

五大特許庁「相互機械翻訳プロジェクト」の概要と韓国特許庁の機械翻訳活用

スマートワークス株式会社代表取締役 **酒井 美里**

PROFILE

1991年奈良女子大学卒。同年セイコーエプソン（株）入社。2005年より現職。

✉ info@1smartworks.com

TEL 03-6868-4924

1 はじめに

近年、英語以外の言語による特許出願の増加が著しい。とりわけ中国特許出願の伸びが顕著であり、2000年の約5万1千件から、2010年には約39万1千件と非常に大きな増加率を示している。出願の増加と同時に特許情報の世界でも、中国語・韓国語など「非英語・非日本語」の文献が存在感を増す、という状況につながっている。

特許審査に関する取り組みにおいては、以前より日米欧の三極（Trilateral）による協力が行われてきたが、2007年からは中韓特許庁を加えた「五大特許庁」による協力がスタートしている。現在、五庁で取り扱う特許出願件数は、世界の特許出願の約80%を占めるとされる。また、五カ国全体の出願件数の約30%にあたる約42万件は、パテントファミリーの関係にあるとされることから、この五庁が協力して取り組むことにより、各庁における特許審査の質と審査効率が向上することが期待されている（資料1）。

五大特許庁の活動として、これまでに「10の基礎プロジェクト」の立ち上げが合意され、各庁が2つの基礎プロジェクトを担当して取り組みをスタートしている。このうち非日本語・非英語文献の機械翻訳に関連する「相互機械翻訳」プロジェクトは、韓国特許庁（KIPO）が担当している。

本稿の前半では、韓国特許情報と、韓国特許庁内のシステムについて概説する。後半では五大特許庁の基礎プロジェクトのうち、KIPOが担当する「共通トレーニン

グポリシー」と「相互機械翻訳プロジェクト」を紹介する。

2 韓国特許情報の概要

韓国特許の出願件数は1990年代から増加を続け、近年は横ばい傾向にある。1998年の特許出願件数は約7万5千件、2009年の特許出願件数は約16万3千件であった。同様に外国人／内国人比率を出願件数ベースで見ると、1998年の韓国特許出願における外国人比率は約33%、2009年の外国人比率は約22%である。同時期の日本特許出願について比較すると、1998年の外国人比率は約10%、2009年は約15%であった。（資料2）

韓国特許庁（KIPO）では、1990年代半ばより庁内向けシステム及び一般向けシステムの両者において、機械翻訳データの整備・提供を進めてきた。現在 KIPO

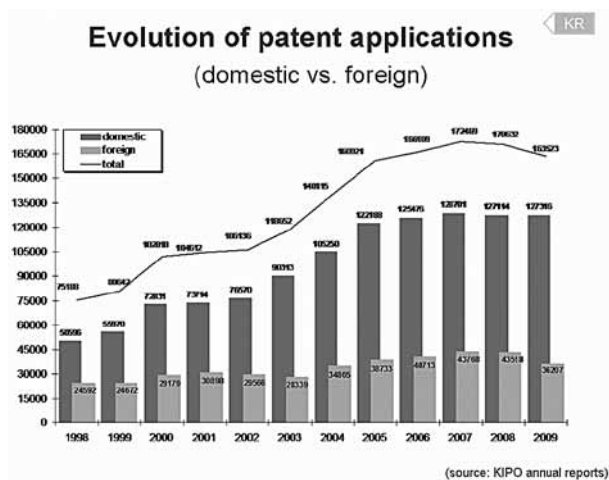


図1 韓国特許出願件数と内国人／外国人内訳の推移
（出典：EPO／データ：KIPO）

の特許文献データベースは、3系統が運用されている。KIPO 審査官向けのKIPO Patent Search System(通称 KOMPASS)、外国審査官向けの Korean Patent Information Online Network (通称 K-PION)、一般利用者向けの KIPRIS である。(この他にも出願管理用のシステム、意匠、商標の検索システム等が存在する。)(資料3、4)

KIPO では 1990 年代には海外特許データ及び機械翻訳に注目してきた。2000 年には KIPO 審査官向けシステムで、日韓機械翻訳データが搭載された。収録範囲は日本語特許文献及び非特許文献であった (J2K Translation)。2003 年には KIPRIS に米国・欧州・日本の特許データを追加し、一般ユーザー向けに提供。2005 年からは海外特許庁向けに韓英特許データの提供を開始した (K-PION)。2007 年以降は毎年のように検索機能の改良・拡充や収録国の追加を実施している。

