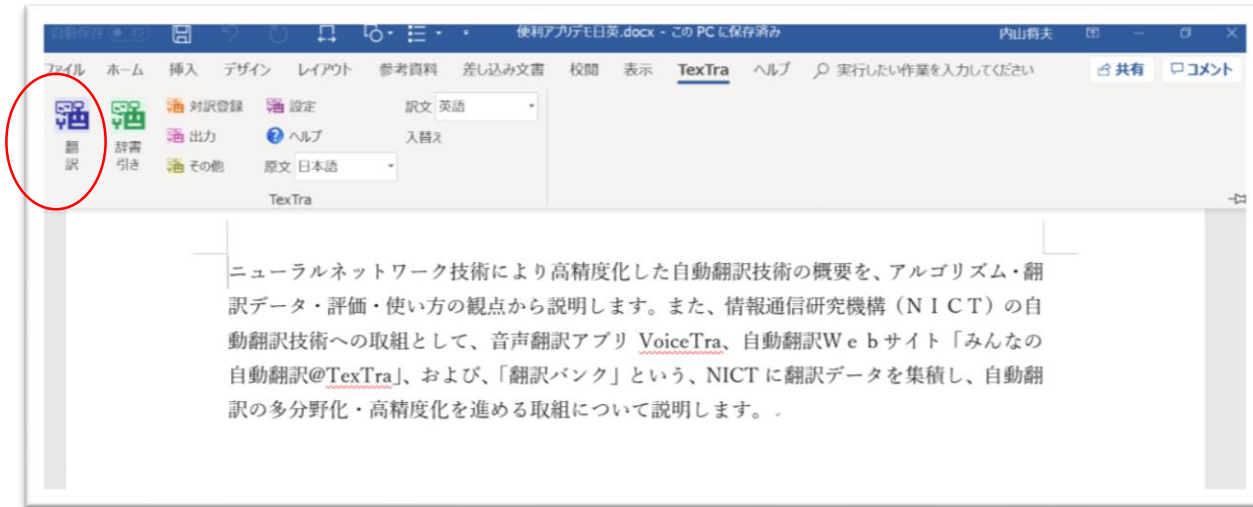
世界最古の女性プードルであるサチコは、昨日、肺手術でOntago犬病院から退院しました。手術はかなり成功しており、28歳のプードルは良好な状態にある。

ワンクリックで切り替えて、対訳形式で翻訳を進めることも可能です

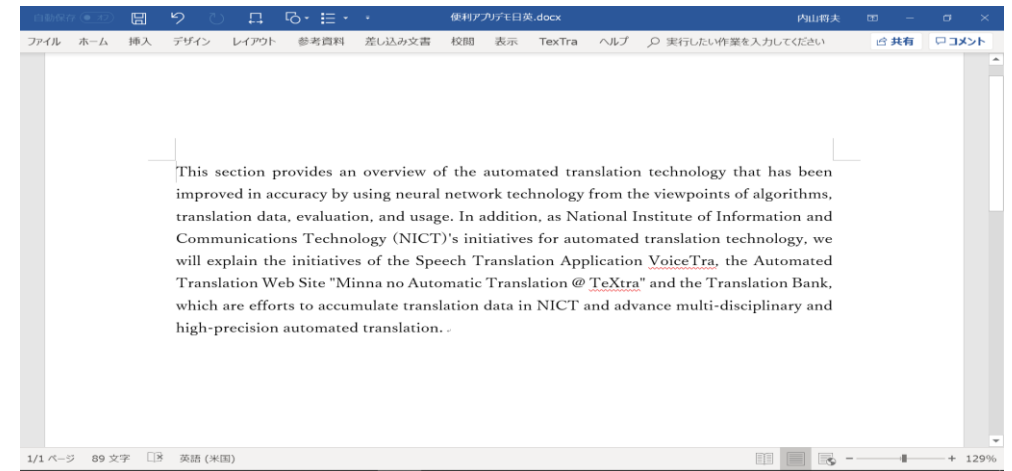
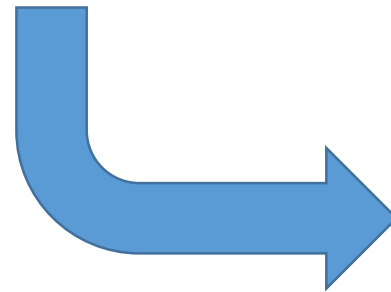


MT	Thank you very much for stopping by our site.
	サイトに立ち寄っていただき、ありがとうございます。
MT	This site provides you
	このサイトでは

