

## 中国特許調査に関する研究

——企業の調査活動のための中国特許調査ノウハウ——

知的財産情報検索委員会  
第 2 小委員会\*

**抄 録** ここ 1, 2 年で中国特許調査環境は急速に充実しつつある。10 年前までは中国から国外に提供される特許情報データには欠落が多く、英文抄録の内容は貧弱であったため、データベースによる中国特許調査は実質不可能に近い状態であった。しかし、中国における知財推進政策の促進が追い風となりこれらの問題は急速に解消され、現在は機械英訳文ではあるものの、特・実、公開・登録公報全文の英文データベースが利用可能となった。

データベースによる調査環境が整った今、企業活動に必要な調査はどの程度まで可能となったのか？多くの調査担当者が悩むデータベースによる調査の精度につき主要な検索項目ごとに検証し、翻訳に起因するいくつかの問題点をうきぼりにしつつ留意事項を整理した。

### 目 次

1. はじめに
2. 中国特許データベースの現状
3. 検索項目毎の検証と留意点
  3. 1 IPC 検索の留意点
  3. 2 出願人検索の留意点
  3. 3 キーワード検索の留意点
  3. 4 翻訳精度の検証
4. 考察とまとめ
5. おわりに

### 1. はじめに

2011 年 2 月の産業構造審議会における国際知財戦略<sup>1)</sup>において、特許庁は新興国における特許出願の増加により日本語・英語以外の特許文献、特に中国特許文献が急増している現状を述べ、企業・特許庁ともに戦略の転換が必要とし、知財紛争を避けるうえでの中国特許、実用新案の調査の重要性を強調した。

中国の特許・実用新案出願件数は、2000 年には合計で 12 万件に満たない数字であったが、こ

の 10 年で急激な増加の一途をたどり、2010 年には特許出願が 409,836 件、実用新案が 391,177 件と、合計すると実に 80 万件を越す件数となった(図 1 参照)。このうち特許出願件数だけに着目しても、日本の 344,598 件を大きく引き離し、2010 年は中国が特許出願件数、GDP ともに日本を抜いて、米国に次ぐ世界二位の知財大国、経済大国となる逆転の年となった。

この急激な出願件数の増加は中国の国家政策に牽引されたものである。2006 年に中国政府が発表した「国家中長期科学技術発展計画綱要(200-2020)」において掲げられた「自主創新政策」では、一定の知的財産権を有する中国企業に対する税制の優遇措置などを定めた「ハイテク企業認定」など、国内企業への財政的支援政策が数多く含まれ、これらが中国国内企業の知財権取得推進の大きな原動力となっていることは間違いないであろう。

\* 2010 年度 The Second Subcommittee, Intellectual Property Information Search Committee

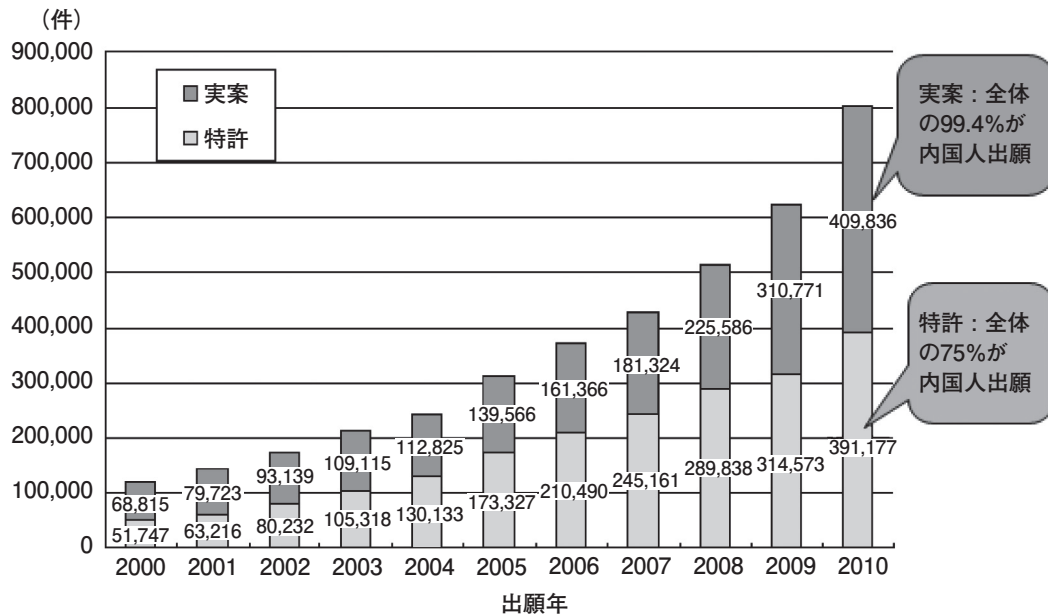


図1 中国特許・実用新案出願件数推移 (SIPO統計データより)

上記背景にも由来して、中国の特・実出願における大きな特徴は、内国人の出願の占める割合が大きい点である。2010年の特許出願についてはその75%が内国人出願、実用新案にいたってはその99.4%が内国人による出願であった(以上の数値はいずれもSIPO統計データ<sup>2)</sup>より)。このことは、第一国出願として最初に公開される情報が中国語であることを意味し、さらに対応外国特許がない場合には中国語公報が唯一の公開情報となることを意味する。

ビジネス市場としてますます中国の重要性が高まる現在、中国特許情報を考慮せずにグローバルな企業活動を行うことは不可能である。事実、中国に重要な製造拠点や市場を有する多くの企業が、現地代理人あるいは現地の調査会社を利用した侵害予防調査を実施している。しかしながら、中国特許情報が「量」「質」とともに急増している現在、新たな開発テーマ着手時の技術動向調査、特許出願の際の先行技術調査、権利化阻止のための公知文献調査など、より多くのステージにおいて中国特許情報が必要とされており、よりスピーディかつ的確に、目的に

応じた方針で実行可能なデータベース(以下DBと略)による調査の必要性が高まっている。

そこで、本稿では多くの調査担当者が悩むDBを用いた中国特許調査の精度、すなわち、今現在どの範囲まで中国特許のDBによる調査が可能なのか?また、どのような点に考慮してDBの選定や調査を行うべきかを検証した。