現在の KIPO の特許文献データベースと対象ユーザー(提供先)、機械翻訳との関係は下記の通りである。(資料5、6)

表1 KIPO の特許文献データベースと対象ユーザー、機械翻訳の内容

システム名	対象ユーザー	対応言語	翻訳方式
KOMPASS	KIPO 審査官	英語・日本語→韓国語	辞書 + コーパス
K-Pion	海外庁審査官	韓国語→英語	辞書 + コーパス
KIPRIS	一般向け	英語・日本語→韓国語 韓国語→英語 (K2E サービス)	辞書 + コーパス

KIPO のシステムで利用される翻訳エンジンは、韓英／英韓はシリウスソフト社又は ETRI 社製、日韓翻訳はシリウスソフト社製、とされる。いずれの翻訳エンジンでも、従来型の辞書ベースの検索に加え、近年はコーパス方式の併用を進めることで、新語への対応や検索スピード・品質の向上を図りつつある。(特許翻訳におけるコーパス方式の応用に関しては(資料6、7)もご参照いただきたい。)

3 韓国特許庁 (KIPO) が運用する特許検索システムの概要

3.1 KOMPASS

KOMPASS は、KIPO 審査官向けの特許／非特許文献システムである。韓国、日本、アメリカ、欧州の特許公報全文を収録しており、日本語と英語の文献は機械翻訳されている。韓国語／英語／日本語での検索が可能である。また、検索結果としては、韓国語と原文(ヒット箇所に対応する英語・日本語)の表示が可能である。併せて IEEE、サイエンスダイレクト等の電子ジャーナル(非特許文献)のメタ検索が可能である。

図2に KOMPASS の画面例を示す。画面例では米国特許がヒットしているが、ワンクリックで韓国語全文表示が可能、とのことである。また、検索項目としてはキーワード(原語・翻訳)の他、US クラス、ECLA(欧州分類)、FI 記号、F タームなどが利用可能である。

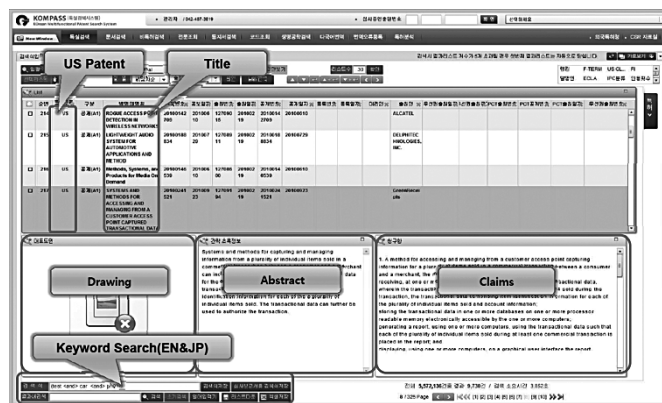


図2 韓国特許庁審査官向けシステム「KOMPASS」の画面例



3.2 K-PION

K-PION は、海外特許庁の審査官向けのシステムであり、現在 37 カ国／機関の特許庁に提供されている。

提供内容は、韓国の特許公報、書誌事項及び経過情報、並びに経過情報書類 (FileWrapper) である。韓国語公報の全文や書誌事項、経過情報項目は英語に機械翻訳される。



図3 K-PIONの包装袋入手画面例。経過情報項目は英語に翻訳されている。

3.3 KIPRIS

KIPRIS は、一般ユーザー向けの検索システムである。主なファイルとして、特許・実用新案／意匠／商標／審判情報／海外特許が提供されている。また、ユーザーインターフェースは韓国語版／英語版が提供されている。

英語版 KIPRIS では、韓国語／英語による検索が可能である。英語での検索対象は主要書誌はもちろん、公報全文の英文機械翻訳データに及んでいる。また情報表示に際しても、韓国特許の出願人名、主要書誌事項・経過情報、要約 (KPA) などが英文表示される。