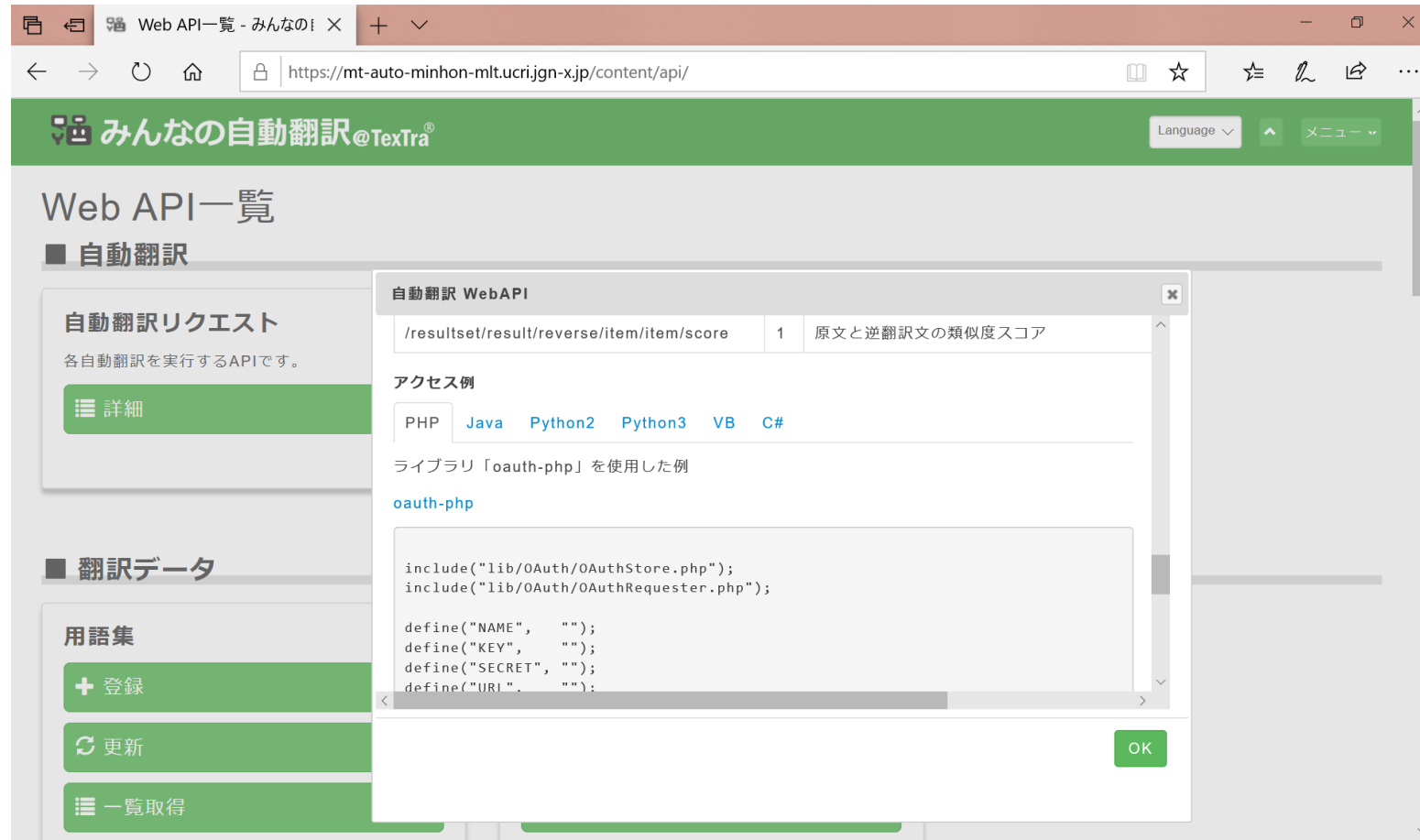
便利アプリで翻訳



- マイクロソフトオフィス
- ・ワード
 - ・エクセル
 - ・アウトルック
 - ・パワーポイント



Web APIで翻訳



みんなの自動翻訳@TexTra

Web API一覧

自動翻訳

自動翻訳リクエスト
各自動翻訳を実行するAPIです。

詳細

翻訳データ

用語集

- 登録
- 更新
- 一覧取得

自動翻訳 WebAPI

/resultset/result/reverse/item/item/score	1	原文と逆翻訳文の類似度スコア
---	---	----------------

アクセス例

PHP **Java** Python2 Python3 VB C#

ライブラリ「oauth-php」を使用した例

[oauth-php](#)

```
include("lib/OAuth/OAuthStore.php");
include("lib/OAuth/OAuthRequester.php");

define("NAME", "");
define("KEY", "");
define("SECRET", "");
define("URI", "");
```

OK

スマホアプリ等と
連携可能

実用化の一方、誤訳の影響の拡大可能性



自動翻訳の特性を理解した
適切な使い方が必要

MTの安全な使い方

前提：自動翻訳においては、自動翻訳結果の訳文の正確性は保証されません。



どのようにMTを使うかが重要

* MTサービスを利用するときには、その利用規約を参照してください。

自分がMTを使う場合

- 自動翻訳結果に間違いが含まれる可能性を理解する
- 理解したうえで、利用する
- 英語の特許等を日本語で読む
- 日本語のメールを英語に自動翻訳して編集する
- 両言語がわかる人がMTを使えば、非常に便利で安全な道具

M T の出力を他人が使う場合

- M T の出力を誰か（企業・ボランティア等）が修正する場合
 - 修正する企業等を信頼して利用する
- M T の出力をそのまま利用する場合
 - 何とかして翻訳を正確にしたい
 - 短い文章を入力する。
 - 主語・述語が明瞭な文章を入力する。
 - 翻訳結果を再度日本語に翻訳して、意味が一致するかを確かめる
 - 複数の自動翻訳エンジンの翻訳結果が一致するかを確かめる
 - 日本語入力を言い換えた日本語入力について上記を実施する
 - 入力文とM T 訳文の双方を利用者に提示する。
 - M T の出力のため、間違いが含まれる可能性があることを知らせる。

まとめ