なお、本検証は2010年度の委員会活動で行われたものであり、各DBの機能や検索結果は当時のものであることを予めお断りしておく。

## 2. 中国特許データベースの現状

中国の特許調査環境はこの10年で大きく変化した。

以下は、特許調査環境の変遷をDBの種類・数と併せて年代で区切った年表である。

- ◆ゼロ期 中国特許情報を得るための有効なツールがない時代。(ニーズも低い)
  - ⇒ 中国特許情報が必要な場合は現地で調査を実施
- ◆第一期 中国特許情報が注目され始める

本文の複製、転載、改変、再配布を禁止します。

(2002年～)

⇒ レイテックPATLIST-CN/WEBリリース  
 (日本初の中国専用DB)

◆第二期 中国特許データベースの増加，全文  
 機械英訳DBの登場，独自機能の追加（2006  
 年～）

⇒ DBの使い方，機能の比較研究が盛ん  
 に（中国特許調査手法手探りの時代）

※中国 自主創新政策，シュナイダー訴訟<sup>3)</sup>

◆第三期 DBは出揃った感あり（2010年～）

⇒ ユーザー側の選択と目的に応じた使い  
 分けへ！

2002年頃より中国特許の英文DBは年々増加  
 し，2007年には機械英訳ではあるものの，全文  
 情報を搭載したDB：CNIPR<sup>4)</sup>が登場した。そ  
 の後，既存の商用DBも収録範囲の拡大，機能  
 増強を続けている。

DBの特徴を理解する際に役立つのは，各DB  
 のデータソース（情報源）とデータの流れであ

る。図2は中国特許情報の流れと代表的なDB  
 を図示したものである。近年リリースされた全  
 文系DBは一部を除き，その情報源を特許中国  
 知識産権局知識産権出版社（IPPH）のCNIPR  
 としており，最近は，この情報にさらに人手翻  
 訳，出願人の名寄せまたは中文併記，ステータ  
 ス情報の付与などの付加価値を付与したDBが  
 提供され始めている。

これらのDBの特徴，機能比較や使い方等を  
 解説した資料，書籍，論説は数多くあり<sup>5), 6)</sup>，  
 関連したセミナーや講演会も多数開催されてい  
 る。個々のDBの特徴や使い方を知りたい場合  
 にはこれらを参照されたい。

### 3. 検索項目毎の検証と留意点

中国特許調査のハードルを高くしている最大  
 の要因は“言語障壁”と“データ収録の信頼性”  
 である。しかし実際のところ，現時点では権利  
 存続期間である過去20年に遡り機械英訳された  
 特許・実用新案の全文DBが存在し，しかも収

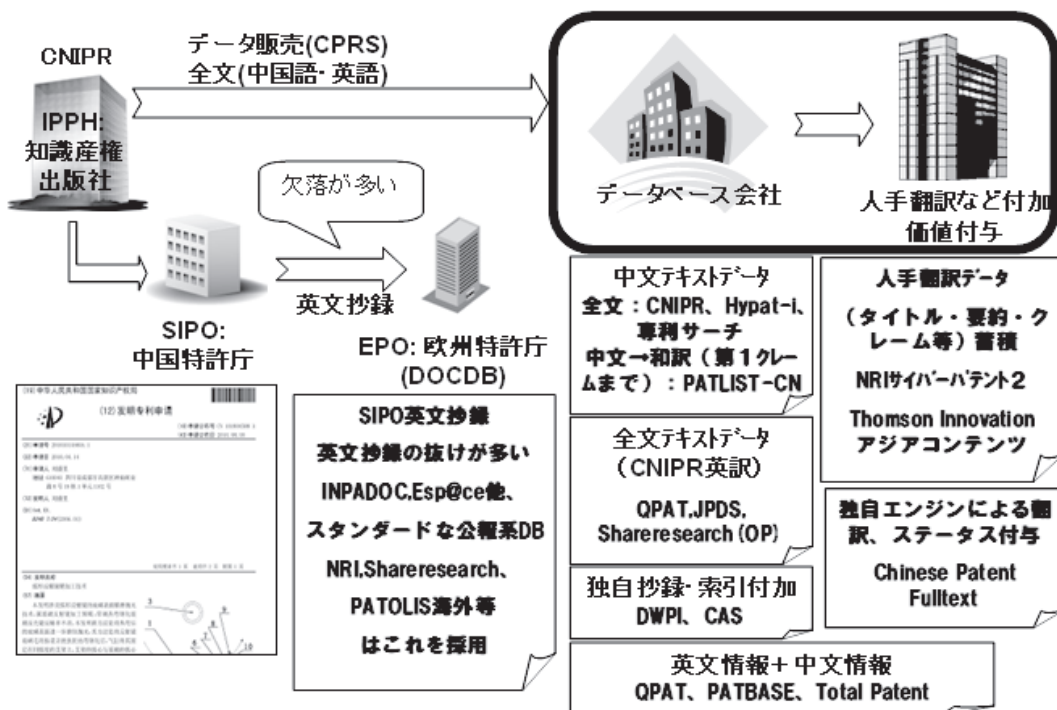


図2 中国特許データのフローと中国データベース

## 本文の複製、転載、改変、再配布を禁止します。

録率、タイムラグ共に欧米特許と比較して大きな差は無くなっている。にもかかわらず、未だに“中国特許調査は難しく困難なもの”という印象がぬぐえない。

企業の調査担当者あるいは知財担当者もつ不安（疑問）を集約すると以下の5点が挙げられる。

- ① 収録期間・収録率は十分なのか？
- ② 特許分類（IPC）は使えるのか？
- ③ 書誌情報は正確か？
- ④ キーワード検索の信頼性は？  
（翻訳に関わる不安1）
- ⑤ 得られた結果はスクリーニング、解析に耐えるのか？（翻訳に関わる不安2）

