

# アジア・新興国特許調査における無料データベースの実力検証

○鈴木愛子<sup>1)</sup>, 涌井利果<sup>2)</sup>, 都築 泉<sup>3)</sup>  
東京応化工業<sup>1)</sup>, ワイゼル<sup>2)</sup>, HIT サービス研究所<sup>3)</sup>  
〒253-0114 神奈川県高座郡寒川町 1590  
Tel: 0467-75-6550 FAX: 0467-75-6420  
E-mail: a-hibino@tok.co.jp

## Ability of free databases in patent search in Asia and emerging countries.

SUZUKI Aiko<sup>1)</sup>, WAKUI Rika<sup>2)</sup>, TSUZUKI Izumi<sup>2)</sup>  
Tokyo Ohka Kogyo (TOK)<sup>1)</sup>, Wisel<sup>2)</sup>, HIT Service Institute<sup>3)</sup>  
1590 Tabata Samukawa-machi koza-gun Kanagawa 253-0114 Japan  
Phone: +81-467-75-6550 Fax: +81-467-75-6420  
E-mail: a-hibino@tok.co.jp

### 【発表概要】

ASEAN など新興国の調査実務では、商用データベースの収録不足を補完する目的で各国特許庁データベースを活用する企業も増えてきた。中国特許調査においても英語情報のみの調査では漏れも発生することから、中国語検索が採用されているように、ASEAN の特許調査においてもインドネシア、タイ、ベトナムなどの調査では各国特許庁データベースでの原語検索、検索結果の査読が行われるようになってきた。そのような中、PATENTSCOPE や Google Patents などの無料データベースに ASEAN や新興国の特許情報が収録され始め、注目を集めている。そこで PATENTSCOPE をはじめとする無料特許データベースの収録状況および検索・表示機能について、商用データベースでの検索事例も交えて、ASEAN や新興国特許情報を調査する担当者が、どのデータベースを選択して調査に当たるべきかの指針を示すことを目的に研究した成果を紹介する。

### 【キーワード】

PATENTSCOPE, 無料特許データベース, アジア, 新興国, 特許調査, 原語検索

## 1. はじめに

外国特許調査においては、欧米および中国、韓国など5大特許庁が関係する特許調査では、従来から商用英語データベースが検索項目、コマンド検索、履歴演算など検索機能だけでなく、検索結果の表示やダウンロード、解析機能などを活用して業務を効率的に進めるためのツールとして多用されてきた。

他方、ASEANなど新興国の特許調査では、それら各国情報を収録する商用データベースも限られている。

そこで、ASEANなどの調査実務では、各国特許庁のデータベースを活用して収録が不十分な商用データベースを補完しようとする企業も増えてきた。中国特許調査においては英語情報のみの調査では漏れも発生するため、中国語検索による調査が補完的に採用されている。ASEANなど新興国の特許調査においてもインドネシア、タイ、ベトナムなどの調査では各国特許庁の原語での調査も行われるようになってきた。

そのような中、PATENTSCOPE や Google Patents などの無料データベースに ASEAN など新興国の特許情報が収録され始め、注目を集めている。その収録は既に商用データベースを凌駕する状況にもなっている。

PATENTSCOPE は、従来、PCT 特許のみを収録していたこともあり、未だに「PCT 特許検索ツール」として捉えている方も多い。2017年8月に ASEAN 各国知財庁が有する特許・実用新案データの基本的な情報のほぼすべてが PATENTSCOPE に収録され、さらに中国など東アジアの情報も逐次収録されている。

2018年9月時点で不十分な収録であった中国特許情報も毎月のように更新され、商用データベースに匹敵する収録

状況となった。

また、これまで US 特許を中心に収録してきた Google Patents も収録国数を増やし、ASEAN など新興国の情報も収録された。ワールドワイドなデータベースとして知られている Espacenet も含め、これら無料データベースを利用して ASEAN など新興国の特許調査がどこまで可能か、実務の観点から商用データベースと比較検証することとした。

## 2. データベース収録と収録検証

### 2.1 データの収録比較

ここでは新興国の特許情報として ASEAN を、アジアの特許情報として中国、韓国を取り上げ、無料、商用データベースの収録状況を概観したのち、具体的な検索用語、特許分類として IPC および出願人検索を比較検証した結果を紹介する。

ASEAN では、10か国のうちミャンマーを除く9か国が特許情報を発行し、データベースに収録されて調査可能になっているが、ブルネイ、カンボジア、ラオスは発行数も少なく、検証母集団も小さいことからインドネシア、マレーシア、フィリピン、シンガポール、タイ、ベトナムの6か国を対象とした。

