

平成 23 年度  
特許庁  
受託事業

# 中国特許情報の実態把握調査

JETRO

2013 年 12 月

日本貿易振興機構（JETRO）

北京事務所 知識産権部

## 2. 中国特許 DB

### 2.1 DBの種類とその関係

#### 2.1.1 国家知識産権局関連 DB と収録状況

国家知識産権局 (State Intellectual Property Office、以下 SIPO という) の下部組織である国家知識産権局知識産権出版社 (Intellectual Property Publishing House、以下 IPPH という) は、専利局関連組織の中で唯一、特許情報の提供・販売を許可された機関であるとされており、SIPO から提供される書誌情報データ、人手英訳抄録データ並びに公報イメージデータと、自己が作成するフルテキストデータを整理・統合・メンテナンスしている。

IPPH は SIPO に整理・統合後のデータを提供している他、国内外の営利・非営利機関に対しても様々な種類のデータを販売している。IPPH は 2011 年に民営化され、ウェブ上で独自に有償 DB の品質改善を実施している。国内外営利・非営利機関に販売されているデータは、CNIPR が運営する DB のそれと必ずしも同一ではなく、同じ種類のデータ項目であっても契約内容により異なるようである。

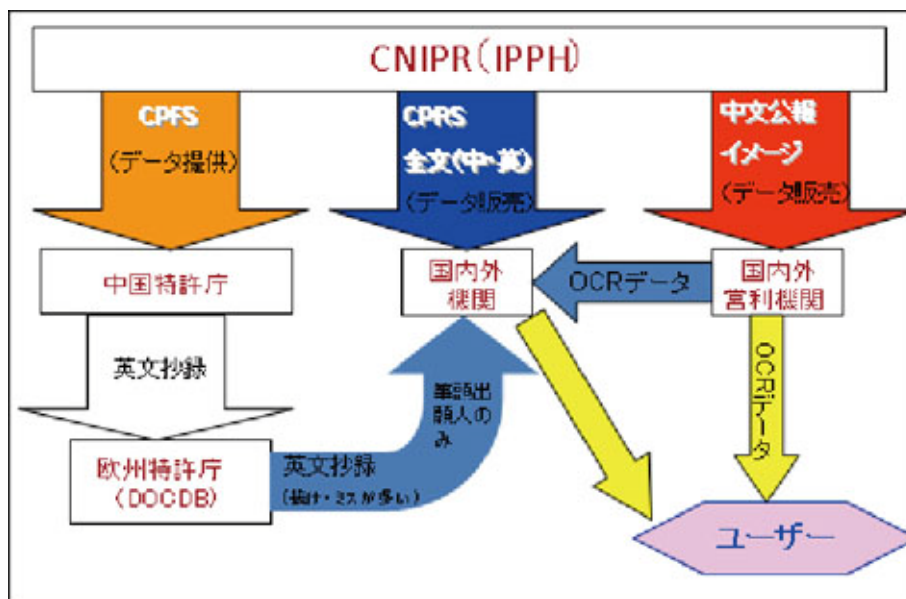


図1. 中国特許データの流れ

CNIPR は中国語版と英語版が提供されている (各 DB の詳細は、第3節参照)。IPPH は、2004 年に開始された電子出願以前のデータ並びに電子出願開始以降も一部の出願人から提出される紙出願のデータについては、独自に OCR で読み込み、人手によるチェック・修正をかけ、中文フルテキストデータとして提供している。なお、2011 年 9 月より、中国語版の新プラットフォームがリリースされた (旧プラットフォームは 2012 年 3 月に廃止予定)。英語版は、前述の中文テキストを独自に確立した手法で機械英訳し

たデータを収録している。これらのデータは、作成元である IPPH 自身が定期的にエラー修正やアップデート等の手を加えているため、本 DB の収録内容が全ての DB の基礎となっており、当然のことながら書誌情報の収録率は最も高いと考えられる。

一方、専利局が独自に運営する SIPO は、IPPH から提供された情報を無料ツールとして公開しているが、技術内容の検索用ツールとしてよりは、特定の特許の書誌情報や法的状態をチェックする機能に特化しており、検索可能なテキスト情報の範囲は中文・英文共に抄録までの情報である。

SIPO、CNIPR 共に、中文 DB については直近まで書誌・テキスト情報共に収録されているようである。これに対し、SIPO の英文 DB は書誌情報を含めた収録のタイムラグが 3 ヶ月以上であり、2010 年以前のデータはかなりの収録斑がある。2011 年にインターフェースが刷新した CNIPR の英文 DB については、書誌・テキスト情報共に収録漏れは見られないものの、中文 DB と比べて一週間のタイムラグが確認された。なお、テキスト情報の収録率については、2009 年の情報プロフェッショナルシンポジウムにて報告された汎用語を用いた方法<sup>1)</sup>にて確認した。