これらの疑問をクリアにし、現状の問題点を理解してこそ、最善の調査を行うことができる。

上記①について、使用するDBの収録範囲は特許調査におけるもっとも基本的な確認事項であるがゆえに、これまでにも十分検証が行われており<sup>4), 7)</sup>、しかも現在進行形で拡充している。そこで、ここでは情報の収録範囲に依存する問題点は追及せず、中国特許情報のデータそのものの特異性に依存する項目、つまり②～⑤の疑問につき取り組むこととした。

つまり、本研究においては、

「どのツールで調査をすべきか？」ではなく、「(与えられたツールで) どのように調査をすべきか？」という視点で(1) IPC, (2) 出願人, (3) キーワード(検索の立場), (4) 翻訳精度(スクリーニングの立場)の4項目について検証し、その結果得られた留意点と対処法をまとめた。

### 3. 1 IPC検索の留意点

#### (1) IPC付与精度の検証

中国特許のIPC精度については2006年にも当

委員会内で検証が行われ<sup>8)</sup>、この時の検証結果として、中国公報に付与されるIPCと日本のIPCとでは振れが存在するため、その振れを認識した上でのIPCの選定が必要、と結論づけられた。しかし、当時の状況と現在とで決定的に違うのは、IPCがリフォームドIPC(以下IPC-R)に変わった点である<sup>9)</sup>。

IPC-Rとは、2006年1月に発効したIPCのことで、その最大の特徴は、すべての特許文献をIPCが改訂されるたびに分類しなおし(再分類)、既発行公報についても最新のIPCで検索可能という点である。また、再分類はファミリーごとに行われる。すなわちファミリーのうちの一つの文献に対して再分類を行い、他のメンバーに同じ分類を付与する。このように再分類されたすべての分類データは、欧州特許庁が構築するIPC同盟国の特許文献の分類データを一元管理するDB MCD<sup>10)</sup>に格納される。多くの商用DBがこのMCDを特許分類の情報源としているため、現在では公報発行時に付与されたIPCだけでなく、公報発行後に再分類されたIPCおよびファミリーに付与されたIPCを併せて検索が可能となっている。

#### 【検証1】

クラス毎に8つの分野のIPCをサンプル抽出し、発行年別に中国特許レコードへの分類付与数と分類の深さ、および適合性を比較・検証した。この時に中国国内出願人の出願と外国人出願のものに分けて比較した。なお、使用したDBはIPC-Rが検索可能な商用DBである。

図3は電池分野のIPC(H01M4/00)が付与された中国公開公報において、付与されているIPCの総数を内国出願人と外国出願人で比較した例である。縦軸はIPCの付与数比率を示し、横軸は発行年を示している。

4年以上前、すなわち2007年以前発行の公報においては概して外国出願人の公報に付与されるIPCの数が内国人のそれに比べて多く、平均



本文の複製、転載、改変、再配布を禁止します。

して1以上の差異が認められる。この理由は、出願してから経過年数が増えるほどファミリー特許が増え、その結果IPC-Rとして収録されるIPCの個数は増加するためである。

従って、このファミリー数の影響を比較的受けにくい過去3年分のデータについて確認したところ、内国人と外国人のIPC付与数の差は1未満となり、2010年の平均付与数は内国人3.53個、外国人3.58個と殆ど差異は認められなかった。

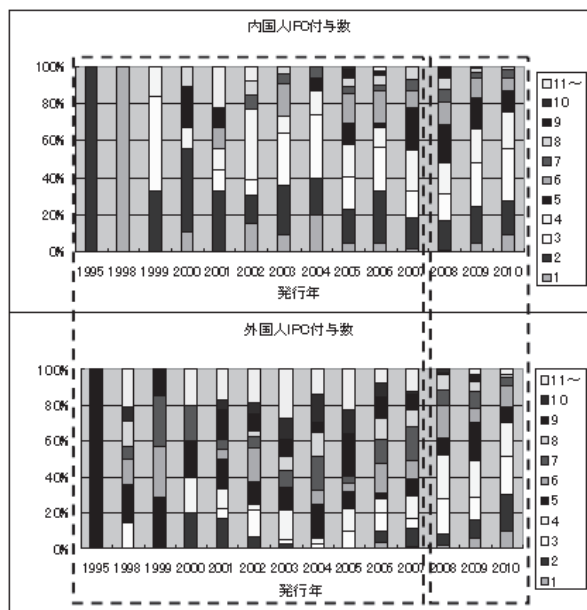


図3 中国特許に付与されたIPC個数比較 (H01M4/00の事例)

上記のようなIPC付与数の検証の他、付与されたIPCの深さ（細かさ）およびその適合性につき全8分野で検証を行った結果、とりわけ2001年以降分については内国人出願のIPCと外国人出願のIPCとは分類の数、深さ、傾向に有意な差異は認められなかった。

【検証2】

日本、中国、米国、EPおよび韓国の2007年～2009年の公開公報を対象とし、請求項にキーワード「電気自動車」を含む集合を抽出した。次いで各集合について付与されるIPCの傾向に差があるかを調べた。使用したDBはIPC-Rを

収録する商用DBである。

図4は縦軸が公報発行年、横軸は付与されたIPCをサブクラスで表したバブルチャートであるが、日本の分類付与の傾向に最も近い傾向を示したのが中国であり、ともにB60L「電気的推進車両の推進装置」、H01M「化学的エネルギーを電気的エネルギーに直接変換するための方法または手段、例. 電池」、およびB60K「車両の推進装置または動力伝達装置の配置または取付け」の3分類が上位を占め、他の分類の付与比率にも大きな差異は認められなかった。

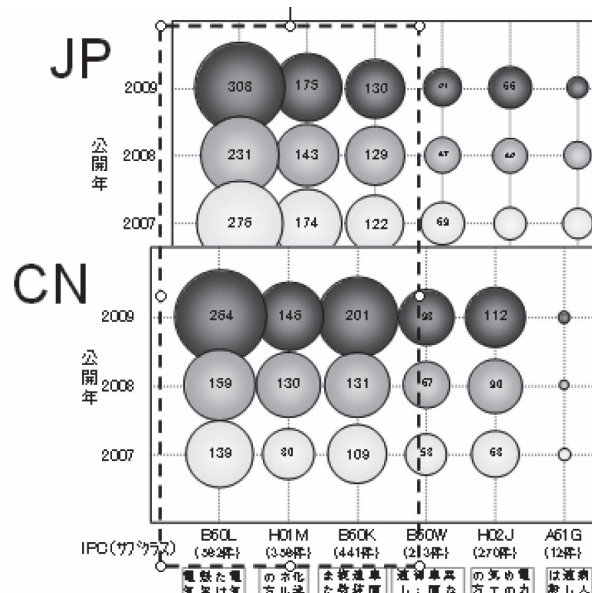


図4 キーワード検索「電気自動車」で得られる特許に付与されたIPC (中国と日本の比較)

