

# TOSHIBA

第15回 長尾賞受賞講演

## 特許庁 機械翻訳システムの開発

東芝デジタルソリューションズ株式会社

西本 俊之

2020.12.02



## エネルギー事業領域

東芝エネルギーシステムズ（株）

## 社会インフラ事業領域

東芝インフラシステムズ（株） 東芝テック（株）

東芝エレベータ（株） 東芝ライテック（株） 東芝キャリア（株）

## 電子デバイス事業領域

東芝デバイス&ストレージ（株）

## デジタルソリューション事業領域

### 東芝デジタルソリューションズ（株）

- インテグレーション  
(業務・各業種向けソリューション)
- IoTソリューション
- AI・アナリティクス
- セキュリティソリューション
- ICT基盤

東芝デジタル&コンサルティング（株）

## 第4回（2009年）長尾賞受賞

「The翻訳シリーズ」

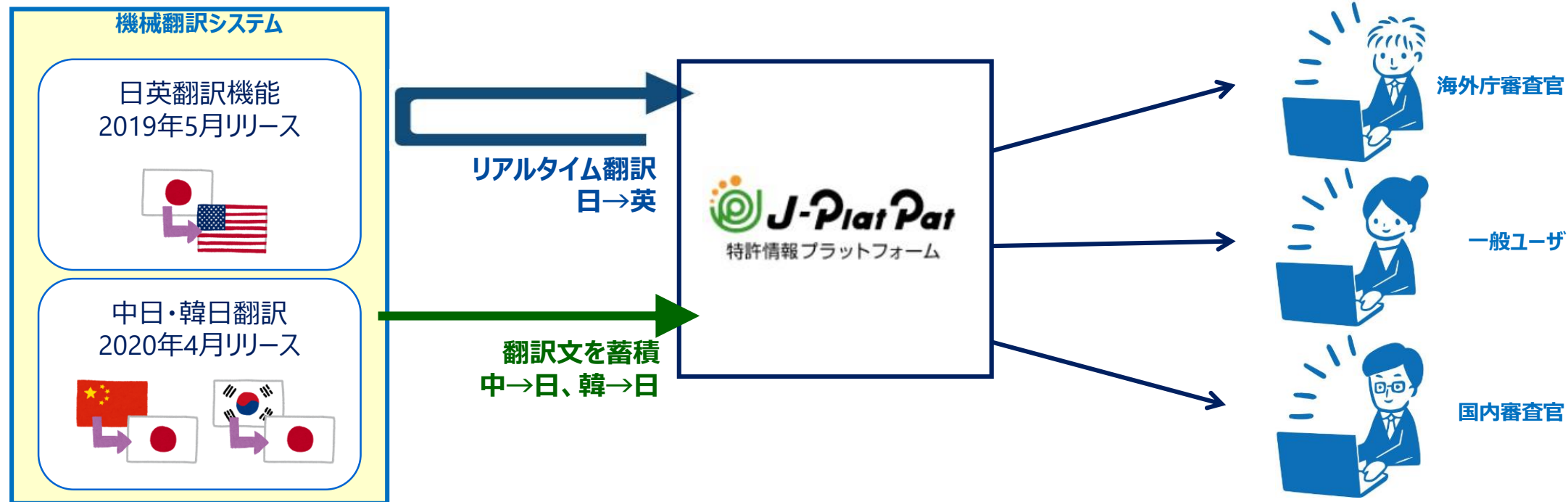


TOSHIBA

英日/日英翻訳ソフト「The翻訳 2009 プレミアム」

# 特許庁 機械翻訳システムとは

国内特許情報を海外に発信する際や、中国語、韓国語の文献を国内ユーザが参照する際に、機械翻訳文を提供するシステム。日英翻訳においては、ピーク時20,000アクセス/時の翻訳処理を行っている。



# ここで見られます。無料です。



<https://www.j-platpat.inpit.go.jp/>

# ここで見られます。無料です。

中日

English × 閉じる

J-Plat Pat  
特許情報プラットフォーム

文献表示

CN-A-111506713

文献表示画面の表示形式： テキスト表示  PDF表示

一次文献 | **機械翻訳**

検索キーワードのハイライトされている文字列:

詳細な説明 閉じる

(54) 【発明の名称】 表示制御システム、プログラムおよび記録媒体

【0001】 本願は、日本国特許出願2019-014418(出願日2019年1月30日)に基づき、この出願から優先的な利益を享受するものである。本願は、この出願を参照することによって、この出願の全内容を含む。

【0002】 【技術分野】

【0003】 本発明は、表示制御システム、プログラム及び記憶媒体に関するものである。

【0004】 【背景技術】

【0005】 ユーザからのダイアログを介してユーザからの質問に対して回答を出力するシステムがある。このシステムに関しては、どのような語句を会話システムに入力したとしても、求められた質問回答ペアリング(question-answer pair)を容易に入手することができるようにユーザが容易に推測できるようにすることが求められている。

