

## 特許情報の翻訳に関する取組について

2024 知財・情報フェア&コンファレンス

2024年10月2日

特許庁特許情報室長 本多 仁



- 1 特許情報提供の目的
- 2 翻訳を活用した特許情報提供
- 3 機械翻訳に関する取組
- 4 令和5年度 機械翻訳に関する調査研究

## 1. 特許情報提供の目的

2

## はじめに

➤ 特許庁の役割の一つとして、「産業財産権情報（特許情報）提供の拡充」。

ホーム > 特許について > 特許庁の役割

### 特許庁の役割

特許庁では、ユーザーのニーズに応えるために以下の方向で取組んでいます。

産業財産権の権利、権利取得、権利行使を適切に産業財産権制度と結びます。また、権利、デザイン、意匠などの知的財産の伝播を促進・活用し、産業の発展に寄与することを目的としており、特許庁が所管しております。産業競争の発展と国民生活の向上のために、産業財産権制度は、21世紀の日本にとって、最も重要な要素の一つです。

特許庁は、創発部、審査部、審査第一課、権利部から組織され、(1)産業財産権の適切な付与、(2)産業財産権制度の企画立案、(3)国際的な制度調和と向上を図るための促進、(4)産業財産権制度の奨励、(5)中小企業・大学等に対する支援、(6)産業財産権情報提供の拡充等、我が国産業の発展に向けた取組を積極的に進めています。

**(1) 産業財産権の適切な付与**

世界各國が取り組んでいる特許法務などの取組について、科学的観点、法律知識などを、審査結果に対する不確実性については、効力裁判所に代わって第一審としての判断

**(2) 産業財産権制度の企画立案**

未来を切り拓く「知的財産立憲」の取組に向け、1)経済・経済特許庁、2)少額出願・共有などによる権利取得の促進、3)地域ブランドなどの醸成、4)権利行使の円滑な

**(3) 国際的な制度調和と向上を図るための推進**

国際調和を目的とした産業財産権制度の取組を推進し、あわせて、我が国企業人の海外での円滑な権利取得や権利行使を支援するため、日・米・欧米連合協力や中露・韓国を軸とした協力協力、向上協力（審査協力、人材交流など）、特許審査ハイウェイ(PHW)の推進、権利部・権利部関係の強化など、積極的に国際活動に取り組んでいます。

**(4) 産業財産権制度の奨励**

産業財産権制度の企画立案や国際交渉の取組などを支援し、特許法、特許法などの関係法令の改正、審査基準などの見直しを進めています。

**(5) 中小企業・大学等に対する支援**

中小企業、大学など取組の強い産業財産権活用を促すため、手続料の負担軽減、特許費用の支援、特許管理体制の強化と支援、審査官連携の促進など、様々な取組を行っています。

**(6) 産業財産権情報提供の拡充**

多様なユーザーニーズに応えるため、インターネット公開の取組などを進めて、産業財産権情報提供の拡充に取り組んでいます。

【更新日 2022年11月4日】

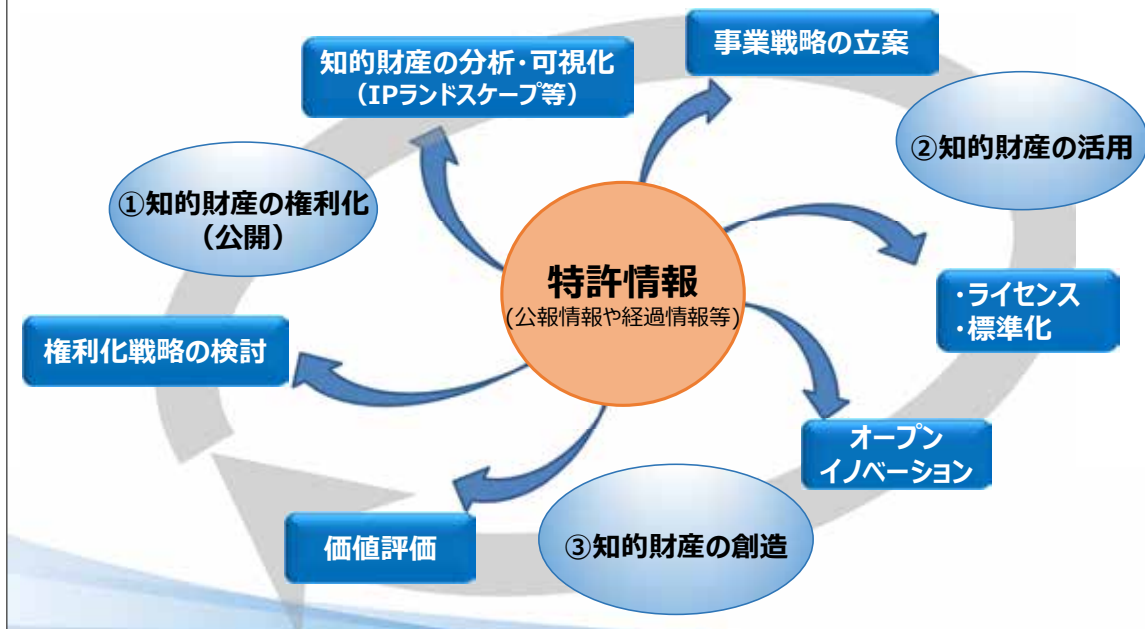
3

16

## 知財エコシステムと特許情報

- 特許情報は知財エコシステムにおける促進剤。

特許情報・特許・実用新案・意匠・商標の出願・権利化に伴って生み出される情報(公報情報や経過情報等)



