

網羅的な中国特許調査のための「中国語の拾い方」

How to pick up Chinese for Chinese patent search exhaustive

アジア特許情報研究会 **伊藤 徹男**

PROFILE

2004～現在：日本知的財産協会セミナー講師、2006～現在：日本特許データサービスセミナー講師、2007～2009：検索競技大会委員、2008年：アジア特許情報研究会設立

✉ patentsearch2006@yahoo.co.jp

1 はじめに

ここ数年の間に中国の特許、実用新案、意匠の出願件数は異常なほどの伸びを示し、PCT 特許出願ランキング TOP10 に中国国内出願 1、2 位を争う中興通信 (ZTE) や華為技術 (Huawei) も入ってきて知財情報の世界を賑わせている。他方、中国国内出願特許や実用新案も登録となる前に、あるいは登録となったのちわずか数年で失効しているものも多く、したがって、権利ある有効な特許、実用新案を把握することが権利行使、紛争の未然防止には重要であり、無用の膨大な出願に戦々恐々とすることはないのである。

最近の中国出願においては、特に 2012 年、2013 年出願特許において、出願から 6 か月以内に公開となる特許がそれぞれ出願総数の約半分を占めるという異常な事態が顕出していることである (図 1)¹⁾。日本特許においても出願から 6 か月以内に公開となる特許がないことはないが、その数は極めて少なく、出願から 18 か月後に公開になるケースがほとんどである (図 2)。

中国では、特許は公開されないと審査が始まらず、登録にもならないから、日本や韓国のように公開前登録特許というものはないが、2012 年 8 月に施行となった「優先審査請求制度」の影響とも相まって早期に公開、登録となるこのような特許をいち早くキャッチして侵害予防などの対応を取ることが求められる。

これら中国特許情報 (実用新案を含む) を把握するツールとしては、欧米特許を調査する際に利用される商用英語データベースが一般的である。商用英語データベースも数年前までは中国をはじめとする東アジア諸国のデータ収録が不十分であり、収録のタイムラグも大きかったが、ここ 2、3 年で、特に中国特許情報の収録状況が目を見張るほど改善され、各国特許庁が提供している原語データベースは利用するまでもない、というような感さえ与えるようになってきた。

しかしながら各国原語から英語への翻訳精度不良などもあって、商用英語データベースだけでは網羅的な調査ができず、権利の抵触関係調査 (侵害予防調査) や無効化資料調査などでは、やはり各国特許庁原語データベースは英語データベースを補完するツールとして手離せな

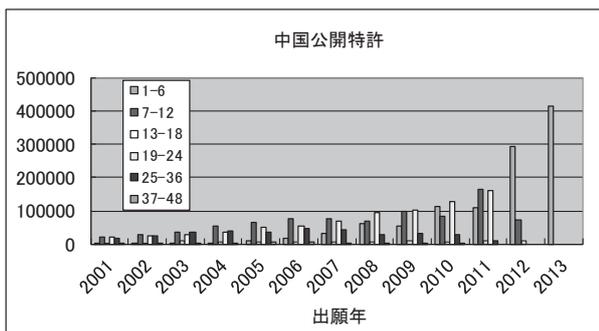


図 1 早期公開特許の推移 (中国)

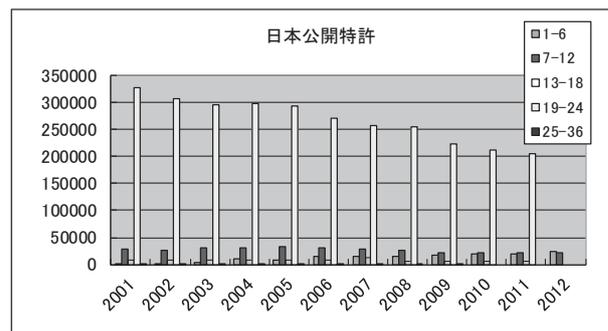


図 2 早期公開特許の推移 (日本)

いのが実情である。ましてやアセアン諸国など商用英語データベースの収録や翻訳精度がさらに問題となっている地域の特許情報については、各国特許庁原語データベースは必須のツールとなっている。

このような観点から本稿では、中国特許調査に焦点を当てて原語検索が必要となる「中国語」を、中国語が読み書きできないサーチャーにも効率的に抽出できる方法について実例を交えながら紹介することとしたい。最初に、中国語（原語）での検索が必要な理由について述べたのち、中国語抽出方法の本題に入ることにする。本稿で取り扱う中国語とは、大陸で使われる簡体字を対象とし、台湾で使われる繁体字については扱わない。