韓国語版 KIPRIS のうち、「海外特許検索」では、英文キーワード検索が基本となっているが、日本特許に限り韓国語検索が可能である。これは庁内に蓄積した日韓機械翻訳データを一般向けに供用しているもの、と考えられる。

韓国語版 KIPRIS 「海外特許検索」の結果表示においても、公報全文の韓国語機械翻訳 (E2K) が無料提供されており、KIPO における機械翻訳データ活用の一端がうかがえる。

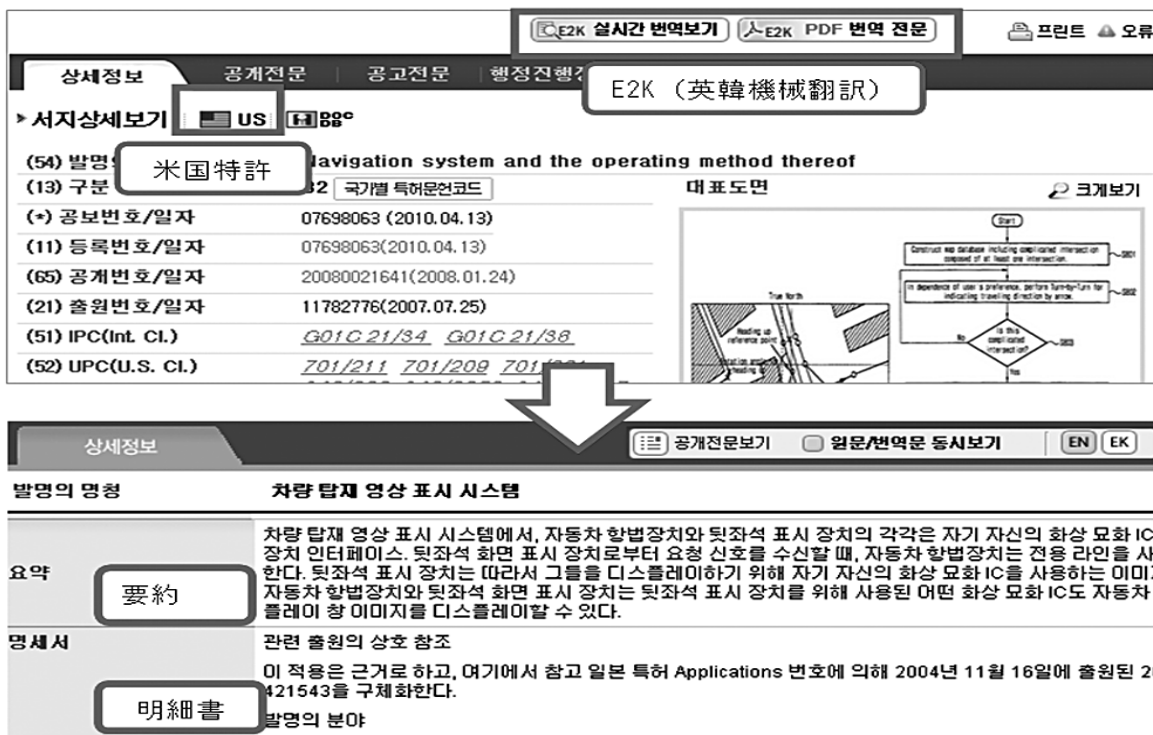


図4 KIPRIS 海外特許検索における、米国特許→韓国語機械翻訳の例

4 五大特許庁協力の概要と韓国特許庁の取り組み

五大特許庁では各庁における特許審査の質と審査効率の向上を目的として、各庁が協力してプロジェクトに取り組んでいる。現在は「10の基礎プロジェクト」が合意され、1庁が2つのプロジェクトを担当している。

10の基礎プロジェクトの内容と担当する庁は、次のとおりである。(資料1)

[作業部会1 特許分類関連]

- 1) 共通ハイブリッド分類 (EPO)

[作業部会2 情報技術関連]

- 2) 共通文献 (EPO)
- 3) 共通のサーチ及び審査支援ツール (USPTO)
- 4) サーチ戦略の共有化と文書に向けた共通アプローチ (USPTO)
- 5) 共通出願様式 (JPO)
- 6) 相互機械翻訳 (KIPO)
- 7) 共通サーチ・審査結果アクセス (JPO)

[作業部会3 審査関連関連]