4

## 特許情報施策の全体方針

- 産業財産権情報利用推進委員会 (長官私的懇談会) 報告 (平成15年3月)

国と民間とのベストミックスで一般ユーザーの多様なニーズに応じていく。

### 国の役割

正確で基本的な一次情報の提供

### 民間の役割

一次情報に高い付加価値をつけた情報の提供

- 産業構造審議会 知的財産分科会 情報普及活用小委員会 (平成28年5月)

我が国ユーザーが享受するサービスの質が全体として世界最高水準となるように特許情報サービスを提供していき、特許情報を広く普及していくための基盤を引き続き整備していくべき。

### 考慮要素

グローバル化、ITの進展  
海外庁のサービスの状況  
民間事業者のサービスの状況  
我が国ユーザーの要望

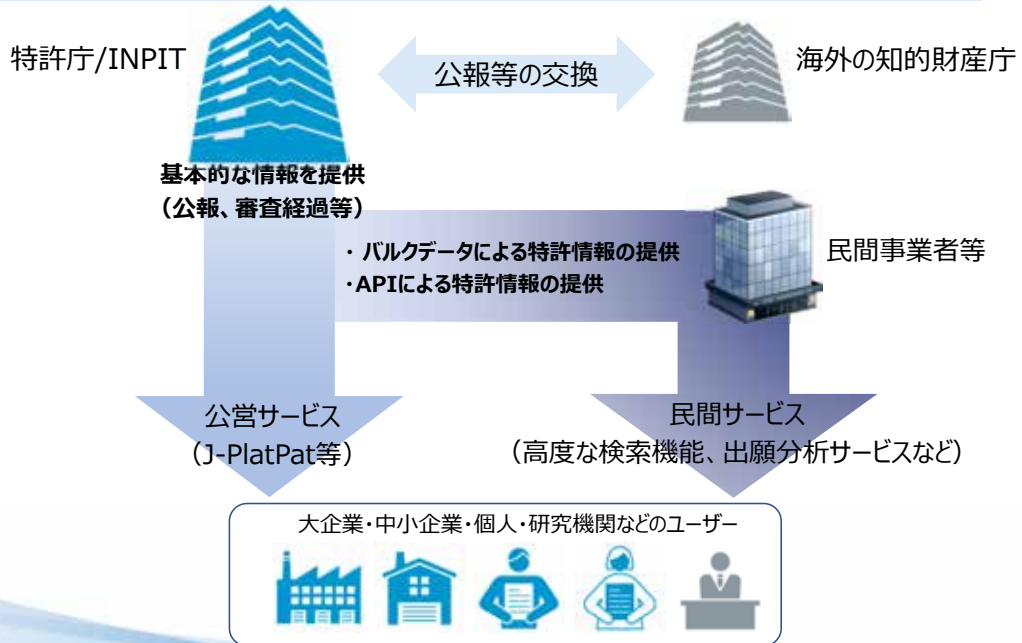
### 環境整備の方向性

特に中小企業等が  
容易に特許情報等を利用できるようにする  
  
海外庁の審査官が  
我が国審査結果を参酌しやすくする

5

## 官民の役割イメージ

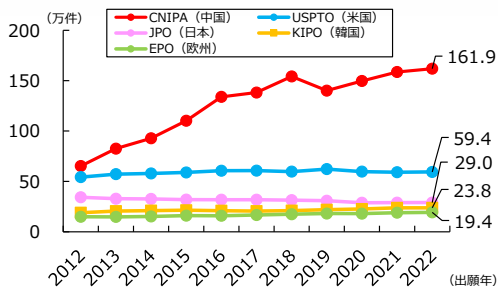
➤ 国が正確で基本的な一次情報を提供し、民間事業者が高度なサービスを提供。



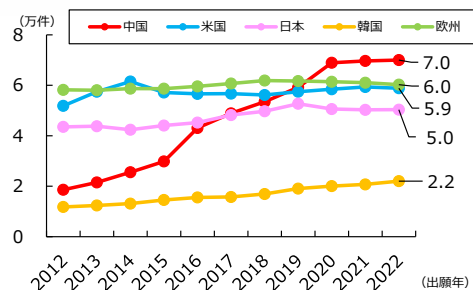
## 2. 翻訳を活用した特許情報提供

## 出願の動向

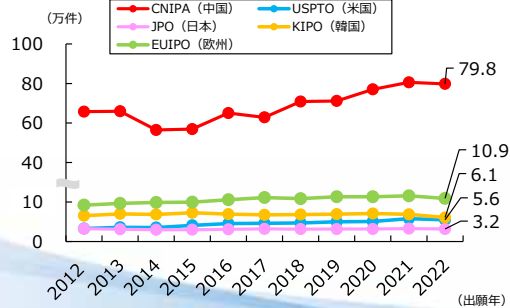
日米欧中韓における特許出願件数の推移



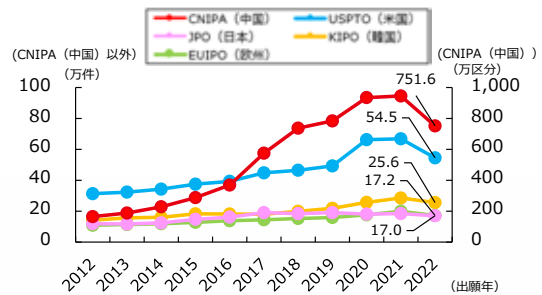
出願人居住国別のPCT出願件数の推移



日米欧中韓における意匠登録出願件数の推移



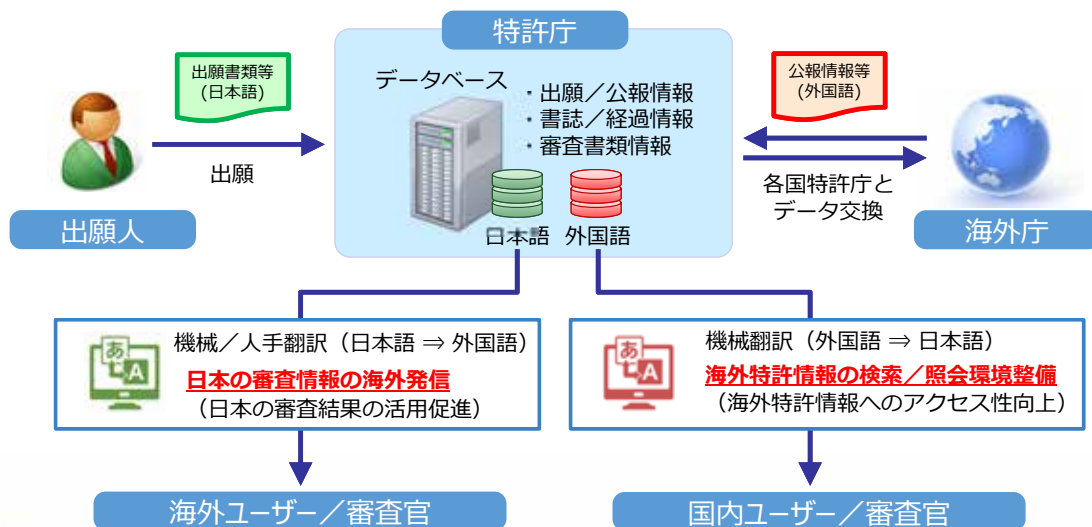
日米欧中韓における商標登録出願件数の推移



\*2023年9月時点の取得可能データに基づき作成

## 翻訳を活用した特許情報提供

- 日本企業等の海外権利取得を支援すべく、日本の審査情報を海外への発信。
- 日本ユーザーの海外特許情報へのアクセス性を向上。