2 中国特許調査で原語(中国語)での検索が必要なわけ

中国特許調査が欧米特許調査と同様に、何故英語データベースだけで処理できないのか、それは偏に中国語から英語への翻訳精度の問題に尽きる。特に、各国特許庁から送られた英語データベースを元にしてデータベースを構築している DOCDB を利用している商用英語データベースの誤翻訳が目立つ。

DOCDB は約 90 の国 / 機関で発行される英語特許文献の書誌情報等を含むデータベースで、各国 / 機関から送付された情報をもとに欧州特許庁が編集・提供しているものである。したがって、各国特許庁で各国原語から英語に翻訳された情報は、翻訳ミス、スペルミスと共に基本的にはそのまま DOCDB として発行される。最近では、スペルミスなどは修正される場合もあるとのことであるが、後述するように中国特許庁英語データベースでのスペルミスがそのまま商用英語データベースに受け継がれていることが多い。

これは発明の名称、出願人などの書誌事項だけでなく、要約、特許請求の範囲なども誤翻訳や不適切な表記で英語に翻訳されていることが英語データベースだけでは網羅的な検索ができない要因の 1 つとなっている。

中国特許庁英語データベースでは、出願人については 2010 年 2 月 17 日公開・公告分までは共願の場合でも筆頭出願人しか表記してこなかったために、ほとんどの商用英語データベースも筆頭出願人しか収録されてい

ない（但し、ファミリー系のデータベースの場合にはいくらか共願人も収録されるなどしているが、誤って異なる共願人を収録したりして混乱も多い）。中国特許庁英語データベースでは翌週の 2010 年 2 月 24 日公開・公告分以降から共願人も併せて表記するようになったために、商用英語データベースの共願人収録が一気に増加することとなった（表 1）。

この点を、毎年、子会社を筆頭出願人として多数の出願をしている HON HAI PRECISION（鴻海精密工業）の例で表 2 に示す。表 2 に於いて、SIPO は中国特許庁英語データベースであり、「DB A」は DOCDB 系の商用英語データベース、「DB B」「DB C」がファミリー系の商用英語データベースである。表最右列が中国語データベース CNIPR による中国語名検索である。この表からも明らかなように SIPO の収録情報がそのまま商用英語データベースの収録に影響を与えている。ファミリー系データベースの 2009 年以前の収録数が SIPO より多くなっているのは外国出願によるファミリーからの情報を収録しているためであり、2010 年以降、中国語データベース CNIPR の値より多くなっ

表 1 SIPO 英語 DB の共願人表記

| |
|--|
| ① CN200810144132.6 申請日 :2008.07.29 公開(公告)日 :2010.02.03 |
| CNIPR : 台湾薄膜电晶体液晶显示器产业协会 ; 中华映管股份有限公司 ; 友达光电股份有限公司 ; 瀚宇彩晶股份有限公司 ; 奇美电子股份有限公司 ; 财团法人工业技术研究院 ; |
| SIPO 英語 DB : Taiwan TFT LCD Association (筆頭出願人のみ) |
| ② CN200810178280.X 申請日 :2008.11.19 公開(公告)日 :2010.06.16 |
| CNIPR : 台湾薄膜电晶体液晶显示器产业协会 ; 中华映管股份有限公司 ; 友达光电股份有限公司 ; 瀚宇彩晶股份有限公司 ; 奇美电子股份有限公司 ; 财团法人工业技术研究院 ; |
| SIPO 英語 DB : Industrial Technology Research Institute ; HannStar Corp. ; Chimei. Innolux. Corp. ; TTLA Taiwan TFT LCD Association ; AU Optronics Corp. ; Chunghwa Picture Tubes Ltd. |
| ③ CN200880124337.6 申請日 :2008.07.30 公開(公告)日 :2010.12.08 |
| CNIPR : 三菱重工业株式会社 ; 常石控股株式会社 ; 日立造船株式会社 ; |
| SIPO 英語 DB : Mitsubishi Heavy Ind. Ltd ; Tsuneishi. Holdings Corp ; Hitachi Shipbuilding Eng. Co. |