公開年	SIPO 中文	SIPO 英文	新旧 CNIPR 中文			新 CNIPR 英文	
	書誌 ・抄録	書誌 ・抄録	書誌	クレーム	全文	書誌	全文
2005 年	100%	40%	100%	100%	100%	100%	100%
2006 年	100%	29%	100%	100%	100%	100%	100%
2007 年	100%	64%	100%	100%	100%	100%	100%
2008 年	100%	100%	100%	100%	100%	100%	100%
2009 年	100%	93%	100%	100%	100%	100%	100%
2010 年	100%	90%	100%	100%	100%	100%	100%
2011/1-6	100%	100%	100%	100%	100%	100%	100%
2011/7	100%	100%	100%	100%	100%	100%	100%
2011/8	100%	19%	100%	100%	100%	100%	100%
2011/9	100%	100%	100%	100%	100%	100%	100%
2011/10	100%	100%	100%	100%	100%	100%	100%
2011/11	100%	0%	100%	100%	100%	100%	100%
2011/12	100%	0%	100%	100%	100%	100%	100%
2012/1/4	100%	0%	100%	100%	100%	100%	100%
2012/1/11	100%	0%	100%	100%	100%	100%	100%
2012/1/18	100%	0%	100%	100%	100%	100%	100%
2012/1/25	100%	0%	100%	100%	100%	0%	0%

表 1. SIPO、IPPH 運営 DB の収録率（専利公開：2012/1/30 現在）

## 2.1.2 各種商用 DB と収録状況

### 2.1.2.1 各 DB のソースタイプ別特徴

図1に示した通り、商用 DB が収録する中国特許データのソースは、収録項目によっても様々であり、殊に、英文データのソースに関しては、同一 DB においても年代によって異なるものさえある。これらは主として以下の7タイプに分類できる。それぞれに特徴を有するが、複数のカテゴリーにあてはまる DB (ex. 厳密には PatBase は④と⑤に該当) や、DB の収録期間によって異なるタイプに該当するもの (ex. Shareresearch Brics は、2011 年以降のフルテキストコンテンツについては③であるが、2010 年以前分については②に該当) があり、①以外の多くの DB では、書誌情報については DOCDB 情報を採用しているなど、その関係性は複雑に入り組んでいる。

なお、以下のタイプ別の分類は、DB 提供各社の資料、説明又はヒアリングに基づくもので、実体を検証した結果に基づくものではない点並びに、本報告に記載した DB は代表的なものであり、全ての DB を網羅している訳ではない点、ご留意願いたい。

①CNIPR 系中文データ採用 DB	SIPO 中文、CNIPR 中文、Orbit.com、PAT-LIST-CN/WEB、専利 SEARCH-i2 等
②CNIPR 系英文データ採用 DB	CNIPR 英文、Orbit.com、JP-NET 等
③独自機械英訳系 DB	TotalPatent、Shareresearch Brics、Discover 等
④独自 OCR 系 DB	PatBase、PAT-LIST-CN/WEB、Chinese Patents Fulltext 等
⑤DOCDB 抄録採用 DB	ESPACENET、PATOLIS 海外、Ultra-patent、Shareresearch 海外等
⑥人手付加価値付与 DB	DWPI、CA (Chemical Abstracts) コンテンツ等
⑦その他	Patent SQUARE (英文)、NRI サイバーパテント 2 等

表2. 中国特許収録 DB (ソースタイプ別)

#### ① CNIPR 系中文データ採用 DB

IPPH が作成している中文テキスト情報を直接あるいは中間業者等を介して採用した DB であり、更に以下の3タイプに分けることができる。日本人が精度の高い検索を目指す上で、特に有用と思われるのは iii) の中英ハイブリッド型 DB であり、書誌情報についても、DOCDB 由来ではなく、信頼性の高い中文系の情報を採用している点でも有用である。但し、機械英訳データの出所や品質は DB によって異なるため、注意が必要である。

i) 中国内 DB・・・中国内の機関・個人によって運営される中国人向け DB

特許庁傘下の DB (CNIPR 中文、SIPO 中文等)

中国地方行政区運営の DB (上海 IPO, 5ipatent 等)

中国個人運営の DB (SooPAT 等)

ii) 機械和訳検索用 DB・・・日本の企業によって運営される日本人向け DB

HYPAT-i (発明通信社)

PAT-LIST-CN/WEB (レイテック)

専利 SEARCH i2 (アイピー・ファイン)

iii) 中英ハイブリッド型 DB・・・中文・英文の両方で検索が可能な商用 DB

Orbit.com (Questel-Orbit)

Discover (CPA Global)

Patent SQUARE (パナソニックソリューションテクノロジー)

Patentics (Easepat)

### ②CNIPR 系英文データ採用 DB

IPPH が中文テキストを独自のシステムにて翻訳した英文データを採用した DB である。CNIPR 英文は 2011 年に全く新しい DB に生まれ変わり、完全に有料化された。複合検索機能、保存機能、解析機能等、様々な機能が付加され、収録漏れや遅れも見られないことから、CNIPR 中文と同等に多機能な DB となっている。Orbit.com は、①にも分類されている、中英ハイブリッド型の DB である。JP-NET は 2011 年より IPPH の英文フルテキストを順次収録中であったが、2012 年 1 月末現在では、遡及分の収録が確認できる。