本報告で各国特許庁データベースと比較検証したデータベースを以下に示した。

無料 A: PATENTSCOPE(WIPO)

無料 B: Espacenet(EPO)

無料 C: Google Patents

無料 D: FOPISER(JPO)

商用 A: Japio GPG/FX

商用 B: Shareresearch

商用 C: Derwent Innovation

商用 D: Orbit Intelligence

商用 E:PatBase  
商用 F:WIPS Global

ASEAN 各国および中国、韓国の収録推移の詳細は最終報告の場で紹介するが、ASEAN 特許調査で原語検索が求められるインドネシア、タイ、ベトナムの概要のみ、この場で紹介しておきたい。

基本的に Espacenet の元となっている DOCDB 系の無料、商用データベースは使い物にならないほど収録が悪く、以下の議論では割愛した。

#### 1) インドネシア

無料では PATENTSCOPE、商用では商用 B および 2010 年公開以降は商用 C の収録がインドネシア庁の収録とほぼ同等。

#### 2) タイ

無料では PATENTSCOPE、2001 年公開以降は FOPISER が、商用では商用 B、商用 D、商用 E がタイ庁の収録とほぼ同等。

#### 3) ベトナム

無料では PATENTSCOPE、2008 年公開以降は FOPISER が、商用では商用 B、2010 年公開以降は商用 C、商用 D がベトナム庁の収録とほぼ同等。

### 2. 2 IPC、用語、出願人等、収録検証

上記、収録状況を化学、機械、電機など各種分野の代表的な IPC 用語、出願人から検索検証し、収録状況に対応する検索結果が得られるか確認した。

ここではインドネシアとベトナムについて IPC は B32B27、用語は multilayer および出願人は UNICHARM を例にして各庁と PATENTSCOPE の検索結果を比較した例のみを示す。用語と出願人についてはスペースやハイフンで表記されるものを選び、それぞれフレーズ検索の差異も確認した。同様に各庁と同様の

収録を示す商用データベースについても検証したが予稿においては割愛した。

表1 インドネシア特許の検索比較

	ID 庁	PS
B32B27	536	528
multilapis*	57	108
"multi lapis*"	49	184
UNICHARM	556	553
"UNI CHARM"	572	572

2000-2018 年公開特許、用語は要約中 multilapis (multilayer)  
PS: PATENTSCOPE

ID 庁 (PDKI) ではフィールド内演算ができないが、PATENTSCOPE では、  
(multilapis\* or "multi lapis\*")  
⇒ 287 件  
("UNICHARM" or "UNI CHARM")  
⇒ 1125 件  
などとフィールド内演算ができる。

表2 ベトナム特許の検索比較

	VN 庁	PS
B32B27	353	257
"đa lớp"	55	86
"nhị lớp"	189	346
UNICHARM	247	140
"UNI CHARM"	273	274

2000-2018 年公開特許、用語は要約中 "đa lớp" (multilayer)  
"nhị lớp" (multi layer)

VN 庁 (IPLib) では要約中から検索ができない、という問題もある。

PATENTSCOPE の IPC や出願人検索件数が各国庁より少ない理由は、若干の収録漏れがある、という点にある。

### 3. PATENTSCOPE の実力検証

データの収録検証でも紹介したように PATENTSCOPE 収録が ASEAN6 か

国だけでなく中国、韓国の収録でも実務に耐えうるツールであることを確認した。

本予稿では、PATENTSCOPE の検索機能、表示機能と共に問題点を取り上げた。

### 3.1 PATENTSCOPE の検索機能

Fields Combination と Advanced Search が存在するが、各種コマンドを駆使して検索できる Advanced Search が実務的には有用である。

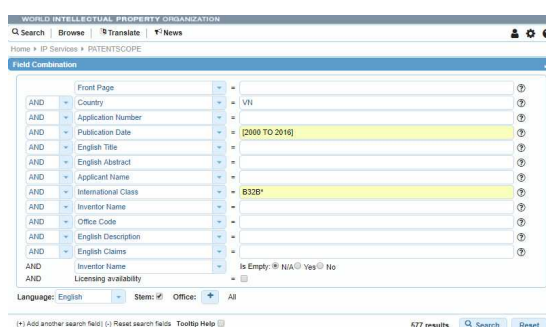


図1 Fields Combination 画面



図2 Advanced Search 画面

Fields Combination では「(用語 or IPC) and 出願人 and 出願年」などのような複雑な検索ができないが、Advanced Search では各種検索コマンドを使って、さらに近接演算子なども組み合わせ、複雑な検索式が可能となっており、この点では商用データベースに匹敵する機能とも言える。