表2 共願人の検索

| 公開年 | Hon Hai + Hong Hai + Honghai Prec% | | | | 鴻海精密 |
|------|------------------------------------|------|------|------|-------|
| | SIPO | DB A | DB B | DB C | CNIPR |
| 1997 | 4 | 4 | 4 | 4 | 4 |
| 1998 | 32 | 32 | 34 | 32 | 33 |
| 1999 | 47 | 49 | 54 | 53 | 52 |
| 2000 | 1 | 1 | 13 | 11 | 37 |
| 2001 | 0 | 0 | 46 | 42 | 124 |
| 2002 | 0 | 1 | 16 | 17 | 64 |
| 2003 | 0 | 1 | 28 | 27 | 112 |
| 2004 | 1 | 1 | 55 | 56 | 328 |
| 2005 | 0 | 0 | 97 | 99 | 593 |
| 2006 | 1 | 1 | 441 | 445 | 791 |
| 2007 | 4 | 5 | 1006 | 1011 | 1372 |
| 2008 | 204 | 10 | 1142 | 1134 | 1233 |
| 2009 | 29 | 29 | 2120 | 2007 | 2155 |
| 2010 | 1898 | 1898 | 2391 | 2347 | 2304 |
| 2011 | 2835 | 2835 | 3006 | 2935 | 2832 |
| 2012 | 3607 | 3607 | 3714 | 3665 | 3607 |

ているのは、鴻海精密工業の100%子会社である鴻富錦精密工業（Hngfujin Precision Industry）と精華大学（Tsinghua Univ.）の共願を鴻海精密工業にカウントしたり、「HONG HAI」で検索すると「Shenyang Honghai Fine Chemicals」や「Shi Honghai」という別出願人まで抽出しているためである。因みに、2010年の「Hon Hai Precision」表記での「DB B」「DB C」の抽出件数はいずれも2230件である。

次いで用語のスペルミスが商用英語データベースに影響している1例を紹介する。2000-2013年の公開特許で「発明の名称」に限定して、上記、DB AおよびDB CとCNIPRについて、シリコン系化合物である「silsesquioxane」についての抽出比較を行った（表3）。

DB Aは後方一致検索ができないため（商用英語データベースの多くは後方一致検索ができない）「polysilsesquioxane」や「polyalkylsilsesquioxane」などが抽出できずヒット件数が少ない結果となっている。

今回の検証範囲内では、スペルミスを含む異表記の例として「sesquialter siloxane」「sesquisiloxane（sesqui siloxane）」「polysiloxane」など14種の表記があり、これらの異表記はSIPO英語DBと同じであることから商用英語データベースのデータソースは中

国特許庁から提供され、DOCDBとして構成されたものであることが推察できる。

表3 SIPO英語DBスペルミスの商用DBへの影響

| | CNIPR | DB A | DB B |
|---------------|-------|------|------|
| 抽出全件数 | 257 | 181 | 224 |
| 各DBのみで得られるもの | 33 | 0 | 3 |
| 後方一致でえられたもの | 36 | 0 | 39 |
| 異表記（スペルミスを含む） | 2 | 41 | 36 |
| 各DBをマージしたもの | 261 | | |

DB A：silsesquioxane*

DB B：*silsesquioxane*（後方一致検索が可能）

CNIPR：%silsesquioxane% or 倍半硅氧烷 or 硅倍半氧烷

「lithium secondary battery」などの電池分野においても「lithium」が「litium, lithim, lithum, lihtium, lithiam」など、「secondary」が「secandary, secondery, secndary, secodary」など、「battery」が「battary, secondary batter, bottery, batery」などのようなスペルミスがあり、これらのスペルミスが商用英語データベースに修正されることなく収録されていることは悲しいことである。正しい情報への修正努力を期待したい。そしてこのような事実（商用英語データベースにはスペルミスもあるということ）に気がついていないサーチャーもおられることと思う。

出願前の先行技術調査や技術動向調査では、数件の漏れが生じて大きな問題とはならないが、権利の抵触性判断や無効化資料調査では、数件の漏れも許されないことが多いために放置できない。

ここでは中国語から英語への翻訳精度については触れなかったが、特に、外国出願していないファミリーのない中国国内のみの出願（中国ではこのような内国のみの出願がほとんどを占める）に翻訳不良が目立つ点については稿を改めて触れたい。

中国特許庁英語データベースの収録方針、スペルミス、翻訳精度の問題などから英語データベースだけでは網羅的な調査ができず、中国語でも検索しなければならない理由がここにある。