## 特許情報提供の主なインターフェース

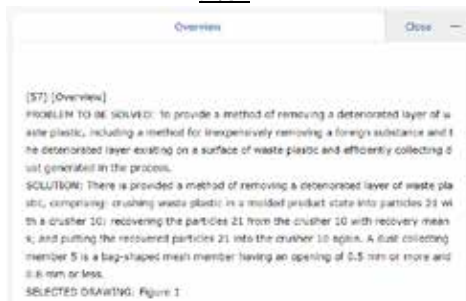
- 翻訳した特許情報は、J-PlatPat、OPD、FOPISER等を通じて对外提供。



## 日本の審査情報の海外発信

- J-PlatPatにおいて、公報や審査書類をリアルタイム英訳にて提供。
- また、公開公報の書誌事項、要約、代表図面を人手により英訳、編集した英文抄録 (Patent Abstracts of Japan: PAJ) も閲覧可能。

### 公報



### 拒絶理由通知

This application should be refused for the reasons mentioned below if the applicant has any opinion(s) against the reasons, a written opinion should be submitted within 30 days from the date on which this notification was dispatched.

Reason:

1. (novelty) The claimed invention(s) in the each claim listed below of this patent application shall not be granted a patent under the provision of Patent Act Article 29 (1)(ii) for the reason that the claimed invention(s) has/have been deemed to be identical with the invention(s) described in the distributed publication(s) listed below or made available to the public through electric telecommunication lines in Japan or other foreign countries prior to the filing of the patent application.