新 CNIPR 英文 (IPPH)

Orbit.com (Questel-Orbit)

JP-NET (日本パテントデータサービス)

### ③独自機械英訳系 DB

IPPH が提供する中文テキストを、独自に又は中間業者を介して機械翻訳した英文データを採用した DB である。IPPH の翻訳データを採用せずに独自の翻訳データを提供しているという意味においては、SIPO の英文データもここに該当すると言える。Discover 並びに Patentics は、①にも分類されている、中英ハイブリッド型の DB である。

Shareresearch Brics (日立製作所)

TotalPatent (Lexis-Nexis)

Discover (CPA Global)

Patentics (Easepat)

### ④独自 OCR 系 DB

中国の PDF 公報を独自に又は中間業者を介して OCR でテキスト化した中文、あるいは当該中文データを機械翻訳した英文データを採用した DB である。① (CNIPR 中文系データ採用) の DB は、IPPH が機械読み込み後に入念にチェックしたデータを

採用しているため、中文テキストの精度は統一化されているはずであるが、本タイプ（独自 OCR 系）の中文テキストの質は各々の OCR 精度に依存するという点で①と異なる。

PatBase は PDF 公報を独自に OCR で処理した中文フルテキストを採用（英文は公開抄録のみ）している。Dialog に収録されている Chinese Patents Fulltext は、オランダの Scipat Benelux 社が独自の翻訳システムにて作成した機械英訳データを提供している。PAT-LIST-CN/WEB は、従来からの CNIPR 中文系機械和訳データに追加して、独自に OCR で読み込んだ中文フルテキストデータを、検索対象として（表示は不可）順次遡及中である。

PatBase (RWS)

Chinese Patents Fulltext (Dialog)

PAT-LIST-CN/WEB (レイテック)

#### ⑤DOCDB 抄録採用 DB

EPO から入手した DOCDB データの人手和訳抄録（英文）のみを採用した DB である。DOCDB 由来のデータは抜けやミスがあるので注意が必要である。

ESPACENET (EPO)

Ultra-patent (Wisdomain)

PATOLIS 海外 (パトリス)

Shareresearch 海外 (日立製作所)

#### ⑥人手付加価値付与 DB

クレーム部などを人手で翻訳したり、発明の特徴となるキーワードを人手で切り出したり等、人手による付加価値が付与された DB である。

DWPI (WPI オンライン) の中国コンテンツと、Thomson Innovation のアジアコンテンツに収録されている英文クレーム情報は同じものであり、2007 年以降の公開特許と実用登録並びに 2011 年以降の登録特許（第 1 クレームのみ）は全件人手翻訳によるものである。なお、DWPI の中国コンテンツとしては、2006 年以前も独自にクレームなどから情報を付加した WPI 抄録をファミリー単位のレコードとして提供しており、マニュアルコードなどの特殊分類と併せて使用することにより効果的な検索が可能となっている。

CA (Chemical Abstracts) は、特許を含む化学および化学工学分野全般の技術関連文献情報を収録した DB であり、SciFinder や STN の CAPLUS ファイル等で検索が可能である。中国特許に関しては、CAS が化学情報と看做す収録保証 IPC については漏れなく収録されており、クレーム情報などから、発明者が強調している点や、化合物情報等に関する索引が人手で付与されるため、化学分野の特許を効率的に抽出すること

ができる。なお、テキスト情報の収録は抄録までである。

DWPI (Thomson Reuter)

Thomson Innovation アジアコンテンツ (Thomson Reuter)

CA (CAS)

#### ⑦その他

中国特許の機械英訳フルテキストコンテンツを提供しているが、データソースが明かされていない DB である。Patent SQUARE は、①にも分類されている、中英ハイブリッド型の DB である。

Patent SQUARE (パナソニックソリューションテクノロジー)

NRI サイバーパテントデスク 2 (NRI サイバーパテント)

### 2.1.2.2 収録状況の比較

上記タイプ②～④の代表的な DB を特定のキーワードで検索した件数について、CNIPR のデータと比較すると、表 3 のようになった。英文 DB の収録率については、DB ごとに辞書が異なり、中には同義語を自動抽出する設定となっている DB もあるため正確な比較はできないが、CNIPR よりも 2 割以上少ない期間については、未収録のレコードや、未翻訳のデータ等の存在が危惧される。