(2) IPC検索の留意点

中国特許調査におけるIPCの精度は日本と大きな差異は無く、IPC検索を行う際には特別な配慮は必要ないと考えられる。

しかし、このことには再付与されたIPC-Rが好影響していることは間違いなく、IPC-Rデータを管理するMCDデータの利用の有無によって大きく影響される。すなわちMCDを利用するDOCDB<sup>11)</sup>データを使用しているDBとそうでないDBとでIPC検索の結果に差が出るので

## 本文の複製、転載、改変、再配布を禁止します。

注意したい。さらには、ファミリー単位のレコード構成となっているDBでは、対応特許のIPCも統合して検索対象となるため、異なる結果が得られる。

このように、検索DBによりIPCの収録ルールが異なり、結果に大きな差が出ることにのみ留意が必要である。以下、主要DBのIPC収録状況をタイプ別に分けて示す（注2011年3月現在）。

### 【タイプA】 公報付与IPCのみを検索

CNIPR, SIPO-DB, HYPAT-i, PATLIST-CN

### 【タイプB】 IPC-Rを検索

TIアジア, JP-Net, Total Patent, CPF, Shareresearch, NRI2, Espacenet

### 【タイプC】 IPC-Rに加えファミリーのIPCを併せて検索

DWPI, PATBASE, QPAT (ORBIT)

## 3. 2 出願人検索の留意点

### (1) 中国企業名称の表記ゆれ

中国出願人の英文表記には注意が必要である。必ずしもHPなどで自称している英文社名でなくブランド名で特許出願している場合もある。さらには、綴りミス、スペース有無、ピンイン表記<sup>12)</sup>の場合も考慮して幅広く検索する必要がある。

#### 【事例1】 フォックスコン

フォックスコンは電子機器の生産を請け負う電子機器受託生産（EMS）では世界最大の企業「鴻海精密工業（Hon Hai Precision Industry；ホン・ハイ・プレジジョン・インダストリー）」のブランド名。台湾に本社を構え、生産拠点は主に中国。（以上、ウィキペディアより、参照日2011年9月1日）

英文DBを用いてフォックスコンを対象とした出願人調査を行うと想定した場合、上記ウィキペディアから得た情報を用いて“Foxconn”（ブランド名）と“Hon Hai Precision”（企業名）とのOR検索を行うことができる。しかし、結論を言えばこれではまだ不十分であり、漏れのない出願人検索としてはフォックスコンの中文表記である富士康のピンイン表記であるFushikang, さらには同社の中国における傘下企業のHongfujinまで検索する必要がある。

表1はEspacenet<sup>13)</sup>を用いてこれら4種の出願人表記で検索を行った結果である。各々の件数を見れば、いずれか一つの出願人名が欠けたとしても不十分であることがわかる。

表1 フォックスコン出願人検索（Espacenet）

	出願人名	件数
1	Foxconn	7,062
2	Hon Hai Precision	24,040
3	Fushikang	1,247
4	Hongfujin	13,094
5	1+2+3+4	40,151

なお、漏れの無い調査を目指す場合は、上記の4種の表記に加え綴りミスやスペースの有無なども考慮し幅広く検索する必要があるが、トムソンロイターの提供するDWPI（Derwent World Patent Index）では、出願人データのクリーニングおよび標準化に加え、関連会社の関係までを考慮したコード化がなされており、今回のケースでは4社とも出願人コード“HONH”で抽出することができる。この出願人コードはDWPI上でしか使用できないが、コード表はトムソンロイターのウェブサイト上<sup>14)</sup>で参照することができる（図5参照）。

本文の複製、転載、改変、再配布を禁止します。

HONH	FOXCONN TECHNOLOGY SHENZHE
HONH	FUSHIKANG KUNSHAN COMPUTER
HONH-	HON CO LTD
HONH-	HON CO
HONH-	HON CORP
HONH	HON FU JIN PRECISION IND CO LTC
HONH	HON FUN JIN PRECISION IND SHEN
HONH	HON HAI JIN PRECISION IND SHENZ

図5 トムソンロイター社 DWPI出願人コード表 (HONHの抜粋)

このウェブサイトを利用して出願人表記の揺れを確認するのも有効な手段である。ただし、コード化されるのは一定の出願件数のある企業であり、全出願人が100%統制されているわけではないので注意が必要である。

また、中国企業名を英文検索する場合は綴りミスやスペースの有無による表記ゆれが多いことに注意されたい。そのため、もし可能であれば中文検索を併用することにより、漏れを防ぐことができる。

中文表記は、英文検索から得た公報番号を中文DBで逆引きすることによって確認可能である。また、ある程度の規模の企業であれば該当企業のウェブサイトあるいはウィキペディアから英文名、中文名称およびピンイン表記まで調べることができる(図6参照)。本事例の場合は、富士康(Foxconn=Fushikang)および鴻海(Hon

フォックスコン	
各種表記	
繁体字:	富士康
簡体字:	富士康
拼音:	Fùshìkāng
発音:	フーシーカン
英文:	Foxconn

図6 ウィキペディアより、フォックスコンの各種表記(参照日2011年9月1日)

Hai)、鴻富錦(Hongfujin)で検索する。

なお、企業のウェブサイトから得た出願人名のみで検索することには検索漏れを起こす危険が伴うので注意が必要である(詳細は後述)。

### 【事例2】清華大学

清華大学は北京市西北郊の清華園に所在する総合大学。国家重点大学の一つであり、近年は中国トップ大学の評価が定着しつつある。中国国内における2010年の特許出願件数は1,509件であり、大学では浙江大学に次ぐ2位。企業との共同出願の占める割合は大きく積極的に企業との共同研究を進めている。