つまり、英語データベース+中国語データベースで一般的な異表記を考慮して検索すれば、スペルミスを気にせずに網羅的な調査が可能となる。

3 中国語特許データベースの進歩

商用英語データベースもさることながら中国語特許データベースも、この1、2年で大きな進歩を遂げている。現在、日本で使われている中国語特許データベースには以下のようなものがある。

- 1) SIPO 中国語データベース (中国特許庁)
- 2) CNIPR (中国知識産権出版社) 中国語版、日本語版
- 3) PSS-SYSTEM (中国特許庁)
- 4) CPIC (中国專利信息中心)
- 5) HYPAT-i (発明通信社)
- 6) 專利 SERARCH-i2 (アイ・ピー・ファイン社)
- 7) PAT-LIST CN/WEB (レイテック社)
- 8) Orbit.com (QUESTEL 社)、PatBase (RWS 社)・・・英語と中国語で検索できる。

CNIPRの一部や5)、6)、7)、8)は有償のサービスである。その他、中国各地の省、地域の知識産権局から、あるいは soopat (<http://www.soopat.com/>)、大为 (<http://www.daweisoft.com/>)、彼速 (<http://www.bizsolution.com.cn/>) といった中国系サイトなどから多くの中国語データベースが発行されている。後2者は、検索式の保存、一括ダウンロードなども可能であり、中国国内で多用されているとのことであるが、収録のタイムラグがある。収録のタイムラグの点では中国特許庁の中国語データベースや中国知識産権出版社のCNIPRにはかなわない。残念ながら中国特許庁の中国語データベースは、簡易な検索は可能であるが、検索、表示、出力機能の点から実務では使い難い。無料でも検索や表示機能など業務に使えるレベルのCNIPRが急速に日本のユーザーに普及しつつある。

そのような中、2014年1月に日本のユーザー向けにCNIPRの日本語とも言えるサービスが有料で始まった。発行元の中国知識産権出版社は「日本語版CNIPR」²⁾としてリリースしたが、検索インターフェースなどが日本語で表示されているだけであり、日本語で検索できる訳ではなく、「日本語で検索できるツール」と勘違いされやすいので、本稿では「日本語版CNIPR」と呼称する。

従来から存在するCNIPR(以下、中国版CNIPRと称する)についての最近の機能強化については1昨年の本誌で紹介したが、その中国版CNIPRと日本語版CNIPRとの主な検索・表示・ダウンロード機能を比較したのが表4-1と表4-2である(差異がある主な機能のみ)。

表4-1 CNIPR 検索項目比較

| 検索項目 | 中国版 CNIPR | 日本語版 CNIPR |
|--------------|-----------|------------|
| 優先権番号 | ○ | |
| 筆頭分類 (IPC) | ○ | |
| 全分類 (IPC) | ○ | ○ |
| 法律状態 | | ○ |
| TI+AB 検索 | | ○ |
| TI+AB+CL 検索 | | ○ |
| 国コード | ○ | |
| ファミリー | ○ | |
| 概念検索 | ○ | |
| 同義語検索 | ○ | |
| クロス原語検索 (En) | ○ | |
| 中英ハイブリッド検索 | | ○ |
| 生死検索 | | ○ |
| 失効検索 | ○ | |
| 法律状態検索 | ○ | ○ |
| ・ 権利移転 | ○ | 法律状態で検索 |
| ・ 質押、保全 | ○ | 法律状態で検索 |
| ・ ライセンス | ○ | 法律状態で検索 |

*TI: 発明の名称、AB: 要約、CL: 請求の範囲

表4-2 CNIPR 表示・出力項目比較

| 表示・出力項目 | 中国版 CNIPR | 日本語版 CNIPR |
|-----------|-----------|------------|
| 結果一覧生死表示 | ○ | |
| ワードハイライト | 検索文字 (1色) | 任意ワード5色 |
| ダウンロード | 中文 | 中文・英文 |
| | 2000件ごと | 5000件ごと |
| ダウンロード項目 | | |
| 国際出願番号 | ○ | |
| 国際公開番号 | ○ | |
| PCT 国内移行日 | ○ | |
| 分割出願番号 | ○ | |
| 生死状態 | | ○ |
| 法律状態 | | ○ |
| 全請求項 | | ○ |
| 解析機能 | ○ | |