DB\発行年月		2012.1	2011.12	2011.11	2011.10	2011.1-9	2010	2009	2008	2007	2006
中 文	①-1	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.10	1.09
	①-2	0.53	0.98	0.98	0.97	1.05	0.98	0.96	0.98	1.08	1.07
英 文	②-1	0.48	0.91	0.89	0.89	0.82	0.85	0.83	0.87	1.09	0.94
	②-2	0.06	0.07	0.75	0.79	0.74	0.75	0.75	0.78	1.01	0.87
	③-1	0.54	0.77	0.90	0.92	1.00	0.88	0.81	0.84	1.07	0.92
	③-2	0.05	0.56	0.92	0.93	0.90	0.79	0.82	0.84	1.10	0.96
	④-1	0.00	0.67	1.10	1.12	1.06	1.09	1.00	0.98	1.13	1.17
	⑦-1	0.00	0.00	0.00	0.05	0.88	0.99	1.02	1.03	1.31	1.12

表 3. 全文収録割合 (専利公開+実用新案: 2012.1.27 現在)

(詳細な説明=半導体, semiconductor\*で検索して得られた件数につき、CNIPR 中文、新 CNIPR 英文の件数を 1 として割合を算出)

このように、同じタイプの DB であっても、収録状況は様々である。但し、上述のように、特に英文データについては入手方法が複雑化しており、ベンダーによってはデータ入手ルートが予告なく変更されるケースも見られることから、現在収録率やタイムラグの点で優劣があっても、数ヵ月後には立場が逆転している可能性も

あるため、特に SDI に用いる DB としては、タイムラグの状況を随時確認する必要がある。

### 2.1.2.3 英訳コンテンツの比較

英文系 DB に関しては、クレームの訳文が検索・スクリーニングの精度を左右するため重要となってくる。近年、検索の観点から機械英訳に頼る危険性が指摘されており<sup>3)</sup>、文法的に精読に耐える DB が存在しないという問題点もあるが、人手英訳である抄録文と併せて、決め手となる概念の有無をチェックする上では有効となる場合がある。以下に特定特許の訳文を、ポイントで比較した。

公開番号CN102089275A (対応日本特許タイトル: エチレン性不飽和化合物からのニトリル化合物の製造方法) の英訳テキストを比較すると、同じソースタイプに属するはずのDBであっても必ずしも同一ではない。原文の少一个烯属不飽和度的烃基化合物 (=少なくとも1個のエチレン性不飽和を有する炭化水素系化合物) に相当する各英文を比較すると、“at least . . . non-degree of saturated compound” のような、意味を取り違えかねない表現がいくつか見られる一方、“at least one ethylenic unsaturation “のように正しく訳されているケースもある。

有機磷配体 (=有機リン配位子) については、“ganophosphorus ligand” machine phosphorus ligand” 等の意味不明な訳文等が見受けられ、単齿有机亚磷酸酯化合物 (=単座有機ホスフィット化合物) については、正しく訳されているものもある一方で、phosphiteと訳すべき部分がphosphateになっているものや、未変換の漢字が残っているものがある。

原文：

包含至少一个烯属不飽和度的烃基化合物通过在包含选自过渡金属的金属元素和至少一个与金属元素形成络合物的有机磷配体的催化体系存在下在液体介质中与氰化氢反应而氢氰化的方法, 其特征在于该催化体系包含至少一种单齿有机亚磷酸酯化合物和至少一种单齿有机磷化合物, . . . .

②-1

1. Contains at least alkenes non-degree of saturation alkyl compound through to contain to elect from crosses the metal the metallic element and at least forms the complex with the metallic element the ganophosphorus ligand under the catalyzed system exists in liquid medium and hydrogen cyanide response, but hydrogen cyaniding method, its characteristic lies in this catalyzed system containing at least one kind of single tooth organic Asia phosphate chemical compound and at least one kind of list tooth organic phosphine chemical



compound, . . . .

②-2

1. Alkyl compound comprising at least one ethylenic degree of unsaturation, through comprising metallic element and at least one ones that form complex with metallic element selected from transition metal as follows But hydrocyanation method of making and reacting with hydrogen cyanide in the liquid medium of catalysis system of the machine phosphorus ligand, characterized by that this catalysis system comprises at least a kind of unidentate organic phosphite ester compound and at least a kind of unidentate organic phosphine compound, . . . .

③-1, ③-2

1. comprises at least one ethylenic unsaturation comprising the hydrocarblyl compound selected from the group consisting of a metal of the metal elements and the at least one metal element to form a complex with the organic ligands in the liquid medium in the presence of a catalytic system in which the method of reacting with hydrogen cyanide hydrocyanation, characterized in that the catalytic system comprises at least one 单齿有机 phosphitylated compound and at least one tooth with a organophosphinic compounds, . . . .

③-3

1. Contains at least one ethylenically unsaturated compound by the degree of alkyl containing metals selected from transition metals and metal element and at least one form of organic phosphorus ligand complex catalyst system in the presence of hydrogen cyanide in liquid medium and the reaction and hydrogen cyanide method, characterized in that the catalyst system comprising at least one organic phosphite compound single-tooth and at least a single tooth organic phosphine compounds, . . . .

