

## 台湾特許調査手法の検討:

### ～網羅的かつ効率的に台湾特許を調査する手法の検討～

○田畑文也<sup>1)</sup>, 石田政司<sup>2)</sup>, 桑原隆秀<sup>3)</sup>, 小山裕史<sup>4)</sup>

富士フイルム(株)<sup>1)</sup>, (株)神戸製鋼所<sup>2)</sup>, 住友金属鉱山(株)<sup>3)</sup>, 電気化学工業(株)<sup>4)</sup>

〒421-0396 静岡県榛原郡吉田町川尻 4000 富士フイルム株式会社<sup>1)</sup>

Tel: 0548-34-5401 FAX: 0548-32-8286

E-mail: fumiya\_tabata@fujifilm.co.jp

## Study of the Taiwan Patent Search.: Study of the Taiwan Patent Search with Completeness and Efficiency.

TABATA Fumiya <sup>1)</sup>, ISHIDA Seiji <sup>2)</sup>, KUWABARA Takahide<sup>3)</sup>, OYAMA Hiroshi <sup>4)</sup>

Fujifilm Corporation <sup>1)</sup>, Kobe Steel Ltd.<sup>2)</sup>, Sumitomo Metal Mining Co.,Ltd.<sup>3)</sup>, Denki Kagaku Kogyo K.K. <sup>4)</sup>

Fujifilm Corporation <sup>1)</sup>, 4000, Kawashiri, Yoshidacho, Hibara-gun, Shizuoka, Japan

Phone: +81-548-34-5401 Fax: +81-548-32-8286

E-mail: fumiya\_tabata@fujifilm.co.jp

### 【発表概要】

昨今、新興国特許調査が脚光を浴びている。本稿では台湾の特許調査について検討を行った。台湾特許について、商用DB(データベース)においては、DBの収録率、収録タイムラグ等に問題がある場合が多い。また、英語キーワードのみを用いた検索では、ほとんどのDBの場合、検索可能な範囲は要約までで、クレーム以下は検索できないことが多く、網羅的に調査することは難しい。台湾經濟部智慧財産局DB(以下、TWPAT)や、その他の商用DBで繁体字中国語を用いた検索での補完を含め、網羅的かつ効率的に、台湾特許を調査する手法について検討したので報告する。

### 【キーワード】

台湾特許、中国語検索、原語検索、繁体字、繁字体、網羅性、文字化け

## 1.はじめに

近年、台湾の技術力の進歩は半導体分野などの先端技術においても著しく、また中国で活躍している企業でも台湾系の企業も多く、台湾特許の動向を正確に調べることの必要性は、分野で状況は異なるものの、高くなっている。しかし、台湾特許の調査については、商用DBの収録も十分でない状況と報告されており<sup>[1][2]</sup>、特に商用DBでの調査の際は収録状況にも留意する必要がある。

まず、台湾特許、実案(実用新案)出願動向を特許庁統計から独自にまとめ、台湾からの出願、他地域(Others)からの出願に色分けし、図1に示す。

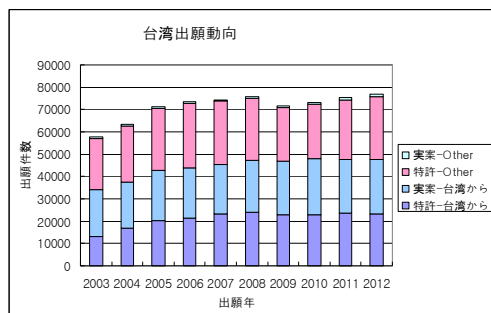


図1. 台湾特許・実案出願動向

2012年度でも、特許+実案では、約6割が台湾からの出願となっており、注意が必要である。

これらの特許をクレーム以下のキーワード検索を含め網羅的に調査する場合は、中国語(台湾の場合は繁体字)を併用した調査が必要になると考える。

また、台湾經濟部智慧財産局DB (TWPAT <http://twpat.tipo.gov.tw>) は、2013年7月に大幅なりニューアルがあり、機能アップが図られている。TWPATには中国語版と、英語版のDBがあるが、英語版では要約までしか検索できず、さらには登録特許については、2008年以降は要約が収録されていないため、検索用途としてはかなり限定的な使用しかできない。

中国語版 TWPAT についても、要約までと、クレーム以下の収録には、タイムラグが有り、特に SDI などに利用する場合は、後日再検索し、補完するなどの注意が必要である。(図2)