日本版 CNIPR の有用な機能のいくつかを挙げてみよう。検索機能では、1) 中国語と英語でのハイブリッド検索が可能になったこと、2) 権利が存続しているか（生きているか）どうかを「有効」「無効」別に検索できるようになったこと、3) 審査経過情報や年金未納などの法律状態からも検索できるようになったこと（図3）、表示機能では、1) 検索結果一覧や書誌・要約画面で英語表示や日本語表示がボタン1つで可能になったこと、2) キーワードハイライトが5ワード（5色）まで可能になったこと（図4）、出力機能では、全請求項まで書誌事項と共に中国語一覧とは別に英語でも csv で生死の情報も付加され、ダウンロード可能になったこと、などを挙げる事ができる。



図3 日本版 CNIPR 検索画面

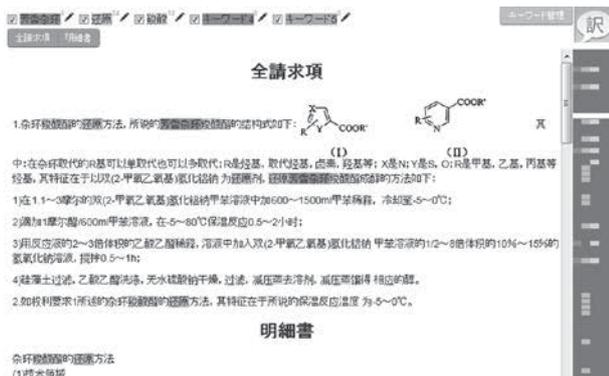


図4 検索キーワードハイライト表示

中国語と英語のハイブリッド検索機能の活用方法については次節で紹介するとして、まず、権利が存続しているかどうかの生死情報の検索について触れよう。中国版 CNIPR では生死情報を確認するためには「失効検索」という別の入口から入るデータベースがあるが、収録のタイムラグが大きいことなどから使いづらい（既に死んでいるものでもかなり長い期間生きている、とされてい

たり、復活したのもも収録されていなかったりするなど、本当に生きているかどうか確認が容易ではない）。

日本版 CNIPR では、図3に示したように検索画面上部に生死状態を検索できる「有効」「無効」のボタンがある。大量の情報が得られた場合には、生きている有効なものから査読し、失効しているものは後回しにするなど効率的な利用が可能である。生死情報のタイムラグについてはアジア特許情報研究会の検討において、「ほとんどタイムラグなく収録されている」ことを確認している（但し、年金未納で失効しているはずであるのに現時点では生きているものとされているなど情報の正確性について問題がある部分もあり、詳細については検討中である）。

中国版 CNIPR では生死状態の検索こそできないが、検索結果一覧画面で「登録になって生きているもの（「有効」：グリーン）」「審査請求前後の公開特許（「在申」：イエロー）」「失効特許（「无效」：レッド）」の表示がある（図5）。この表示機能は日本版 CNIPR にも欲しい機能の一つであり、中国知識産権出版社に要望を出しているところである。

その他、日本版 CNIPR には、翻訳機能として日本語と英語への翻訳機能もあるが、中国語から日本語への翻訳は難しく、英語翻訳の方が精度が高いと言われているが翻訳精度については未検証である。出力機能では全請求項まで中国語とは別に英語でも csv でダウンロードできることは大きな進歩である。



図5 中国版 CNIPR 検索結果一覧



図 8 ChemYQ (siloxane の例)

4) 中国語版言選 Web⁶⁾

キーワードの自動抽出システムとして知られている「言選 Web」に中国語版が用意されている。日本語版と同様に中国特許・実案の抄録、クレームなどを貼り付けると(図 9-1) キーワードを切り出してくれるので、そこに示されるワードを検索用語として活用する。特許文書中からのワードであるので確実に検索できるようにも思われるが、異表記までは示してくれない。

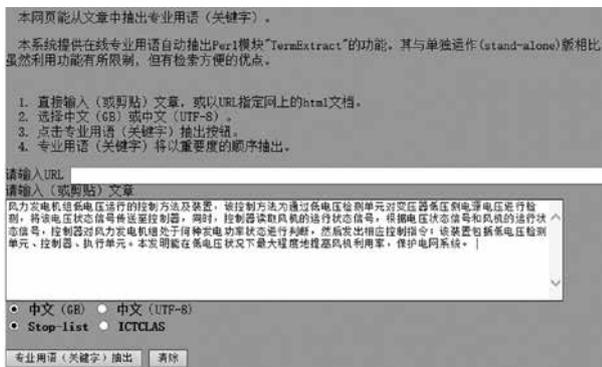


図 9-1 言選 Web 中国版



図 9-2 言選 Web 中国版で抽出した用語を Google 翻訳

5) 特許データベースから直接、抽出する方法

以上紹介した 1) ~ 3) の英中 Web 辞書は、英語を放り込めばそれなりの中国語を抽出してくれるが、所詮、特許データベースで使われている用語を網羅してくれる