④-1, ⑦-1, ⑦-2

1. To 1. at least non-degree component of the saturated compound is a 11-13, ?wherein is selected from a is 50-150 transition metal element and at least one of the complex the ligand ganophosphorus catalytic system is present in the metal element in the liquid medium and acrylonitrile reaction, ?and

hydrogen cyaniding's, ?characterised is a catalytic system comprises at least one of a tooth of organic phosphate compound and to at least one of preparing organic phosphate compound, . . .

⑥-1

1. A method for the hydrocyanation of a hydrocarbon-based compound comprising at least one ethylenic unsaturation, by reaction, in a liquid medium, with hydrogen cyanide in the presence of a catalyst comprising a metal element chosen from transition metals and an organophosphorus ligand, wherein the organophosphorus ligand is composed of a mixture of at least one monodentate organophosphite compound and at least one monodentate organophosphine compound; . . .

このように、DB ごとに使用している文法辞書、技術用語辞書の精度に差があるようであり、比較ポイントとした3点全ての訳を満足できたのは⑥-1（独自付加価値系）のみであった。また、本件の技術主題が、“hydrocyanation method”に関するものであることが文法的に読み取れるのも⑥-1のみであった。中国特許の英文抄録はほぼ全てが人手翻訳であるため、分野によっては一次・二次スクリーニングまでを機械英訳 DB のみで賄える場合もあるが、検索の観点からは、できるだけ機械英訳 DB の検索だけで終わらせず、中文 DB で補うべきであり、抽出の観点からは、できるだけ抄録の英文、対応特許の内容並びに原文と対比させつつ、もっともらしい文章でも鵜呑みにせずに慎重にスクリーニングすべきである。

### 2.1.3 各種DBの活用場面

#### 2.1.3.1 CNIPRを用いるケース

上述の通り、現在 IPPH によって運営されている新 CNIPR 中文は、コンテンツの観点からは最も信頼性の高い DB であると言える。また、複合演算や、フルテキストデータを対象とした検索が可能であり、同義語検索機能、保存機能、概念検索機能、監視機能、解析機能等を兼ね備えるなど、特許庁関連の DB の中ではトップレベルであると言っても過言ではない。無料部での利用はあくまでもβ版としての位置付けであり、2011年に IPPH が民営化されて以降は、ログイン時に個別の ID、パスワードが必要な有料部分の機能と大幅に差別化される予定であったが、民営化後の最大の改善点として期待されていた表示・ダウンロード機能に関しては、未だに従来の10件ずつのままである。また、テキスト情報として簡便に確認できるのは第一クレームまでであり、商用 DB のように全クレームを連続閲覧できる機能や、栞機能、コメント挿入機能等、自己の判断を DB 上に残す機能は備えていない（2012年1月末現

在。

従って、中国語を理解できるユーザーが、自身で検索・スクリーニング・判断までを専任で完結できる場合であっても、注目特許の抽出には公報の印刷やメモを余儀なくされる。中国語を理解できないユーザーが、英文 DB や和訳系 DB での抽出漏れを補う目的で利用する場合には、表示・ダウンロード機能が現在のレベルである以上、多大な困難が伴う。

CNIPR 英文については、従来の C-PAT はあくまでもベータ版としての位置付けとのことで、収録率、機能共に制限されていたため、あまり利用価値がなかった。これに対し、完全有料化された新 CNIPR 英文については、収録率・機能共にかなりの改善が見られるものの、収録内容は所詮機械英訳である。言語の相性の観点から、機械英訳は機械和訳と比較すると精度は高いが、2.1.2.3 に記載のような問題点があるため、侵害予防調査を機械英訳系 DB のみで完結させるのは望ましくない。また、本 DB はリリースされたばかりで検証が不十分であるため、下記の目的別の使用例としては、CNIPR 中文のみを対象とした。

#### ●中国語を理解するユーザーが CNIPR を用いるケース

“中国語を理解するユーザー”が、特許調査の知識を習得しており、検索、スクリーニング、判断の流れを一ユーザーで完結する場合には、現在の CNIPR 中文は、技術動向調査、無効化資料調査、侵害予防調査、出願人調査等、あらゆる技術内容からの調査に耐えるコンテンツ・機能を備えている。但し、未だ表示・ダウンロード機能が貧弱であるため、オフラインの状態で閲覧を行う場合には非常に不便であり、葉・コメント挿入機能を有していない等の難点はある（詳細は第3節参照）。

#### ●中国語を理解しないユーザーが CNIPR を用いるケース

現在の CNIPR 中文は解析機能が充実しているため、中国語を理解できないユーザーであっても、技術動向調査であれば、非常に有効に活用できる場合がある。

以下に、コンフリクト予防調査の事例を示す。代理事務所や代理人名を検索項目に含むのは上記①に属する DB のみであるが、中でも簡易に多量のデータ処理が可能な解析機能を有する CNIPR 中文（但し、無料部にて解析可能なレコード数は1万件以下）は、この種の調査に適している。図2、図3は特許事務所“隆天”が扱う日本企業クライアントの分析を目的とした事例であるが、1分程度（但しネット環境による）で完了することができる。