清華大学を対象とした出願人調査を行う場合、大学のHP情報をもとに下記の中文、英文表記を確認できる(図7参照)。



図7 清華大学HPより同校ロゴ  
(<http://www.tsinghua.edu.cn>) 2011年9月1日

- ・清华大学(中文表記)
- ・Tsinghua Univ(英文表記)

各々についてCNIPRの中文・英文DBで検索を行った。

### 【結果】

CNIPR(中文)で“清华大学”を検索すると10,125件の結果が得られた。次いで、CNIPR(英文)で“Tsinghua University”を検索した場合、8,710件となった。

英文検索の場合に1割以上回答が減少する理由は、約15%が公式英文表記“Tsinghua”ではなく、ピンイン表記である“Qinghua”が用いられていたためである。Qinghua University

本文の複製、転載、改変、再配布を禁止します。

で検索の結果、1,415件の回答が得られた。

表2に主なDB毎の英表記件数を示す。EspacenetやPATBASEなどDOCDB系のデータを用いた多くのDBで“Tsinghua”が約8割，“Qinghua”が約2割の比率となっているが、DWPIだけは異なる傾向を示した。(DBによって件数に差異があるが、これは各々のDBの収録期間やレコード構成に違いがあるためである。今回は英文表記の違いを示すための検索であるため特に検索期間などは揃えていない。)

表2 英文DB上の清華大学の出願人英表記

	CNIPR	Espacenet	PATBASE	DWPI
Tsinghua Univ	8,710	8,468	8,604	315
Qinghua Univ	1,415	2,320	2,376	9,634

注) TsinghuaとUnivの語順指定せず検索、Univは前方一致検索した。Qinghua Univも同様

DWPIは、前述の通り出願人名の標準化がなされており、清華大学の特許データは全て“UNIV QINGHUA”に統制されて収録されている(図8)。Tsinghuaで抽出された315件は“RES INSTITUTE OF UNIV TSINGHUA IN SHENZHEN”や“UNIV TSINGHUA GRADUATE SCHOOL”等いずれも別の機関であった。

```

L2 ANSWER 1 OF 10690 WPIX COPYRIGHT 2011 THOMSON
Full
Text
STN
AN 2011-J95272 [201153] WPIX
ED 20110818
TI Time division multiplexing light differential phase :
code generating device, has intensity modulators con:
delay wires, where modulating speed rates of data sin:
DC W01; W02; X26
IN HUO L; LOU C; XING Y
PA (UYQI-C) UNIV QINGHUA
CYC 1
PI CN 102118336 A 20110706 (201153)* ZH 8[6]
ADT CN 102118336 A CN 2011-10050604 20110303
PRAI CN 2011-10050604 20110303
IPC1 H04L0027-18 [I,A]; H04Q0011-00 [I,A]
AB CN 102118336 A UPAB: 20110818
NOVELTY - The device has a continuous wave laser, m
sources, differential coder, multiple intensity modu
light delay wires. Each intensity modulator includes
connected with an output end of a continuous wave la
    
```

図8 清華大学DWPIのレコード例

なぜこのように英文表記にばらつきがあるのか?それは中国特許庁から提供される英文データソースの出願人表記に原因があると推測される。図9はThomson Innovationの解析機能を用いて、年代ごとの出願人表記推移をグラフ化したものである。

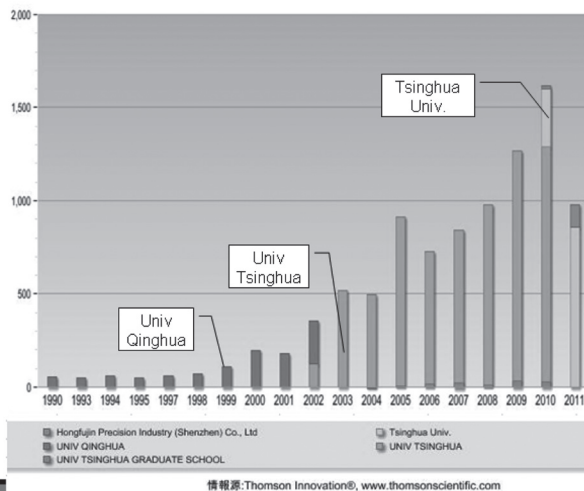


図9 清華大学の英文表記の移り変わり (Thomson Innovationによる統計解析より)

グラフからは2001年まではピンイン表記である“Qinghua”が使われ、2002年の途中から表記を“Tsinghua”に切り替えたことがうかがえる。なお、最近、“Univ Tsinghua”から“Tsinghua Univ.”に変わった点も検索の際は留意する必要がある。

フォックスコンや清華大学のように多くの出願件数を有する著名で大規模な出願人ですら、特許情報においてその英文出願人表記は統制されていない。調査担当者の立場から見ると、中国企業あるいは大学・機関の特許出願を調査する場合、該当する企業・大学のホームページで用いられている英文表記を使用するのが最も早道であるが、それだけでは不足であることは明確である。

その他に考慮すべきなのは、①ピンイン表記、②ブランド名や関係会社、③過去に名称(表記)の変更が無い、④綴りミス、スペース位置な



**本文の複製、転載、改変、再配布を禁止します。**

どの単純な表記ゆれ、などがある。しかし、どれだけ注意を払っても、出願ごとに都度異なる英文表記が使われているケースや機械翻訳をかけたような出願人名もあり、とりわけ新しい研究機関や小規模な企業に関しては英文DBで出願人検索を行うには限界がある。

その場合有効なのは、出願人を原語すなわち中文で検索する機能となる。現在、中文で出願人検索が可能なDBはCNIPR, HYPAT-, PAT-LIST CN/WEB等の中文DBの他、英文DBに原語情報を付加したハイブリッド型DBとしてPATBASE, ORBIT.comなどがある。

**(2) 中国外企業名称の表記ゆれ**

続いて、欧米出願人の中国出願時の中文表記につき検証する。

表3はSIPO統計データより得られた2010年の中国特許出願件数（外国企業）上位5社のデータである。英文名を見れば誰もが知る大企業ばかりであるが、中文名からは簡単に企業名を連想できないものも多い。