2. (inventive step) The claimed invention(s) listed below of this patent application shall not be granted a patent under the provision of Patent Act Article 29 (2) for the reason that the claimed invention(s) could have easily been made by persons who have common knowledge in the technical field to which the claimed invention(s) pertains, on the basis of the invention(s) described in the distributed publication listed below or made available to the public through electric telecommunication lines in Japan or other foreign countries prior to the filing of the patent application.

Note: [For the cited documents, please refer to the list of cited documents below.]

— Reason 1 (novelty), Reason 2 (inventive step)

## 海外特許情報の検索／照会環境整備

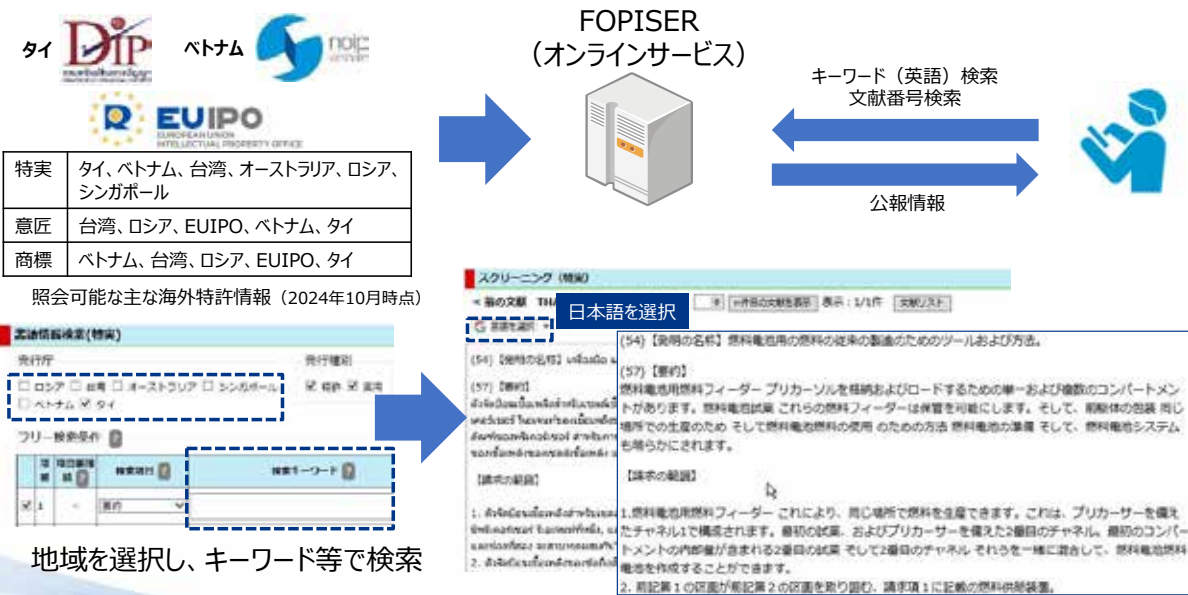
➤ J-PlatPat及びFOPISERにおいて、機械翻訳を活用した海外特許情報を提供。

国・地域 (言語)	US (英語)	EP (英語／独語 ／仏語)	CN (中国語)	KR (韓国語)	WO (英語)	FOPISER提供国 (タイ語／ベトナム語 ／ロシア語等)
提供サービス	J-PlatPat	J-PlatPat	J-PlatPat	J-PlatPat	J-PlatPat	FOPISER
翻訳対象	要約	要約	全文	全文	-	要約 (検索) 全文 (照会)
翻訳手法	特許特化エンジンによる機械翻訳 (R5fy以降)	特許特化エンジンによる機械翻訳 (R5fy以降)	特許特化エンジンによる機械翻訳	特許特化エンジンによる機械翻訳	-	Google翻訳
備考	原文に対して、英語検索可	原文に対して、英語検索可			原文に対して、英語検索可	

12

## 外国特許情報サービス (FOPISER) の概要

- 新興国の特許情報を日本語ユーザーインターフェースで提供 (2015年8月より開始)。
- 文献番号による照会、キーワード (英語) ・分類による簡易検索が可能。



機械翻訳の利用により、日本語での閲覧が可能。 13

## 外国特許情報サービス（FOPISER）での中国審決の提供

- 中国の特許・実用新案に関する審決の和訳を提供（書誌情報検索で照会）。

外国特許情報サービス (FOPISER)

特許庁 審決番号索引照会(中国審決)

審決番号索引照会(中国審決)