侵害予防調査においては、抄録系 DB や、機械英訳系 DB しか利用できない場合に、補足目的で有効に活用できる。また、2.3 で後述する通り、書誌情報に DOCDB データを採用している英文系 DB においては、漢字表記の出願人名の検索に大きな不備があることから、それを補う上でも価値がある。しかしながら、中国語を理解しないユーザーの場合は本 DB 上での閲覧ができないため、得られた結果の番号を、人手翻訳系又は機械英訳系の DB に移行して閲覧・判断するしかない。この際、本 DB の表示・

ダウンロード機能は現状 10 件ごとであるため、抽出件数が数千件にのぼる場合には番号移行は非常に手間がかかる作業となる。



図 2. CNIPR 中文によるコンフリクト予防調査検索事例

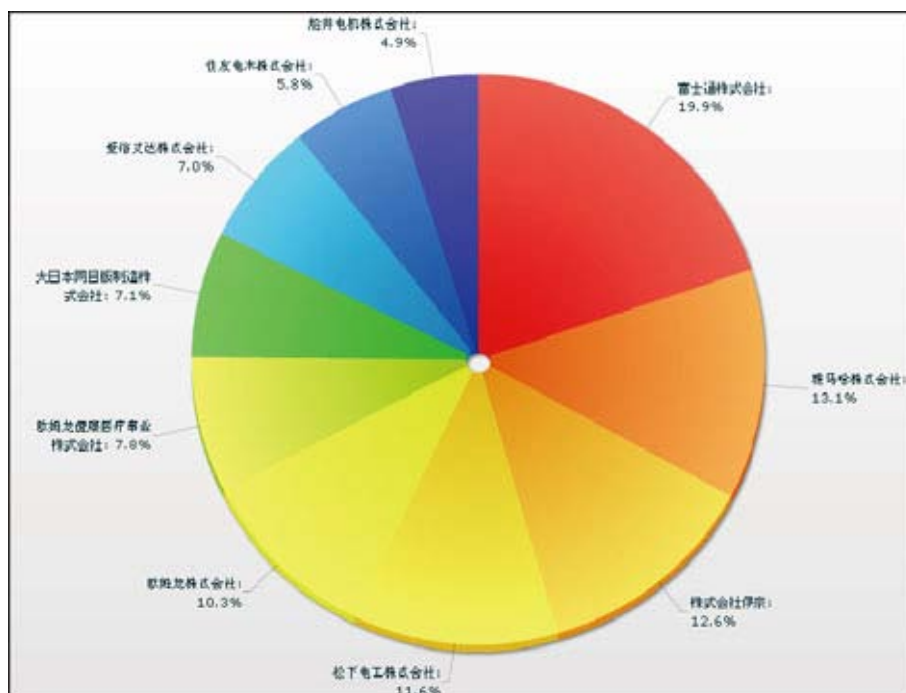


図 3. CNIPR 中文によるコンフリクト予防調査解析結果

### 2.1.3.2 その他 DB

商用 DB はいずれも、中国語を理解しないユーザーにとって価値があるものである。各々独自の強みがあることから、ユーザーの調査環境（全社で展開している DB、予算、中国特許を重視する度合い等）、調査目的、調査対象（技術分野、出願人等）に応じて、様々な利用例が考えられるが、特に中国の特許調査において最も重要かつニーズが高いと思われる、侵害予防調査について、中国語を理解しないユーザーが用いた場合について考察する。

1 つの DB で調査、スクリーニングまで行う場合は、DWPI（化学分野の場合は CAS）または中英ハイブリッド型の DB が最も望ましい。

#### ●DWPI の活用

DWPI はマニュアルコード等の独自の技術分類が活用でき、特に高分子分野等の化学分野の分類が充実している（但し、特殊分類の使用には別途会員料金が必要となる）ため、フルテキスト型の DB が登場する以前から十分に精度の高い検索が可能であった。

中国特許に関しては、2007 年以降はクレーム全文の人手英訳を収録しており、遡及分については発明者抄録よりも情報量の多い独自の抄録を有している。2006 年以前のクレーム全文を検索対象とできない点がよく問題視されるが、2006 年以前は中国特許の年間発行件数が現在の 25%以下であり、かつ中国 1 ヶ国のみの出願は全体の半数以下であった<sup>5)</sup>。出願から 7 年以上権利が維持されている特許の割合は内国人出願（＝中国 1 ヶ国出願である可能性が高い）については全体の 30%以下であり<sup>5) 6)</sup>、複数国に対応特許を有する残り半数のレコードは、他国に付与された抄録・分類からも抽出可能である。即ち、DWPI ファイルにおけるクレーム検索対象外の中国特許レコードの中で気をつけなければならない特許の絶対数は僅かであり、2006 年以前の発行分については出来るだけ分類のみで抽出する等の手当てをすれば足りる。