**表3 2010年中国特許出願件数ランキング  
(外国企業) SIPO統計データより**

順位	英文名	中文名	国	件数
1	Sony Corporation	索尼株式会社	JP	2,036
2	Panasonic Corporation	松下电器产业株式会社	JP	1,723
3	Qualcomm Inc.	高通股份有限公司	US	1,446
4	Samsung Electronics Co., Ltd	三星电子株式会社	KR	1,219
5	Sharp Corporation	夏普株式会社	JP	1,202

中国に市場を有する多くの国外企業が中国向けのウェブサイトを持つため、ここで企業名の中文表記を調べることができる。また、中国に1件以上の出願があれば、英文DBで中国の特許番号を特定し、その公報に記載された中文

表記を確認できる。

この2つの方法で欧米の大手製薬企業3社の中文表記を調べ、さらにCNIPRの英文中文各DBで、英文・中文各出願人表記で得られる公報数を調べた。この結果を表4に示す。

**表4 欧米製薬大手企業の中文表記**

CNIPR(英文)	公報数	CNIPR(中文)	公報数
Novartis	943	诺华(HP)	32
		诺瓦提斯	1019
Eli Lilly	351	礼来(HP)	0
		伊莱利利	507
Hoffmann la Roche	524	霍夫曼-拉罗氏(HP)	7
		霍夫曼-拉罗奇	383

※IPC：A61K31とAND演算した回答数

この結果より、ウェブサイトに記載された中文表記が必ずしも出願時の表記に使用されていないことがわかる。ここからは推察となるが、出願時の表記は発音の当て字に近く、しかも代理人によって表記にばらつきが見られる。これに対し、企業が自らウェブサイトで称する公式中文名は、中国の消費者・ユーザー向けに文字から連想させるイメージも考慮したものとなっているようである。

なお、このことは欧米出願人だけでなく、日本出願人の中文表記の際にも共通して留意すべき点である。社名にカタカナやアルファベットを有する企業名の場合は、当該企業が公表する中文企業名のみを検索するのでは不足であり、実際の中国公報で使用された表記を、年代を分けて複数件確認するのが無難であろう。また、漢字表記の日本企業であっても、日本で使用される漢字と簡体字とは異なるケースが多いため、中文表記を複数件確認する必要があることは言うまでもない。

さらには、国外親会社の中文表記を中国現地法人の社名に利用したケースで、その現地法人の中国出願情報が英文に訳された際に二度翻訳

本文の複製、転載、改変、再配布を禁止します。

を経たことを原因として、元とは異なる英文表記となるという複雑なケースも見受けられた。

例)  
 三菱 (Mitsubishi) → SanLing (ピンイン);  
 SANLING ELEVATOR CO., LTD. (CN)  
 DuPont → 杜邦 → DuBang; DUBANG  
 HUAJIA CHEMICAL CO., LTD. (CN)

以上のことから、カナ・アルファベットの出願人名検索の際にも英文と中文の両方での検索が望ましいといえる。

(3) 共同出願人データの欠落

これまでは中国語の翻訳に由来する出願人検索時の留意点を述べたが、中国特許調査における出願人検索において留意すべき事実として、英文DBにおける共同出願人データの欠落の問題がある。

図10は、予め2社以上の共同出願特許を選び、それらがDB上どのように出願人表記されているかを調べた結果である。中文DB上は複数の

共同出願人データが漏れなく収録されている案件がSIPO, CNIPR, Shareresearch等の英文DB上では、筆頭出願人だけしか収録されていないことが確認できた。図10に記載の英文DBの他にもEspacenet等、計5種類のDOCDB系の英文DBを検証したが、結果は全て筆頭出願人1社だけの収録であった。

この原因を図2記載のデータのフローを遡って検証すると、IPPHが作成する英文データの時点で2社名以降の出願人名に抜けが生じていると推測できる。このため、「共同出願人データの欠落」は殆ど全ての英文DBにとって共通の問題として存在するといえる。

侵害予防調査として特定の企業が保有する知財権を調査する場合は言うまでもなく、出願動向調査、あるいは大学等の共同研究先を知る上でも共同出願人情報が欠落していることは致命的な「調査漏れ」につながるリスクがあり、このような調査漏れを防ぐためには、出願人をキーとした調査を行う際には、中文DBを使用する必要がある。

公報番号	出願人1	SIPO(中文) 出願人2	SIPO(英文)	CNIPR(英文)	SHARE標準化出願人
CN1130135A	刘本林	常州市电子工业职工大学	Liu Benlin	Liu Benlin	LIUBENLIN (CN)
CN1202708A	住友 wiring 株式会社	住友电气工业株式会社	Sumitomo Wiring Systems, Ltd.	Sumitomo Wiring Systems, Ltd.	SUMITOMO WIRING SYSTEMS (JP)
CN1178402A	住友 wiring 株式会社	住友电气工业株式会社	Sumitomo Wiring Systems, Ltd.	Sumitomo Wiring Systems, Ltd.	SUMITOMO WIRING SYSTEMS (JP)
CN2625185Y	武汉力兴测试设备有限公司	武汉力兴绿色动力科技有限公司	Wuhan Lisun Testing Equipment Co., Ltd.	Wuhan Lisun Testing Equipment Co., Ltd.	WUHAN LISUN TESTING EQUIPMENT (CN)
CN2638255Y	清华大学	中国南车集团株洲电力机车研究所	Tsinghua Univ.	Tsinghua Univ.	UNIV TSINGHUA (CN)
CN2647706Y	上海燃料电池汽车动力系统有限公司	万帮	Shanghai Fuel Battery Vehicle Dynamic System Co., Ltd.	Shanghai Fuel Battery Vehicle Dynamic System Co., Ltd.	SHANGHAI FUEL BATTERY VEHICLE (CN)
CN2476126Y	范汉强	国家电动汽车试验示范区管理中心	Fan Hanqiang	Fan Hanqiang	FAN HANQIANG (CN)
CN1535307A	日产汽车株式会社	新日本石油株式会社	Nissan Motor Co., Ltd.	Nissan Motor Co., Ltd.	NISSAN MOTOR (JP)
CN1322029A	松下电器产业株式会社	丰田汽车株式会社	Matsushita Electric Industrial Co., Ltd.	Matsushita Electric Industrial Co., Ltd.	MATSUSHITA ELECTRIC IND CO LTD (JP)
CN1397447A	上海燃料电池汽车动力系统有限公司	万帮	Shanghai Fuel Battery and Car Power System Co. Ltd.	Shanghai Fuel Battery and Car Power System Co. Ltd.	SHANGHAI FUEL BATTERY AND CAR (CN)
CN1417874A	清华大学	北京市世纪博纳能源技术有限公司	Tsinghua Univ.	Tsinghua Univ.	UNIV TSINGHUA (CN)
CN117321A	C.R.F.阿西安尼顾问公司	国家研究议会	C. R. F. Societa consortile Per Azioni	C. R. F. Societa consortile Per Azioni	FIAT RICERCA (IT)
CN1452793A	霍克能源产品公司	联邦科学和工业研究委员会	Hawker Energy Products Inc.	Hawker Energy Products Inc.	HAWKER ENERGY PRODUCTS INC (US)
CN2600270Y	同济大学新能源汽车工程中心	万帮	New Energy Vehicle Engineering Centre, Tongji	New Energy Vehicle Engineering Centre, Tongji	NEW ENERGY VEHICLE ENGINEERING (CN)