出願・審決番号 番号クリア

審決決定日 201101 ~

出願日 201101 ~

種別  不特審判  特許審判

決定の言語 日本語で入力

照会本文 (A=中国專利法, B=実用新案) A-022-03

発明の名称 日本語で入力

請求人 中国語で入力

照会本文 中国語で入力

検索 部分一致検索 条件クリア

照会

入力番号フォーマット

- 出願番号 例: nnnnnnnn.x
- 審決番号 例: yyyyyyyyyy.n

[https://www.foreignsearch2.jpo.go.jp/cnapp/cnapp\\_bangousyoukai.php](https://www.foreignsearch2.jpo.go.jp/cnapp/cnapp_bangousyoukai.php)

14

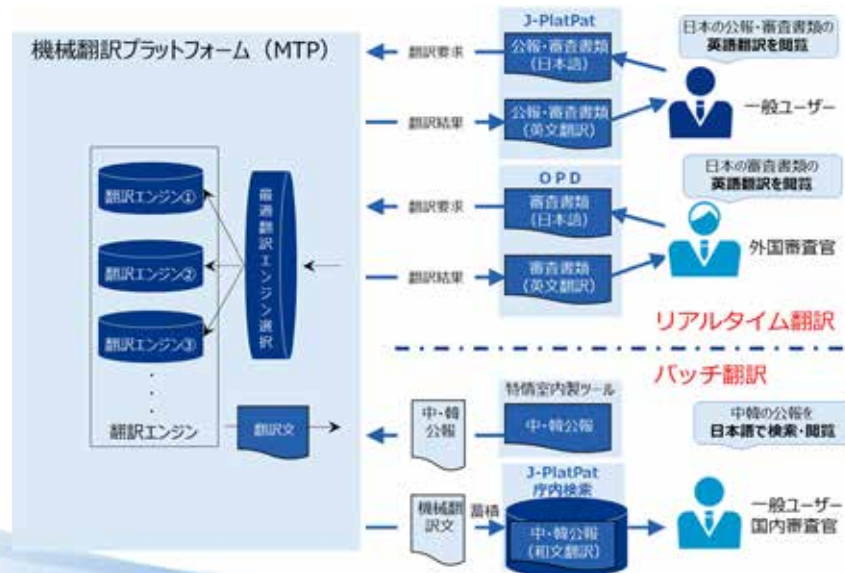
## 3. 機械翻訳に関する取組

15



## 機械翻訳プラットフォーム（MTP）の概要

- 機械翻訳プラットフォーム（MTP）により各種機械翻訳を実施。
- MTPは複数の翻訳エンジンを搭載しており、令和元年5月に日英翻訳機能、令和2年4月に中日・韓日翻訳機能をリリースし、現在運用中。



16

## MTPにおける翻訳エンジン

- MTPの翻訳エンジンは、NMTをメインに利用しつつ、SMT及びRBMTを併用。
- 翻訳前後の文章に対する各種処理等を通じて、翻訳品質を最適化。

### MTPにおける翻訳エンジンの使い分け

日英翻訳	翻訳対象	メイン翻訳エンジン
	特許公報（明細書等）	NMT（公報用カスタマイズ）
	特許公報（発明の名称）	SMT
	審査書類等	NMT（審査書類用カスタマイズ）
	書誌情報・特許以外の公報 等	RBMT
中日翻訳	翻訳対象	メイン翻訳エンジン
	特許公報（明細書等）	NMT
	特許公報（発明の名称）	SMT
韓日翻訳	翻訳対象	メイン翻訳エンジン
	特許公報	SMT

17

## 4. 令和5年度 機械翻訳に関する調査研究

18

### 調査研究の背景

- 現状、ASEAN言語等の特許文献については、汎用翻訳エンジンを採用。
- 海外特許情報の機械翻訳について、日英中韓以外の言語への拡大を検討。

国・地域 (言語)	US (英語)	EP (英語/独語 /仏語)	CN (中国語)	KR (韓国語)	WO (英語)	FOPISER提供国 (タイ語/ベトナム語 /ロシア語等)
提供サービス	J-PlatPat	J-PlatPat	J-PlatPat	J-PlatPat	J-PlatPat	FOPISER
翻訳対象	要約	要約	全文	全文	-	要約 (検索) 全文 (照会)
翻訳手法	特許特化エ ンジンによる機械 翻訳 (R5fy以降)	特許特化エ ンジンによる機械 翻訳 (R5fy以降)	特許特化エ ンジンによる機械 翻訳	特許特化エ ンジンによる機械 翻訳	-	Google翻訳
備考	原文に対して、 英語検索可	原文に対して、 英語検索可			原文に対して、 英語検索可	