訳ではないので最終的には中国語特許データベースから拾わないと網羅的な特許用語は集められない。

かつては SIPO などの中国英語データベースの「発明の名称」に英語ワードを入力し、その出願番号を基に中国語データベースで検索して中国語の「発明の名称」と対比させて抽出していたが、最近では、3. で紹介した日本版 CNIPR のハイブリッド検索機能を活用して容易に抽出できるようになった。有料データベースなので会社などで契約していない場合には、少々手間はかかるが、無料の英語データベースと中国語データベースを対比させて抽出する方法でももちろん問題ない。

ここでは日本版 CNIPR のハイブリッド検索による中国語ワード抽出の具体例を紹介する。

エネルギー分野の技術として最近注目されている配電技術として「smart grid」がある。この中国語抽出例を示そう。「smart grid」は Google 翻訳ツールでは「智能电网」と訳される(鄭州大学の辞書では出てこない)。これを日本版 CNIPR の発明の名称で「smart grid」と検索すると「智能电网」の他、「智能化电网」「智能网格」などが抽出できる。因みに、1985 年~2013 年公開特許の「発明の名称」で「smart grid」と検索すると 42 件の公開特許が検索される。「smart grid」を元に得られた「智能电网 or 智能化电网 or 智能网格」で「発明の名称」を検索すると 167 件の検索結果が得られる。したがって、英語と中国語のハイブリッド検索ができると言っても英語ワードだけでは不十分であることがわかる。

また、「智能电网」の異表記は、これら 3 つだけではない。「智能 (smart)」というワードを基に「smart grid」に付与されている IPC 分類 (H02J3、H02J13、G06Q50/06) を掛け合わせると以下の関連用語(異表記)を抽出でき、これらの用語で 1985 年~2013 年公開特許の「発明の名称」を検索すると、629 件の特許が抽出できた。

「智能电网、智能化电网、智能网格、智能配电网、智能配电、智能化配电、智能化电力配电、智能综合配电、智能供电、智能型变配电、微电网智能、智能微电网、智能变电、智能变电站、智能化变电站、智能变压器、智能用电、智能选相、智能环网、智能漏电、智能电力监控、

智能配变终端、供电智能化系统、智能充放电、智能交流配电、智能一体化电源、智能本安型电源网、供电智能系统、智能电气网络、网络智能型继电器、智能电源系统、智能电源监控、智能发电控制」

このように特許データベースからは Web 翻訳ツールでは抽出できない、しかも特許文書中に使われている用語の異表記を網羅して集めることができる。ここでは確認が容易なように「発明の名称」での抽出例を示したが、要約、特許請求の範囲、詳細な説明中から拾えば、さらに多くの異表記も集めることができるが、抽出の効率は悪くなる。

もう 1 例を化学分野のワードで紹介しよう。「HETERO AROMATIC (HETEROAROMATIC)」という用語がある。これも同様に日本版 CNIPR のハイブリッド検索で 1985 年～2013 年公開特許の「発明の名称」で検索すると 101 件がヒットし、「芳香杂环、芳香族杂环、芳杂环、杂芳、杂芳环、杂芳香、杂环芳香、杂芳族、杂芳基」などが見つかる。ここで得られた中国語で英語ワードと同様の条件で検索すると 783 件が得られる。

このように日本版 CNIPR のハイブリッド検索機能で英語ワードを入力することによって容易に該当する中国語異表記を見つけることができようになり、業務の効率もアップすることとなった。近接演算子を使ってさらに多くの同義語や類義語を集めることもできる。

5 中国語抽出時の留意点

ここでは以上のような英語ワードをもとに集めた中国語について、さらに異表記や同義語はないか、という観点から中国語を集める場合の留意点を挙げておこう。

中国語には、日本語と同じ表記、同じ意味でしかも日本語で入力可能な漢字がある。液体、温度、反射、密封材料などがそれである。次は意味内容が日本語と同一で簡体字となっているために最初はわからないにしても慣れれば読めるものがある。日本語 - 中国語で示すと、回転 - 回转、表示 - 显示、発光 - 发光、樹脂 - 树脂、紫外