【0006】 【発明の開示】

【0007】 本発明の実施形態は、文に対する応答をユーザが推測しやすい表示制御システム、プログラム及び記憶媒体を提供する。


【0008】 本発明の一実施形態によれば、表示制御システムは、処理部を備える。処理部は、ユーザにより入力された第1の問い合わせキーワードを含む第1の文を受け付ける。処理部は、前記第1の問い合わせキーワードを含む第1の質問と前記第1の質問に対する第1の回答とを含む第1の質問回答ペアの少なくとも一部と、前記第1の質問に含まれる第1の質問キーワードを含む第2の文とに対して、前記第1の問い合わせキーワードおよび前記第1の質問キーワード以外の単語を第1のパターンで表示し、前記第1のパター

※中日、韓日はバッチ翻訳のため、概ね2020年3月以降の発行分が、新しい機械翻訳システムで翻訳されたものです。

<https://www.j-platpat.inpit.go.jp/>

## 多種多様な書類に、それぞれ最適な翻訳エンジンを組み合わせ

公報







**特実**  
請求項・明細書・要約書

**意匠**

**商標**

**審判**

審査書類（300種類以上）



**拒絶理由  
通知書**

**意見書**

**手続補正書**

**特許査定**

主要な特許公報（請求項、明細書、要約書）だけでなく、意匠・商標・審判公報や、多様な審査書類。

それぞれ語彙や表現が異なる。

## 多種多様な書類に、それぞれ最適な翻訳エンジンを組み合わせ

拒絶理由通知書

特許出願の番号 特願2011-091×××  
起案日 平成25年 4月20日  
特許庁審査官 東芝 太郎 9901 4N00  
特許出願人代理人 大谷 翔平(外 2名) 様  
適用条文 第29条第1項、第29条第2項

この出願は、次の理由によって拒絶をすべきものです。これについて意見がありましたら、この通知書の発送の日から3か月以内に意見書を提出してください。

理 由

1. この出願の下記の請求項に係る発明は、その出願前に日本国内又は外国において、頒布された下記の刊行物に記載された発明又は電気通信回線を通じて公衆に利用可能となった発明であるから、特許法第29条第1項第3号に該当し、特許を受けることができない。

2. この出願の下記の請求項に係る発明は、その出願前に日本国内又は外国において、頒布された下記の刊行物に記載された発明又は電気通信回線を通じて公衆に利用可能となった発明に基いて、その出願前にその発明の属する技術の分野における通常の知識を有する者が容易に発明をすることができたものであるから、特許法第29条第2項の規定により特許を受けることができない。

記 (引用文献等については引用文献等一覧参照)

文献1には、蛍光物質がリンカーを介して結合した修飾ヌクレオチドが記載され、それをリアルタイムでの核酸検出システムに用いることも記載されている(特許請求の範囲、段落0002-0005)。そして、表1には各種のリンカーが、表2には調製された各種の修飾ヌクレオチドが記載されており、これらには請求項1、2の特定を満たすものが含まれている。

文献2にも、蛍光物質がリンカーを介して結合した修飾ヌクレオチドが記載され、それを核酸のシーケンシングに用いることも記載されている。よって、請求項1-10に係る発明は、いずれも当業者が容易に構成し得たものである。

書類ヘッダ

本文

そして、これらの発明を構成したことによる効果も当業者が予測し得ない程格別顕著なものとは認められない。  
ゆえに、請求項1-10に係る発明は進歩性を有しない。

引用文献等一覧

1. 国際公開第2009/0146××号  
(注) 法律又は契約等の制限により、提示した非特許文献の一部又は全てが送付されない場合があります。

先行技術文献調査結果の記録

調査した分野

IPC  
C12N15/00-15/90  
C12Q1/68

DB名  
WPI  
MEDLINE/CA/BIOSIS(STN)

この先行技術文献調査結果の記録は拒絶理由を構成するものではありません。

この拒絶理由通知の内容に関するお問い合わせ、または面接のご希望がございましたら下記までご連絡ください。  
特許審査第三部 生命工学 審査官 東芝 太郎  
TEL 03(3581)11XX 内線 9988  
FAX 03(3501)0XXX