19

## 調査研究の目的と概要

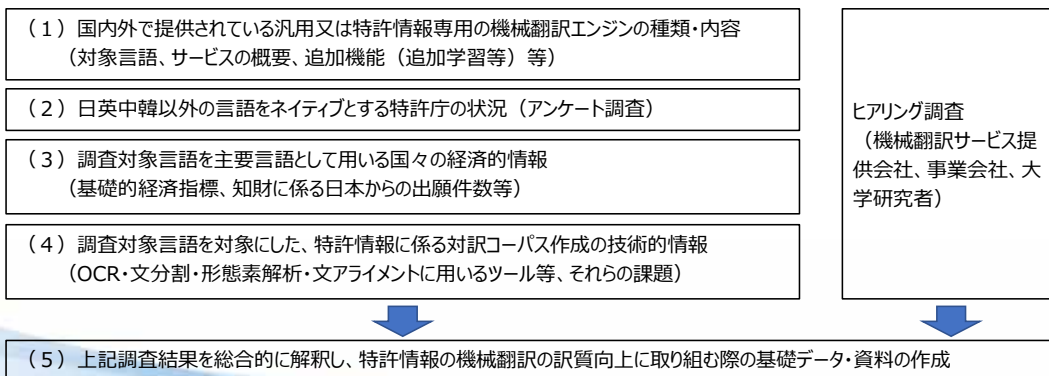
### <目的>

- ✓ ASEAN、中南米、インド等への日本からの進出・投資が増加傾向にある現状を踏まえ、**英中韓以外の言語**を対象として、特許情報の機械翻訳の訳質向上に取り組む際の基礎データ／資料を作成。

### <対象言語及び国>

- ✓ スペイン語、ポルトガル語、イタリア語、インドネシア語、マレー語、ベトナム語、タイ語、アラビア語、ヒンディー語、トルコ語の10言語
- ✓ スペイン、メキシコ、ポルトガル、ブラジル、イタリア、インドネシア、マレーシア、ベトナム、タイ、サウジアラビア、インド、トルコの12か国

### <調査概要>



20

## 汎用的な機械翻訳エンジンについて

- 汎用的な機械翻訳エンジンは、幅広い言語に対応しており、調査対象言語をカバーしているものが多数。

調査対象	対応言語数	スペイン語	ポルトガル語	イタリア語	インドネシア語	マレー語	ベトナム語	タイ語	アラビア語	ヒンディー語	トルコ語
民間エンジン・サービス	134言語	○	○	○	○	○	○	○	○	○	○
	128言語	○	○	○	○	○	○	○	○	○	○
	30言語	○	○	○	○	×	×	×	○	×	○
	45言語	○	○	○	○	○	○	○	○	○	○
	28言語	○	○	○	○	○	○	○	×	○	○
	14言語	○	○	○	○	×	○	○	×	×	×
	106言語	○	○	○	○	×	○	○	○	○	○
23言語	○	○	○	○	○	○	○	○	○	○	

※具体的なエンジン・サービス名は、報告書をご参照ください。

21

## 特許情報に特化した機械翻訳エンジンについて

- ▶ 特許情報に特化した機械翻訳エンジンは、対応言語数は少なく、調査対象言語の一部のみカバー。

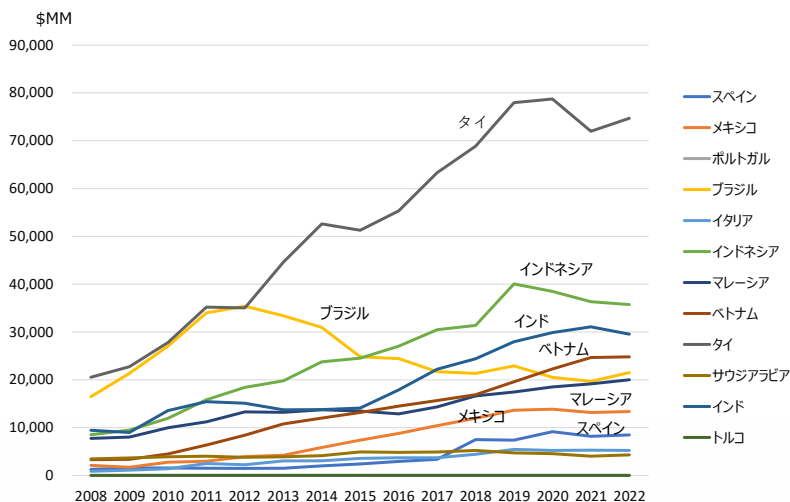
調査対象	対応言語数	スペイン語	ポルトガル語	イタリア語	インドネシア語	マレー語	ベトナム語	タイ語	アラビア語	ヒンディー語	トルコ語
海外知財庁のエンジン・サービス	32言語	○	○	○	×	×	×	×	×	×	×
	3言語	×	×	×	×	×	×	×	×	×	×
	13言語	○	○	○	×	×	×	×	○	×	×
民間エンジン・サービス	6言語	×	×	×	×	×	×	×	×	×	×
	9言語	○	○	×	×	×	×	×	×	×	×
	6言語	×	×	×	×	×	×	×	×	×	×
	6言語	×	×	×	×	×	×	×	×	×	×
	6言語	×	×	×	×	×	×	×	×	×	×