線 - 紫外线のようなものである。そして日本語に存在する漢字であるのに意味内容が異なるものがある。やはり日本語 - 中国語で示すと、出力 - 输出、入力 - 输入、送信 - 发送 (發送)、受信 - 接收、自動車 - 汽车 (汽車) などである。「汽车」が「自動車」のことを指すというのは有名でお分かりの方も多いと思う。また、当然に中国だけにしかない漢字、意味内容のワードもある。こうなると全く手に負えないようにも思われるが、中国だけにしかない漢字、意味内容のワードも日本版 CNIPR の英中ハイブリッド検索や英中 Web 辞書で抽出して検索式に組み入れればよいのである。

本項では、中国語に独特の表記 (表現) 方法があることを紹介しておこう。それらを 3 パターンに分類した。

パターン 1: 用語の反転 (表 5)

表記された漢字が反転しているが、ほぼ同じ意味のもの。名詞的表記か動詞的表記かの違い。

パターン 2: 用語の短縮 (表 6)

用語の一部が省略されているが、ほとんど同じ意味のもの。

パターン 3: 中国語で多用される表現 (汎用語) (表 7)

同じ意味のワードであるが、日本ではほとんど使われないが中国語ではむしろ多用されるもの。

以下、パターンごとに、2000-2013 年の公開特許で「発明の名称」に現れる件数を表中に示した。

表 5 用語の反転

| | |
|----------------------|----------------|
| laminate(multilayer) | |
| 积层 / 层积 | 315 / 265 |
| 叠层 / 层叠 | 1823 / 2040 |
| 层合 / 合层 | 520 / 316 |
| 夹层 / 层夹 | 873 / 29 |
| bronchoscope | |
| 气管支镜 / 支气管镜 | 3(詳細説明中) / 25 |
| heat curable | |
| 可热固化 / 热可固化 | 36 / 4 |
| skin care | |
| 皮肤护理 / 护理皮肤 | 113 / 12 |

表 5 において、「叠」は「たたむ」、「夹」は「はさむ」という意味。「叠层」(層をたたんだもの)とか、「夹层」(層を挟んだもの)は名詞的で、反転した「层叠」(層をたたむ)、「层夹」(層を挟む)という用法は文法的には動詞的な使い方である。層を積み重ねたものが「积层(積



層)」というふうに理解すれば何となく multilayer であるというのも感覚的にわかる。しかし、「laminate」や「multilayer」を日本語で「積層、多層」などと表現しても「層積」などとは表現しない。

表 6 用語の短縮

| | |
|---------------------|--------------|
| aromatic | |
| 芳香族 / 芳族 | 781 / 910 |
| carbon electrode | |
| 碳素电极 / 碳电极 | 5 / 97 |
| endoscope | |
| 内窥镜 / 内镜 | 1722/139 |
| photoinitiator | |
| 光敏引发剂 / 光引发剂 | 30 / 262 |
| photoresist | |
| 光致抗蚀 / 光抗蚀 | 429 / 4 |
| polyimide | |
| 聚酰亚胺 / 聚亚胺 | 1309 / 6 |
| polymer | |
| 聚合物 / 聚物 | 16978 / 6347 |
| separator | |
| 隔离板 / 隔板 | 27 / 1037 |
| transesterification | |
| 酯基转移 / 酯转移 | 7 / 18 |

表 6 の短縮形は日本ではあまり使わないが、必ずしも短縮形の方が中国で多く使われているとは限らない。ある中国語を抽出した時には、このような短縮されたワードが使われているのではないかという疑問を持って異表記を抽出するとよい。

表 7 中国での汎用語

| | |
|----------------------|--------------|
| composition | |
| 组合物 > 组成物 | 82232 / 1017 |
| (photo)curable | |
| (光) 固化性 > (光) 硬化性 | 890 / 85 |
| solar cell(battery) | |
| 太阳能电池 > 太阳电池 | 7803 / 1278 |
| degradation | |
| 降解 > 分解 | 5734 / 2598 |
| chemically amplified | |
| 化学放大 > 化学增幅 | 92 / 47 |
| endoscope | |
| 内窥镜 > 内视镜 | 1722 / 62 |

表 7 の composition における「组合物」は、装置やシステムなど部品を組み合わせる、という意味での用法もないとは言えないが、樹脂组合物、農薬组合物、分散体组合物、化粧品组合物など、ほとんどの使い方は、日本での「組成物」と同一である。