- 8) 共通トレーニングポリシー (KIPO)
- 9) 審査実務と品質管理の共通ルール (SIPO)
- 10) 審査のための共通統計パラメータ (SIPO)

これより、本稿の後半では、KIPOが担当する「共通トレーニングポリシー」と「相互機械翻訳」の概要・進展について紹介する。

5 「共通トレーニングポリシープロジェクト」の概要

共通トレーニングポリシープロジェクトの目的は、各庁の研修コースや研修内容に関する情報を共有し、研修リソースの利用率を向上させることによって、審査官の研修の実効性を高めることにある。また、研修実践の相互利用を進めることによって、各庁のトレーニングコースのベンチマークを設けることや、将来的には各庁の審

査官の決定における相互矛盾を最小限に抑える狙いもあるとされる。

共通トレーニングの取り組みのひとつに「IP5 審査官ワークショップ」が挙げられる。ワークショップは5庁のいずれかで毎年開催される事となっているが、2010年度には全部で13個の研修が相互参加設定され、活発な取り組みがうかがわれる。2010年に中国の国家知的産権局で行われた研修の様子は(資料9)に、韓国特許庁における研修は(資料10)にそれぞれ紹介されている。

6 「相互機械翻訳プロジェクト」の概要と進展

「相互機械翻訳プロジェクト／Mutual Machine Translation (MMT)」は、機械翻訳を他のプロジェクトに利用し、言語障壁を克服することによりワークシェアリングを促進するプロジェクトである。

本稿の冒頭でも述べた通り、五大特許庁で取り扱う特許出願件数は、世界の特許出願の約80%を占め、その約30%にあたる約42万件は、パテントファミリーの関係にあるとされる。言語障壁を克服することで、各特許庁の間でサーチ・審査の共有が期待されている。一方で、各特許庁はこれまでも機械翻訳サービスの品質向上に努めてきたものの、現段階では人手翻訳での品質に及ばず、継続的な改善が必要とされる。

このような理由から、このプロジェクトでは五大特許庁それぞれで蓄積された知見やノウハウを活用し、翻訳品質の改善を通じて先行技術文献調査及び審査情報の流通・交流を促進する事を目的とするものである。

相互機械翻訳プロジェクトは、KIPOが担当庁となっている。2012年までの推進スケジュールは下表のとおりである。2012年秋には、主に日・韓・中→英の各機械翻訳システムがアップグレードされ、次いで米・欧でアップグレード結果の評価が予定されている。



表2 相互機械翻訳プロジェクト 推進スケジュール

Date	Item	Detail
~ 2008	Each MT system	Each offices developed their own MT systems
2009	Foundation Projects	IP5 agreed that MMT is one of the foundation projects for WG2
2010	MT Quality Assessment	IP5 conducted the quality assessment
2010	Error Review Pilot	IP5 conducted the pilot for error review and delivered feedback
2011	Full-Scale Error Review	IP5 conducted full scale error reviews 4 times and delivered feedback
2011	Automatic Assessment Pilot	JPO and KIPO conducted the pilots for automatic MT quality assessment
2011	Corpora Pilot	KIPO conducted the pilot for Corpora DB creation
Sep. 2012	MT Upgrade Report	JPO, KIPO, and SIPO will upgrade their MT systems and report the upgrade status
Oct. 2012	MT Quality Assessment	After MT upgrades, EPO and USPTO perform the MT quality assessment

元々 KIPO は、早い段階から日韓機械翻訳を進めていた。相互機械翻訳プロジェクトにおいても、実質的な取り組みは KIPO・JPO 間の「自動機械翻訳の品質評価に関する共同作業」からスタートしている（2011）。また、KIPO 内ではコーパス翻訳（統計翻訳）のパイロット版作成に着手した、とのことである（2011）。

2012 年秋以降には、日・韓・中の機械翻訳システムのアップグレード、及び、アップグレード結果の評価が予定されている。現在、五大特許庁がそれぞれ保有する機械翻訳システムは下記の通りである。