※具体的なエンジン・サービス名は、報告書をご参照ください。

## 経済的指標（日本→対象国）

- ▶ 対外直接投資残高（日本→対象国）について、上位国（2022年）は、タイ（747億ドル）、次いでインドネシア（357億ドル）、インド（295億ドル）と続く。

対外直接投資残高の推移（日本→対象国）

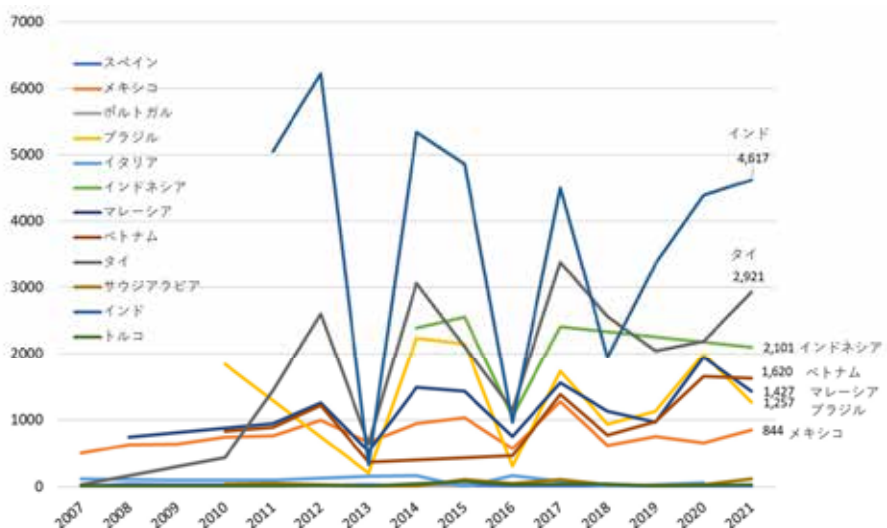


出所：JETRO直接投資統（<https://www.jetro.go.jp/world/japan/stats/fdi.html>）より作成。

## 特許出願動向（日本→対象国）

- 特許出願件数（日本→対象国）について、上位国（2021年）は、インド（4,617件）、次いでタイ（2,921件）、インドネシア（2,101件）と続く。

特許出願件数（日本→対象国）の推移（単位：件数）



出所：特許行政年次報告書(WIPO IP Statistics Data Center)より作成。

24

## 特許出願動向（PPH：日本→対象国）

- PPH申請件数（日本→対象国）について、直近10年の平均で上位国は、タイ（363件）、インドネシア（194件）、マレーシア（161件）、メキシコ（101件）、ブラジル（67件）であった。

PPH申請件数（第1庁：日本、第2庁：対象国）（単位：件数）

	スペイン	メキシコ	ポルトガル	ブラジル	イタリア	インドネシア	マレーシア	ベトナム	タイ	サウジアラビア	インド	トルコ
2011		1										
2012	1	21										
2013		39				25						
2014		50					17		108			
2015	0	85				265	122		168			
2016	1	111	0			137	166	0	390			
2017	0	90	0			0	178	0	455			
2018	5	200	0	41		251	247	0	463			
2019	1	169	0	83		373	252	0	545			0
2020	0	101	0	71		112	224	0	411	0	0	2
2021	0	78	0	66		0	84	91	348	0	0	0
2022	0	89	0	72			0		381			1

注：特許庁PPHポータル「統計情報」では、「統計情報の提供がない知財庁、集計期間中にPPH申請が0件であった知財庁の統計情報は掲載されていません。」と記述されている。PPH申請件数データにて、①「国名あり、PPH申請件数なし」は原則、PPH申請件数0と見做して上表に記入。②「国名なし」は空欄とした。

出所：特許庁 PPHポータル PPH申請件数（2011年～2022年）より作成。

25

## ヒアリング内容の抜粋（事業会社）

### <ASEAN/南米/中東地域に進出する事業会社>

- ✓ FTO調査（侵害予防調査）については、日米欧中等の主要国以外を調査する場合の課題として、**商用データベースにおける公報の収録範囲**が十分であるか、**英語未対応の現地庁サイト**の使いにくさ、**機械翻訳の精度**が課題となる。
- ✓ ASEAN地域については、IPランドスケープ調査を行い、現地のプレイヤーや顧客候補企業の調査を行う場面がある。
- ✓ 特許出願・権利化にあたって、日本語と現地語の翻訳においては、作成された翻訳が真に正しいのか検証しづらいことが課題。翻訳精度の関係で**日本語→英語→現地語の順に翻訳**することがある。**翻訳費用の増大**に繋がるため、予算面での課題が生じる。
- ✓ 権利の維持状況を調査する必要が生じるが、著名なデータベースであってもステータスが収録されていないことが多く、現地特許庁のシステムにアクセスする必要に迫られることがあるが、**言語の壁も含めて利用のハードルが非常に高く、目的とする知財にたどり着けない**ことも多い。