収録データの点からは問題のない多くの特許庁サイトではこのようなコマンド検索ができないばかりかフィールド内での演算ができないものも多い。

また、Advanced Search では各国原語での検索が可能である。アジアや

ASEAN など新興国情報を収録する(商用)英語データベースでは機械翻訳英語であることが多く、誤訳などによる検索漏れを防ぐためにも原語検索をして情報を補完する点からは有用である。

インドネシアにおける多層フィルムの検索式例(原語と英語のハイブリッド検索)  
ID\_ALLTXT:(FILM NEAR5  
(LAMINA\* or "MULTI\* LAPIS\*" or  
"LAPIS\* MULTI\*" or  
MULTILAPIS\* or "MULTI\*  
LAPIS\*")) or ((EN\_ALLTXT:(film  
NEAR5 lamina\*) or IC:(B32B27\* or  
"B32B 27\*")) and CTR:ID) and  
DP:[2010 TO 2018]

NEAR などの近接演算子が使える他、ワイルドカード(\*)では前方一致、中間一致、後方一致、前後一致が機能する。

尚、原語検索するには国コードではなく、言語コードを使う。ASEAN 各国においては国コード=言語コードであるが、中国、韓国の言語コードは、それぞれ ZH, KO である。中国語で3D プリンタなどを検索するには「CN\_TI:(“三维打印”)」ではなく、「ZH\_TI:(“三维打印”)」と言語コードを使う。

さらに中国特許の場合には、英語で 3D printer や 3D printingなどを抽出する目的で「3D \*print\*」のようにワイルドカードを使うと語尾変化などの表記ゆれも抽出できるが、中国語で「三维打印机(3D printer)」を抽出する目的で「ZH\_TI:(“三维打印\*”)」とやると 0 件となり、用語も部分一致検索できないなど検索の癖がある。

### 3.2 PATENTSCOPE の表示機能

検索結果はデフォルトでは 10 件であるが、200 件まで表示可能であり、検索結果も 10000 件ダウンロードできる。

No.	Title	IPC Class	Applicant	Inventor	Pub Date
1. 051 4637	FILM MULTI LAPIS YANG DAPAT DIBILAS	B32B 27/00	Wipo Translate	SCHUHMAN, Michael	05.10.2011
2. 70 187045	FILM MULTI LAPISAN FILM TERLAMINASI UNTUK DIGUNAKAN SEBAGAI MATERI PENGEMAS, TAS PENGEMAS, DAN KANTUNG BERCIK	B32B 27/02	Google Translate	Hiroaki MATSUBARA	09.03.2018
3. 051 5300	FILM MULTI LAPISAN DAN KANTONG YANG DIBENTUK DARI FILM	B32B 27/02	Bing/Microsoft Translate	MORI, Tadashi	22.12.2011

図3 検索結果の表示

検索結果はインドネシア語、タイ語、ベトナム語、中国語などの原語であっても備え付けの Google 翻訳ツールなどにより英語はもちろん 108 か国の言語に瞬時に翻訳できる。

さらに Analysis 機能により、国別検索数、IPC、発明者、出願人、公報発行年別に上位10のランキングが簡易解析結果を表示できる。

Countries	IPC	Inventor	Applicant	Pub Date
China	40099	STATE GRID CORPORATION OF CHINA	37347	2009 448007
	4618K	THE INVENTOR HAS INVAIED THE RIGHT TO BE MENTIONED	17319	2009 481035
	A61P	华为技术有限公司	17319	2010 631064
	G01N	WANG WEI	12313	19894
	H04L	ZHANG WEI	11552	2012 1002391
	H01L	LI WEI	8568	2013 1320365
	B01D	WANG LEI	8525	11044
	F21V	LIU WEI	7969	2015 1451970
		ZHANG LI PEI	7911	2015 1505518

図4 Analysis による簡易解析

この点については既に紹介もされているように、出願人名の名寄せがされていないので名寄せをしたうえで再検索しないと正確な情報が得られない。

表3 中国公開特許ランキング

	Analysis	英語検索	中国語検索
①	37,347	52,763	60,754
②	18,542	27,341	40,884
③	16,614	31,690	35,151
④	15,043	20,069	26,439
⑤	11,120	20,013	26,819

(2019/3/20 現在)

図4の Analysis により示した中国の 2005~2017年公開特許のランキング上位5位までをそれぞれ再検索してみると上記表3のようになる。

- ① STATE GRID (国家电网)
- ② Huawei Technologies (华为技术)
- ③ ZTE Corporation (中兴通讯)
- ④ Samsung Electronics (三星电子)
- ⑤ Zhejiang University (浙江大学)