以上のように、反転したワード（文法的には同一とは言えないが）や短縮したワードが異表記として使われ、日本で使う漢字表現とは異なる表記が異表記または同義語として汎用されるケースもあるので中国語を抽出する際には注意が必要である。

6 おわりに

本稿では、中国特許調査において、(商用) 英語データベースだけでは網羅的な調査ができず、検索ワードや出願人名に中国語を使った原語調査が必要であることを述べた。その中国語を抽出するには（中国語が読み書きできないサーチャーでも）、Web にある英中辞書から抽出する方法、英語と中国語の両特許データベースのワードを比較対照しながら抽出する方法などについて触れ、日本版 CNIPR のハイブリッド検索機能を活用すれば英語ワードから中国語ワードを容易に抽出できることも紹介した。

さらに、中国語を抽出する際には、中国語独特の表記があり、そのような点に留意して異表記、同義語を拾う必要があることにも触れた。

しかし、英語データベースより中国語データベースの方が網羅性が高いからと言って中国語のみで調査すればよい、というものではない。英語情報であれば技術者や研究者は理解できても中国語で査読できる日本人技術者や研究者は少ない。

したがって、英語データベースでの情報を基本にしつつ、英語情報では洩れる（検索できない）情報を中国語データベースで検索して抽出し、英語情報とマージすれば、中国語で査読しなければならないものはかなり少なくなる。

さらに、現状の商用英語データベースの中国特許収録

(実用新案も含め)は100%と言っても過言ではないので、英語情報とマージした結果、中国語のみから得られた中国特許の出願番号を英語データベースに再入力して検索すれば、ほとんどのものが英語データベースで得られる。これは上述したように中国語から英語に翻訳する際に生じた翻訳不良やスペルミスによるものである。このような機械翻訳英語による英語情報でも調査対象の特許であるのか、ノイズであるのかぐらいはほぼ判別できる。もちろん、権利の抵触関係などの微妙なものについては、中国語クレームなどを中国語に堪能な方に翻訳を依頼して査読する必要があることは当然である。

最後に、中国語での調査における外注について若干補足しておきたい。

中国語を抽出する場合には中国語の知識があった方がよいことは当然であり、可能であれば技術内容も理解できる中国人スタッフなどにサポートしてもらうのがベストであるが、中国語が読み書き、理解できない者でも上述したような方法で中国語を抽出して調査は可能である。しかし、調査に必要となる中国語を網羅して検索するような、そんな暇はない、それほど手間をかけるなら調査を外注した方が短時間で安上がりだ、と思われるかもしれない。

もちろん、1回限りの単発的な調査であればそれもいいでしょう。しかし、SDIのような継続して調査するテーマについては、一度作成した検索式を修正して再検索する場合など、社内で実施した方が小回りが利いて経済的でもある。継続して検索する場合でも最初の検索式だけ外注で作成してもらおう、という手もある。

そのような外注の場合には、外注先のスタッフが技術内容を理解して検索式を作成したり、調査しているかが重要である(「技術内容を理解して調査する」ということは、どんな調査の場合にも当てはまることではあるが)。したがって、中国語が理解できるスタッフまたは中国人スタッフに頼めば網羅的な中国特許調査ができるというものではない。そのような点では言語の障碍もない中国現地に調査依頼した方が網羅性の点でも完璧!、と思われがちである。しかし、調査スキルの点で、特に商用(英語)データベースを使った経験が少ない中国の代理人事務所や調査会社に依頼しても(それなりの報告

書を高額な請求書と共に送付してくれるが)、現状では満足な調査結果を得ることは難しい。むしろ、調査スキルに長け、中国特許調査を得意とする日本の調査会社に依頼した方が権利の抵触関係や無効化資料調査などは有効なことが多い。

参考文献

- 1) 伊藤 他: パテント Vol.67 No.8 (2014)
- 2) 日本版 CNIPR (<http://www.cnipr.jp/>)
- 3) Google 翻訳 ツール (<http://translate.google.com/>)
- 4) 鄭州大学: 科学用語辞典 (<http://www3.zzu.edu.cn/zzjdict/>)
- 5) ChemYQ (<http://www.chemyq.com/xz.htm>)
- 6) 中国語版言選 Web (http://gensen.dl.itc.u-tokyo.ac.jp/gensenweb_cn.html)

上記 URL および本文中の URL は、いずれも 2014 年 7 月 21 日に確認したものである。