実用新案については、2006 年以前は収録されておらず、ほとんどが対応特許を有さない内国人の出願によるものであるため、実用新案も調査する必要がある場合には別 DB で補う必要があるが、権利期間が 10 年間と短いため、DWPI のみの検索でカバーできるようになるのは時間の問題である。

#### ●CA の活用

化学、製薬、バイオに特化した調査の場合は、CAS のケミカルアブストラクトコンテンツ（STN の CAPLUS ファイルや SCIFINDER）が推奨できる。収録保証 IPC に関しては、中国特許の収録漏れはなく、タイムラグは書誌が数日～1 週間、索引情報が最短 2～3 週間程度である。実用新案については 2006 年以降分しか収録されていないが、合成方法や化合物の新規性調査などの場合には問題ないと思われる。特に化合物名に関しては、機械英訳の質の悪さが問題視されている<sup>2)</sup>ため、機械英訳

系の DB より CA や DWPI ファイルが有効である。但し、化学分野の調査であっても、機械・エレクトロニクス分野の用途も含めた調査の場合は収録されていない対象特許がある可能性があり、製造プロセスの調査であっても、装置に関する実用新案が権利化されている場合もあるため注意が必要である<sup>2)</sup>。

#### ●中英ハイブリッド型 DB の活用

特殊な検索コードを必要としない機械・金属分野などのケースや、DWPI や CA を利用できる環境にない場合は、中英ハイブリッド型の DB が推奨される。現在、Orbit.com、Discover、Patent SQUARE、PatBase などがこのタイプの DB としてリリースされている。機械英訳の不備と、適切な中文を選択する能力不足を互いに補完することによる効果が期待できる。例えば、漢字表記の出願人名を特定した検索を実施し、英訳と中文を両方表示させて閲覧することも可能である。但し、各コンテンツの収録状況や中英検索によって補えるカバー率は DB によって異なり<sup>4)</sup>、前述の通り英文コンテンツの収録状況は今後不安定化する可能性があるため、自社で導入済みの DB の質については定期的なチェックが必要である。

## 2.2 対応特許情報調査

#### ●特定の中国特許の対応特許を調査する場合

SIPO や CNIPR で表示されるファミリー情報は、INPADOC ファミリーの情報である。INPADOC ファミリーは主要国をカバーしているものの、台湾や韓国等のマイナー国情報については、年代によって収録されていない場合もあるため、できる限り包括的な対応特許調査を実施するには DWPI のファミリー情報も確認した方が良い。また、DWPI ファイルは優先番号が共通でない NON CONVENTION 特許も対応特許として収録しており、独自に書誌情報を更正しているため、序のデータや INPADOC データより信頼できる場合がある。

#### ●特定の国内外特許の対応中国特許を調査する場合

中国特許の書誌情報の収録に最も信頼性が高いのは SIPO または CNIPR であるため、これらの優先権項目で対象特許の出願番号を検索するのが望ましい。CNIPR のファミリー情報は、SIPO から提供された INPADOC データである。しかしながら、ウェブ DB 上での更新は、SIPO より CNIPR の方が早いケースが確認されているため、直近の対応中国特許の有無まで確認したい場合には、SIPO と CNIPR の両方を確認した方が良いと思われる。

調査事例：特願平 6-283415 の対応中国特許調査

<SIPO 中文>

現在の SIPO 中文の優先権項目では、項目内での演算ができず、マスクの%を用いると検索に異常に時間がかかりエラーになることが多い。また、西暦の年号は 4 桁表示の場合と 2 桁表示の場合があるため、推奨の入力形式は

「283415」であり、他国の番号や他の年代の出願番号に関するノイズは目視で除くしかない。

<CNIPR 中文>

前述の理由から、優先権項目に「JP and 283415」のような形式で入力するのが推奨される（他の年代の出願番号に関するノイズは目視で除く）。

### 2.3 生死分け情報の信頼性、活用方法

中国特許 DB 上のステータス情報は、全て SIPO 由来の情報であり、その信頼性は低いとされている（従って、ステータス情報に関しては、速報性の優劣はあっても DB による信頼性の優劣はない）。特に、DB 上で失効となっており、権利回復の情報が掲載されていない場合でも、原簿を確認すると権利が存続しているケースが実際にあるため、DB 上の生死分けの機能に頼るのは危険である。とはいうものの、中国特許の半数以上又は実用新案の 9 割以上を占めている中国在住の個人・法人による出願は権利維持率が低く<sup>4) 5)</sup>、膨大な抽出件数に対して全件を確認するのは非常に非効率である。従って、ヒット率の高い集合や、注目企業・大学等の出願については、出願年から 20 年前以降分については全件確認し、念のために実施した網羅性重視の集合については DB の生死分け機能を利用して効率的にチェックする等の工夫が必要である。