出願人が複数(共願データ有り)

出願人が単独(共願データ無し)

SIPO(中文)やCNIPR(中文)では共願人のデータが入っているが、SIPO(英文)、CNIPR(英文)、Espacenetその他英文DBでは筆頭出願人データしか入っていない。

図10 DBによる出願人収録状況の違い

本文の複製、転載、改変、再配布を禁止します。

なお、英文DBであっても以下の2通りのDB検索で共願情報を得ることができる。

① 英文／中文ハイブリッドタイプDBの使用  
PATBASE, QPAT (ORBIT.com) および Total Patentがこのタイプにあたる。これらは英文DBであるが、原語データも有しており(図11参照)、PATBASEおよびQPAT (ORBIT.com) においては検索も可能である。Total Patentについては中文データを有する点は同じ

であるが、残念ながら今のところ表示機能のみであり、検索は不可能である。

② データの付加価値付与機能を有するDBの使用

トムソン・イノベーションおよびDWPIが相当する。前者に関してはちょうど本研究が進められていた2010年の12月に共同出願人情報の追加収録を開始した。順次遡及の予定であり、またDWPIのレコードにも情報の反映を開始した。

5) PatBase 番号: 40874722 (CN101017917A)	
タイトル:	Electrolyte for the lithium ion battery
優先権:	CN200710064066 20070226
優先権タイプ:	
ファミリー:	発行番号 発行日 出願番号 出願日 リンク
ファミリーエクスプローラー:	CN101017917 A 20070815 CN200710064066 20070226
譲受人:(標準):	USTC UNIV SCIENCE TECH CN
譲受人:	[EN] UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA [CN] 中国科学技术大学; 合肥国科行高科技动力能源有限公司
発明者:(標準):	CHEN CHUNHUA XIANG
発明者:	[EN] CHUNHUA XIANG CHEN [CN] 张承平; 陈春华; 项定炎
国際分類(IPC 8-9):	C09K21/06 H01M10/40 H01M6/16 (Advanced/Invention); C09K21/00 H01M10/36 H01M6/16 (Core/Invention)

図11 PATBASEレコード表示例

### 3. 3 キーワード検索の留意点

#### (1) データソース毎の留意点

中国特許調査の精度は、2009年の全文英訳DBの登場により劇的に向上した。それまで多くのDBが、SIPOがDOCDBに提供する英文抄録を収録していたが、現在は多くのDBがCNIPR系全文データを収録するようになった

表5 データソースと中国特許DB (英文)

種別	代表DB	要約	請求項	全文	特徴
SIPO系DB	SIPO, espacenet, PATBASE, PATOLIS 海外, Shareresearch (スタンダード)	○	×	×	SIPOからDOCDBに提供される英文要約データを収録する。古い年代には極端に簡素化された要約があったり、抜けが見られるなど注意が必要。
CNIPR系DB	C-PAT, QPAT, Total Patent, Shareresearch (オプション), JP-Net, NRIサイバーパテント2	○	○	○	IPPHが販売する中国特・実全文の英文機械翻訳データを収録する。1985年以降の登録／公開特許、実用新案全件の英文データがある。全文系DBのほとんどがこのソースを採用しているが、DBによっては収録開始時期に違いがあったり、データに抜けがあるため注意。
独自翻訳系DB	Chinese Patent Fulltext (Dialog File 325)	○	○	○	SciPatが独自に作成する英文機械翻訳データ。1985年以降の登録／公開特許、実案の全クレームをステータス情報と共に収録する唯一のDB。ただし、OCRによる紙公報の読み取り後機械翻訳にかけているため、誤読み取りが多い、ファミリー情報なし、出願人タイムラグが大きいなど若干の課題有り。
独自抄録系DB	DWPI, CA, Thomson Innovation (TI) アジアコンテンツ	○	△	×	要約は各々の作成基準で索引あるいは抄録化されている。クレーム情報を有するのはTIアジアのみであるが、2007年以降の公開および実用新案を人手翻訳している唯一のDBであり、翻訳精度は最も高い。



**本文の複製、転載、改変、再配布を禁止します。**

(図2参照)。この2つの情報源の他にも、独自抄録系DBとOCRで読み込んだ中文公報を独自の機械英訳で英文化した全文DBがある。表5に主要英文DBをデータソース別に分けて示した。なお、ここでは検索対象として要約・請求項・全文が含まれるか否かを示す。

表5に示すようにデータソースによって検索対象となるテキストの量、質に大きく違いがあるので注意が必要である。

2008年以前はSIPO系DBが殆ど唯一の英文情報源であったが、2004年頃までは要約の抜けが多く、内容も1～2行の極端に簡素なものが見受けられたりと、特許検索に耐えうるものではなかった。これが中国特許調査におけるキーワード検索の信頼性を低下させた一因と考えられるが、その後SIPO要約の抜けはCNIPR要約で補填され、要約の収録率・情報量は大きく向上している。