26

## ヒアリング内容の抜粋（翻訳サービス提供企業）

### <特許情報の機械翻訳サービスの提供企業>

- ✓ 特許情報はその文章や用語が一般的な文章と異なる部分も多く、**汎用機械翻訳エンジンを用いた場合の翻訳精度への信頼は必ずしも得られていない**。
- ✓ 今回の調査対象言語の特許情報に対する**顧客からの翻訳ニーズは、潜在的には存在している**といえ、顧客ニーズが**顕在化して高まりを見せれば、民間レベルでのサービス提供・開発機運が高まる**可能性もあるといえる。

27

## 特許情報に係る対訳コーパス作成の技術的情報

➤ 以下の観点で、調査対象の10言語について、特許公報から対訳コーパス（※）を作成するための技術的情報を調査。

- ① パテントファミリーデータの取得元およびデータの状態
- ② 多言語OCRツール
- ③ 文分割ツール
- ④ 形態素解析（単語分割）ツール
- ⑤ アライメントツール
- ⑥ 対訳辞書作成ツール
- ⑦ パテントファミリーデータからの対訳コーパス作成手順
- ⑧ コーパス試作、精度評価、作成可能な対訳コーパス量
- ⑨ その他の関連技術

※「対訳コーパス」とは、異なる言語の文と文を対訳にしてまとめた文書のあつまり。特許公報由来の対訳コーパスを、機械翻訳エンジンに学習させることで、特許情報に特化したエンジンを作成可能。

28

## 特許情報に係る対訳コーパス作成の技術的情報

➤ （1）－（3）までの調査結果全体を踏まえて、インドネシア語、タイ語、ベトナム語について、対訳コーパスの試作を実施して、課題を洗い出し。

### <インドネシア語>

- ✓ 一定の範囲の公報で、明細書を含む全文テキストデータが入手可能。
- ✓ 文分割・文アライメントの手法については改善の余地がある。

### <タイ語>

- ✓ 取得可能な公報データが画像データであってOCRを要するが、得られるテキストデータの精度が悪い。
- ✓ タイ語の言語的な特性から、文分割の精度が悪い。
- ✓ 文アライメントについては、上記状況なので評価できない。

### <ベトナム語>

- ✓ 明細書を含む全文テキストデータが取得可能な期間が限られており、十分なテキストデータが存在するとはいえない。
- ✓ 文分割・文アライメントの手法については改善の余地がある。

29

## まとめ（総合分析より抜粋）

### ＜現状の機械翻訳エンジンの観点＞

- ✓ 本調査での対象言語は、現状の特許特化型機械翻訳エンジンでは対応していない言語も多く、**汎用機械翻訳エンジンで対応せざるを得ない**状況。
- ✓ 一方、汎用機械翻訳エンジンを用いた場合の翻訳精度への**信頼は必ずしも得られていない**。
- ✓ ここで、対訳コーパスの作成手法に関して機械翻訳の活用や文のベクトル化ツールを活用するなど、**自然言語処理の技術の進展**によって、ASEAN言語等の言語資源が比較的少ない言語であっても特許特化型機械翻訳を**実現できる可能性が高まってきている**。

### ＜経済的情報の観点＞

- ✓ 日本からの特許出願件数や投資状況において**上位の国は、機械翻訳サービスの潜在的なニーズが高い国**と考えられる。
- ✓ また、ヒアリング調査の結果、FTO 調査（侵害予防調査）や先行技術調査にて、ASEAN等を対象とする場合、十分なデータベースが整っていないことや、**既存の翻訳エンジンでの翻訳精度が十分でない**点が課題として挙げられた。

### ＜対訳コーパス作成の技術的情報の観点＞

- ✓ インドネシア語、タイ語、ベトナム語について、対訳コーパスを試作した結果としては、現状の言語資源・技術において、**十分な対訳コーパスを作成できる可能性が高いのは、インドネシア語**といえる。
- ✓ 試作においていくつかの改善点を見出したので、今後、上記の言語の対訳コーパスを自動的に作成する場合には、**これらの改善点も活かす**ことができる。

30

## （参考）本調査研究の報告書について

- ✓ 本調査研究の報告書は、以下の特許庁ウェブサイトで公開しております。

[https://www.jpo.go.jp/system/laws/sesaku/kikaihonyaku/kikai\\_honyaku.html](https://www.jpo.go.jp/system/laws/sesaku/kikaihonyaku/kikai_honyaku.html)

The screenshot shows the official website of the Japan Patent Office (特許庁). The page title is '機械翻訳に関する調査報告書' (Investigation Report on Mechanical Translation). The main text discusses the patent office's efforts to improve foreign language patent search and the availability of machine translation services. It mentions that while machine translation is available, the quality is not always reliable, and the patent office is working to improve it. The report is dated '令和5年度' (Reiwa 5th year, 2023).

31



ありがとうございました

---

特許庁 特許情報室

[PA0630@jpo.go.jp](mailto:PA0630@jpo.go.jp)