### 2.4 出願人表記ゆれ対応

中国特許の出願人名を調査する際には、中文 DB を使用する場合と、英文 DB を使用する場合とで、考慮すべき対象が異なる。

#### ● 中文 DB を使用する場合

中国の出願明細書において、出願人名は中国語で記載されるが、カタカナ表記又はローマ字表記の出願人名は、代理人によって様々な表記のバリエーションがある。例えば、中国語にて、出願人名としてブリヂストンを網羅するには、“Bridge Stone” の意味を漢字に当てはめた「桥石」「石桥」の他に、ピンインを漢字に当てはめた「普利司通」「布里奇斯托恩」「布里吉斯通」「普利司通运动」「普利斯通股份」なども検索する必要がある<sup>6)</sup>ため、中文 DB でこのような出願人名を検索するのはあまり推奨できない。但し、英文 DB の出願人情報は、基本的に共願人情報が含まれていないため、厳密に調査する場合には、中文、英文両方で検索するのが望ましい。英文系の DB として特に望ましいのは、出願人コードを有する DWPI ファイルであり、発行件数 500 件以上の法人に対して付与されているようである。

漢字の出願人名（特に中国企業名）については、中文 DB での検索で問題ないが、中国企業の関係会社や中国研究機関の関連企業まで網羅したい場合には、事前に同社の中文ホームページを訪問するなどして関連組織情報をチェックする必要がある。

る（特に、規模の大きい中国の大学は、自己実施可能とも判断できる関連企業を有しているため注意が必要である）。なお、新 CNIPR 中文には、出願人名補助ツールが付帯されているため、活用を検討するのも一考である。

●英文 DB を使用する場合

英文 DB に収録されている出願人情報には以下のように問題点が多い。

a) 共願人情報が収録されていない<sup>1)</sup>

b) 中国法人名のアルファベット表記を網羅するのは非常に困難である

c) データ抜け、入力ミスが多い

a)の課題については、2010年12月よりDWPIが改善を進めているが、遡及分の手当ては完了していないようである。b)の課題については日本法人についても当てはまる場合がある（例えば、三菱化学は”Mitsubishi kagaku” “Mitubishi kagaku” “Mitsubishi chemical”等で表記される）が、中国の出願人については、日本企業以上に可能性のあるバリエーションを予測しにくい。c)の課題については、出願人の国籍を問わずに考慮すべき問題である。例えば、昭和電工の場合、英語表記は”SHOWA DENKO”に統一されているものの、中国語の昭和電工で抽出される1274件を、DOCDB系の英文DBに移行したところ、空白25件、昭和電工以外53件、ミス入力18件のレコードが確認された（2012.1.30現在）。ミス入力の事例を表4に示す。



SHAWA DENKO K. K.
SHOA DENKO K. K.
SHOW DENKO K. K.
SHOWA DENDO K. K.
SHOWA DENDO KABUSHIKI KAISHA
SHOWA DENK K. K.
SHOWA ELECTRIC PLASTIC PRODUCT
SHOWA VENKO K. K.
SHOWA WENKO K. K.
SHOWAN DENKO K. K.
SHOWDA DENKO
SHOWO PENKO PLASTIC PRODUCT
SOWA DENKO K. K.

表 4. 英文出願人データの誤表記事例

但し、日・欧・米の企業については、中国単独で出願しているケースは少ないと思われるため、ファミリー単位の検索が実施できる DB (DWPI、Patbase、Orbit.com 等) が有効である。例えば前述の DWPI ファイルにて出願人コードを用いた検索をファミリー単位のレコードに対して実施することにより、かなりのカバー率が期待できる (実際、上記昭和電工の事例で抽出された 1274 件を DWPI に移行した場合、PACO=SHOW or PA=" SHOWA DENKO" で抽出されなかったのは古い年代の 3 件のみであった)。一方、中国法人の出願人名に関しては、例えローマ字の商標が英語圏で浸透している BYD のような大企業であっても、中文 DB にて中国語表記の検索を実施した方が無難である<sup>7)</sup>。

#### 注記

- 1) 伊藤徹男 他「中国・台湾および韓国特許庁データベースの全文検索機能とその活用」(情報プロフェッショナルシンポジウム, 2009)
- 2) 知財管理 2012, Vol. 62, No. 1, pp. 67-83 「中国特許調査に関する研究」
- 3) 赤壁幸江「化学企業のための中国特許情報調査」(化学工業日報社, 2011 年) 説明資料
- 4) 田畑文也 他「英語・原語によるハイブリッド検索」(情報プロフェッショナルシンポジウム, 2011)
- 5) 西内盛二「ますます高まる中国特許調査の重要性」(日本アルミニウム協会外部講師講演会, 2010) 説明資料 pp. 15-16
- 6) 赤壁幸江、小山裕史「ここがポイント! 中国特許調査」(日本特許データサービス, 2009)
- 7) 知財管理 2011, Vol. 61, No. 4 pp. 521-531 「中国特許調査環境の実情」