表3 五大特許庁の保有する機械翻訳システム

特許庁	サポート言語	提供サービス/システム	提供対象	翻訳方式
EPO	仏・独・西・伊・葡→英語	esp@cenet	一般向け	辞書(人手) + 統計ベース
		EPOQUE	EPO審査官向け	辞書(人手) + 統計ベース
JPO	日本語 → 英語	AIPN	海外審査官向け	辞書(人手)
		IPDL	一般向け	辞書(人手)
KIPO	英語・日本語 → 韓国語	KOMPASS	KIPO審査官向け	辞書(人手) + 統計ベース
	韓国語→英語	K-PION	海外審査官向け	辞書(人手) + 統計ベース
	英語・日本語 → 韓国語 韓国語→英語 (K2E)	KIPRIS	一般向け	辞書(人手) + 統計ベース
SIPO	中国語→英語	CPMT	一般向け	辞書(人手)
USPTO	仏・独・西・伊・葡・瑞	ASR	USPTO審査官向け	辞書(人手)

2012 秋のアップグレードでは、アジア言語→英語の機械翻訳をより人手翻訳結果に近づける事が目標と位置づけられている。人手翻訳に近づけるための手法としては、1) 各国言語とその特許文献に特化した、未登録語辞書及びパターン辞書の作成。2) 辞書ベース翻訳とコーパス翻訳を併用した、ハイブリッド機械翻訳方式の模索、の併用が進められている。

7 まとめ

韓国特許庁 (KIPO) では、かねてより英韓／韓英、日韓／韓日の機械翻訳を積極的に活用しており、その一端は英語版 KIPRIS における「英文全文検索」や、韓国語版 KIPRIS の「海外特許」の韓国語検索、結果表示の英韓機械翻訳などから窺い知る事ができる。

KIPO が担当庁となる、五大特許庁・相互機械翻訳プロジェクトでは、2012 年秋以降の活動として「アジア言語から英文への翻訳モジュールアップデート」と「欧・米特許庁による翻訳品質の評価」を進める予定、とのことである。アジア地域、特に中国特許・実用新案の出願件数が増加する中、今後の進展が期待される。

参考資料

資料1. v 五大特許庁 (JPO)

http://www.jpo.go.jp/torikumi/kokusai/kokusai2/godai_patent.htm

資料2. EPO Asian patent information

<http://www.epo.org/searching/asian.html>

資料3. KIPO's IP Information Services

<http://www.japio.or.jp/fair/files/2011/2011kipo01.pdf>

資料4. KIPO's MT Activity & IP5 Mutual MT Project

<http://www.mt-archive.info/MTS-2009-Kim.pdf>

資料5. シリウスソフトパンフレット

<http://siriussoft1.cafe24.com/global/brochure/SS Brochure.pdf>

資料6. Etri 社 業務報告

http://www.etri.re.kr/eng/pub/pub_01010000.etri

資料6. 特許の機械翻訳～翻訳コストの低減を目指して (JapioYearBook 2009)

http://www.japio.or.jp/00yearbook/files/2009book/09_5_01.pdf

資料7. 機械翻訳結果の自動評価と人手評価の比較

(JapioYearBook 2011)

http://www.japio.or.jp/00yearbook/files/2011book/11_4_05.pdf

資料 8. 五大特許庁 Web サイト (英文)

<http://www.fiveipoffices.org/index.html>

資料 9. SIPO で行われた五庁審査官相互派遣研修の紹介 (特許懇誌 262 号)

<http://www.tokugikon.jp/gikonshi/262/262tokusyu2.pdf>

資料 10. 韓国特許庁の研修に参加して (特許懇誌 262 号)

<http://www.tokugikon.jp/gikonshi/262/262tokusyu3.pdf>

資料 11. IP5 MMT Project (英文)

<http://www.fiveipoffices.org/documentation/industry/IP5MT.pdf>

資料 12. Current Machine Translation Status for Patent Domain Users in Korea

[http://aamtjapio.com/kenkyu/files/kenkyu03/AAMT_Japio_sympo\(20101210\)-04.pdf](http://aamtjapio.com/kenkyu/files/kenkyu03/AAMT_Japio_sympo(20101210)-04.pdf)

資料 13. 特許庁の特許情報提供における機械翻訳の活用と今後の展望 (日本語)

[http://aamtjapio.com/kenkyu/files/kenkyu03/AAMT_Japio_sympo\(20091127\)-01.pdf](http://aamtjapio.com/kenkyu/files/kenkyu03/AAMT_Japio_sympo(20091127)-01.pdf)