前記したように中国語検索では出願人検索においても出願人名をダブルクォーテーション(“”)で囲って検索しないと and 演算することになる。

表3における浙江大学を PA=(浙江大学)とすると「浙江 and 大学」となって「浙江工业大学」なども抽出される。

検索結果一覧にリンク表示されている出願番号をクリックすると「Biblio Data」タブに出願番号、IPC、出願人、発明の名称などの書誌的事項と共に抄録が表示される。

ASEANなど新興国の詳細情報は、この「Biblio Data」までの情報であるが、中国や韓国の詳細情報では、「Claims」「Description」「Drawings」タブが用意され、それぞれ「全クレーム」「全文」が原語ではあるが表示され、「全図」も確認することができる。

National Biblio. Data | Description | Claims | Drawings | Documents

Note: Text based on automatic Optical Character Recognition processes. Please use the PDF version for legal matters.

权利要求书

1. 打印头，用于三维打印机，包括安装座、第一雾化器及第二雾化器，所述第一雾化器的下端设有第一打印喷嘴，所述第二雾化器的下端设有第二打印喷嘴；  
其特征在于：  
所述第一雾化器上固定连接有第一滑块，所述第二雾化器上固定连接有第二滑块，所述第一滑块及所述第二滑块相对所述安装座均可在垂向往复移动；  
切换单元，所述切换单元包括滑板、第一弹簧及第二弹簧，所述滑板相对所述安装座可在水平方向往复移动；

National Biblio. Data | Description | Claims | Drawings | Documents

図5 クレームや図面情報

### 3.3 PATENTSCOPE の有用性

各国特許庁データベースに極めて近い収録を示す PATENTSCOPE であるが、これも既に指摘されているように IPC

や要約、出願人の欠落など一部不具合も存在することは認識する必要がある。

しかし、各国特許庁データベースではフィールド内演算などができないため、以下のような検索(インドネシア特許検索例)では16回もの検索をして、各データをマージし、重複除去するなど、かなり手間を要するが、これが PATENTSCOPE では1回の検索で済むことは極めて大きな利点である。

CTR:ID and DP:([2000 to 2018]) and (EN\_TI:(laminare or multilayer\* or multi-layer\* or "multi\* layer\*") or EN\_AB:(laminare or multilayer\* or multi-layer\* or "multi\* layer\*") or ID\_TI:(laminasi\* or multilapis\* or multi-lapis\* or "multi lapis\*") or ID\_AB:(laminasi\* or multilapis\* or multi-lapis\* or "multi lapis\*"))

また、VN 庁 (IPLib) でフレーズ検索すると、表4に示すように用語のフレーズ検索では要約中からの検索ができない。

PATENTSCOPE では発明の名称、要約の一部欠落のため若干少ない数値を示すものの要約中からも検索できおり、IPLib では抽出できない情報を抽出できるなど利点も大きい。

表4 ベトナム特許検索

	IPLib		PS	
	TI	TI+ AB	TI	TI or AB
"da l@p" (multilayer)	55	55	46	86
"nhi@u l@p" (multi layer)	189	189	170	346

#### 4. おわりに

東アジアおよび ASEAN 特許調査における無料および商用データベースについてデータの収録、検索や表示機能などを検証し、実務で利用するツールとして現状では何を選択すべきかの指針となる情報を提示することを目的に検討した。

その結果、DOCDB 系データベースでは、無料、商用データベース共に東アジア、ASEAN の収録が極めて悪く、実務では使用できないことが明らかになった。

無料の PATENTSCOPE (今なお、「PCT 特許検索ツール」として捉えられている面もあるが) や一部の商用データベースが東アジアや ASEAN 特許調査に利用できることもわかった。

各国特許庁データベースでは複雑な検索や表示機能はないものの全クレームや全文などが PDF で用意され、詳細情報として法律状態情報なども収録されているので、PATENTSCOPE で検索した集合を基に、登録や生死情報を各国特許庁データベースで確認する、という方法で権利侵害調査も可能と考える。

予稿においては、インドネシア、タイ、ベトナムなど主として原語検索・表示のデータベース収録状況や中国を含めた検索ノウハウ、表示機能などを紹介したが、マレーシア、フィリピン、シンガポールおよび韓国についても同様に検証して詳細を紹介する。

#### 5. 参考文献

- [1] 伊藤: PATENTSCOPE による ASEAN 特許調査概要 (Japio Year book2018)  
[http://www.japio.or.jp/00yearbook/files/2018book/18\\_2\\_04.pdf](http://www.japio.or.jp/00yearbook/files/2018book/18_2_04.pdf)