部長/代理 審査長/代理 審査官 審査官補  
特許 一朗 東芝 太郎  
9988 9901

書類フッタ

一つの書類の中でも、部分によって特徴が変わる。

それぞれ語彙や表現が異なる。

## 多種多様な書類に、それぞれ最適な翻訳エンジンを組み合わせ

### 1 ニューラル翻訳エンジン (NMT)

特長

正確で流暢な訳文。  
言語特性の異なる日英、中日においても高品質。



担当

特許公報（要約、請求項、明細書）本文、審査結果本文

### 2 統計翻訳エンジン (SMT)

短文、体言止めなどに適する。  
言語特性が近い韓日において高品質。



韓日翻訳、「発明の名称」の翻訳

### 3 ルールベース翻訳エンジン (RBMT)

原文に忠実、定型文の翻訳に適する。  
語彙数に制約なし。  
大規模な学習データが不要。



書誌情報、意匠・商標・審決公報の翻訳



## 当社の自然言語処理の活用

### 1 レイアウト解析処理

原文のスペースや記号、キーワードを元に前後行の文章を結合または分割する。また、不要な飾り文字を除去する。



### 2 特殊表現退避処理

化合物表現や数値表現を一旦、原文から退避し、別々に翻訳を行った上で再度結合する。



### 3 長文分割

長文を分割し、別々に翻訳を行った上で、再度結合する。できるだけ文意を損なわないよう最適な位置を判定して分割。



### 4 誤訳判定

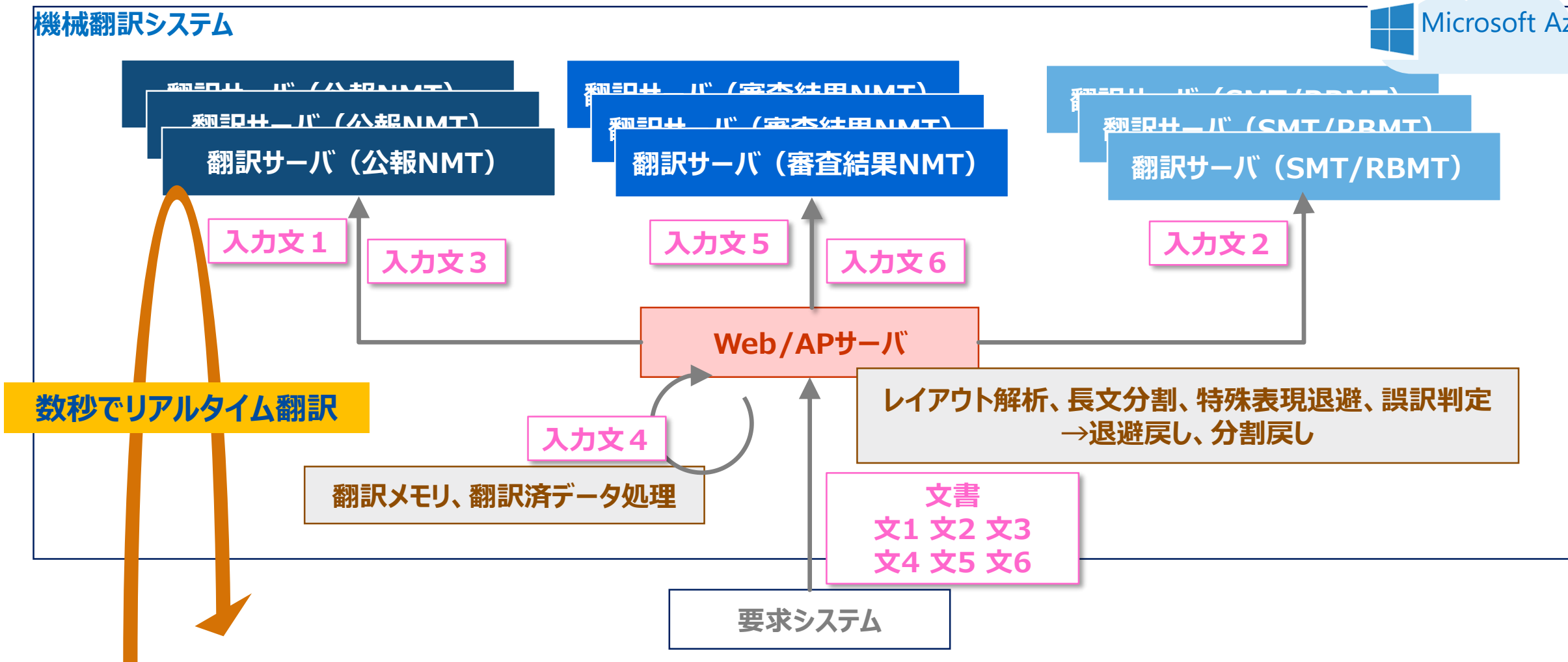
NMT特有の訳抜け・湧き出しの発生を検知し、自動的に再翻訳または他エンジンへの振り替えを行う。



# 機械翻訳システムの特徴③

## 大規模パブリッククラウドによる分散処理でリアルタイム翻訳を実現。

機械翻訳システム



## 終わりに

### ニューラル翻訳とアプリケーションの融合


ニューラル翻訳は既に実用化レベル。更にうまく扱うには、前後処理アプリケーションで工夫することが効果的。

### 特許翻訳の最適解の追求

ニューラル翻訳、クラウドサービスの技術進展は目覚ましい。今後も、特許庁、NICTと連携し、最新の技術を取り入れて、特許分野の機械翻訳ソリューションの最適解を追求していきたい。

### 他分野へも

本システムの開発で培った技術やソリューションを活かして、他の分野においても機械翻訳の更なる実用化に貢献していきたい。

A hand is shown from the bottom, holding a glowing, translucent globe. The background is a warm, golden sunset or sunrise over a landscape, with a rainbow visible in the distance. The overall mood is hopeful and forward-looking. On the left side, there are three overlapping geometric shapes: a blue rectangle at the top, a grey triangle in the middle, and a red triangle at the bottom.

東芝グループ経営理念

人と、地球の、明日のために。