ただし、要約の収録率がほぼ100%に改善された現在であっても、検索対象が要約のみでは特許調査として不十分といわざるを得ない。

要約と請求項の内容を対比したところ、多くの場合、要約は請求項1の内容を反映したものであるという知見も得られており、よって、請求項の数が多いものほど要約では権利内容をカバーしきれていないと判断できる。このことから調査の精度を向上させるためには請求項が検索対象となるDBの使用が望ましい。

**(2) 英文検索と中文検索**

次に英文DBと中文DBを同じ意味のワードで

検索した際の違いを検証する。

事例1) カーボンナノチューブ

「カーボンナノチューブ」に相当する英文と中文とで、英文DB、中文DB各々の検索を行った。比較対象としたのは、2008年および2009年の中国公開および実用新案件数である。英文DBは前項データソースごとに1つずつ選択し、中文DBとしてはCNIPRを使用した。

英文は“CARBON NANOTUBE” および“CARBON NANO TUBE”(複数形含む)を用い、中文は“碳納米管”を用いた。結果を表6に示す。

中文DBと各英文DBは全て異なる件数を示したが、概して中文DBの再現率が高く、タイトル、要約、請求項、全文とテキストの量が多くなるほどその差異は広がる傾向にある。

なお、「独自抄録系DB」におけるタイトルおよび要約は、独自の作成基準で発明の内容を簡潔に示すために作られたものであり、公報のタイトル、要約よりも情報量が多い。タイトル、要約についてこの再現率が最も高いのはそのためである。

中文で抽出されたにも関わらず英文で抽出されなかったレコード内の表記としては、“carbonaceous millimicro tubule”, “carbon nanometer pipe”, “carbon nano pipe” などが見受けられた。

このように、中文から英訳される際に汎用とは言い難い英単語に翻訳されることは多く、さらに、英文DB同士でも機械英訳システムが異なると回答件数に大きな差が生じる。

**表6 同意のワードによる中文DBと英文DB件数比較 (カーボンナノチューブ)**

	CNIPR (中文)		SIPO系DB (英文)		CNIPR系DB (英文)		独自翻訳系 (英文)		独自抄録系 (英文)	
	2008年	2009年	2008年	2009年	2008年	2009年	2008年	2009年	2008年	2009年
タイトル	265	289	235	284	234	283	273	290	379※	545※
要約	426	581	342	540	396	545	409	561	593※	926※



## 本文の複製、転載、改変、再配布を禁止します。

同様の検証を「自動車」「太陽電池」および幾つかの医療用語にて行ったが、結果はいずれの場合も中文検索の方が回答数が多く、「太陽電池」においては中文検索が約25%増の結果を得た(表7参照)。中文検索で得られて英文検索で得られない案件の表記を確認すると、“cell in solar module”, “Solar energy battery”, “silicon photocell”等の語句が用いられていた。

表7 中文検索と英文検索の比較 (太陽電池)

キーワード	検索範囲：2009年公開	
	検索対象：名称	検索対象：要約
太陽電池	特許：692件	特許：1,210件
太陽電池	実案：269件	実案：1,062件
solar cell	特許：647件	特許：1,024件
solar battery	実案：228件	実案：661件

以上のことから、再現率が重要な検索においてキーワードを使用する場合、可能であれば中文検索が望ましく(中国語を母国語とする専門家により吟味・選定されたワードを用いるのが望ましいことは言うまでもない)、英文検索を行う場合には、翻訳により起こる独特の表記や誤訳を予想した上で通常使用される語句+ $\alpha$ の用語検索が必要といえる。なお、中文検索と英文検索の結果を統合できるとなると望ましく、その場合には出願人検索時と同様に、英文/中文検索の結果を統合可能なハイブリッドタイプのDBが有効である。

### 3.4 翻訳精度の検証

これまでに述べた各項目の留意点に注意しながら検索を実行したとして、最後の重要ステップとして回答のスクリーニングが必要となる。

そこで、翻訳された英文DBの内容がどの程度スクリーニングもしくは解析に耐えうるかを検証・評価した。

#### 【検証方法】

(1) 自動車, 化学, 医薬, 電気の4分野におけ

る中国公開特許を抽出した。

- (2) 各々につき、いくつかのDBで英語/日本語の要約を取得し、対応日本出願の要約と対比しながら翻訳精度についてワーキングメンバーの評価を纏めた。
- (3) 無料翻訳サイトを使用した中→日訳についても評価を行った。

使用したDBおよび翻訳ソフトを表8に、取得した抄録一例を表9に、翻訳精度の検証結果を表10にまとめた。

表8 検証に用いたDBおよび翻訳ソフト

DB	翻訳	翻訳エンジン
SIPO英文	中→英	CPMT (CPIC)
CNIPR英文	中→英	華建多語IAT
WPI抄録	中→英	人手翻訳
PatList-CN/WEB	中→日	KODENSHA
HYPAT-I/CN	中→日	クロスランゲージ
JP-Net	英→日	(東芝製のThe翻訳+一部カスタマイズ)
Sharesearch	英→日	PAT-TRANSEER (クロスランゲージ)
無料翻訳ソフト	原語	翻訳エンジン
Google翻訳	中→日	Google独自
Excite翻訳	中→日	KODENSHA
Infoseek翻訳	中→日	クロスランゲージ
livedoor翻訳	中→日	AMIKAI

検証したのは4件と少なく、分野よりの傾向も様々と思われるが、共通の認識として以下の知見が得られた。

- (1) 翻訳は中→日より中→英→日の方が良い。(英語を介すことにより文章として読みやすい日本語となる。用語の翻訳精度に大きな差異は生じない)
- (2) 技術内容理解は英語で行うのがベター(WPIまたはCNIPR由来の英文DBを利用)
- (3) 英文レコードをスクリーニングする際にも、化学式や構造図などのイメージ図を確認するために、原文公報にすぐにアクセスでき