早期公開特許資料範囲		
資料項目	資料範囲	
特許書目資料	民國92年5月1日-民國102年7月16日(第11卷/第14期)	4
摘要 要約	民國92年5月1日-民國102年7月16日(第11卷/第14期)	4
說明書全文資料	民國92年5月1日-民國102年7月1日(第11卷)	4
公報	民國92年5月1日-民國102年7月16日(第11卷/第14期)	4

図2. 中国語版 TWPAT 収録

## 2.検討内容

台湾特許(実案を含む)を収録する商用DBが数多くあり、特に台湾に特化した中国語DBとしては、LearningTech社のWEBPAT (<http://webpat.tw/>)がある。また、英語のみのDBとしては、WPI(トムソン・ロイター社)やShareresearch((株)日立製作所)などがあり、英語だけでなく原語(中国語)も併用しハイブリッド検索<sup>[3]</sup>可能なDBとしては、PatBase(RWSグループ社)やOrbit.com(Questel社)などがあり、これらのDBについて、TWPATと比較し検証した。

### 2-1) 商用DB収録確認

各DB(A~Eと表記)について、TWPAT中国語版DBの収録を基準とし、英語DBについては、要約が収録されているものを、AB=YESのコマンドや、汎用語を用いた収録確認方法<sup>[4]</sup>を用い、公開特許、公告・登録特許、実案について収録を確認した。

### 2-2) TWPAT文字化け

台湾特許については、TWPATそのもので文字化け(図3)が発生する(表示フォントを設定変更することにより、文字として認識できる場合もあ

る)。萘(ナフタレン)、鎊(オニウム)、肱(オキシム)、呔喃(フラン)、醜(キノン)、喹啉(キノリン)について、TWPATでの文字化け挙動を発明の名称、要約フィールドで調べた。

申請番号	発明名称
092123296	環戊[b]D衍生物 CYCLOPENTAB[D]NAPHTHALENE DERIVATIVES
093113356	2-甲基D及D衍生物作為結晶改質劑之用途 USE OF QUINOLINE AND NAPHTHALENE DERIVATIVE MODIFIERS
093139948	新型含D團基之高折射率液晶化合物及其液晶組合物 NOVEL HIGH BIREFRINGENT LIQUID CRYSTAL COMPOSITIONS CONTAINING THE SAME
098103760	鎊-醜互嵌半導體聚合物 NAPHTHALENE-IMIDE SEMICONDUCTOR POLYMER

図3. 文字化け例(ナフタレン)

### 2-3) 中国語繁体字、簡体字混用

中国語には、台湾地域で使われている繁体字と、中国本土で使われている簡体字があるが、携帯電話を示す手機(繁体字)と手机(簡体字)の検索時の字体混用について調べた。

### 2-4) ハイブリッドDB原語収録

ハイブリッド型DBのDB-C, DB-Dについて、台湾公開特許(2003年～)、について、各年度 n=10 でクレーム収録を目視で確認した。

### 2-5) 台湾特許中国語クレーム検索

2010年～2013年発行の台湾公開特許を対象に、透明質酸(ヒアルロン酸)、糖醛酸(ウロン酸)、萘(ナフタレン)、膜(膜)について、中国語検索可能なDBで、中国語版TWPATを基準とした網羅性について調べた。

## 3. 検討結果

### 3-1) 商用DB収録確認

図4～図6に、台湾公開特許、公告・登録特許、実案の商用DB収録をまとめた。(2013.7.12収録確認)

収録タイムラグを許容するとしても、DB-A, B, C以外は、収録に問題が

あり、台湾特許を調査する場合は、用いるDBの選択も重要である。

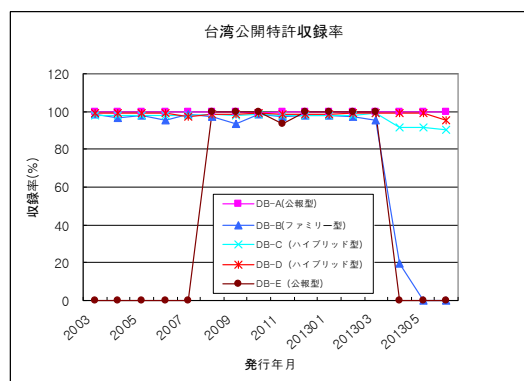


図4. 商用DB公開特許収録確認

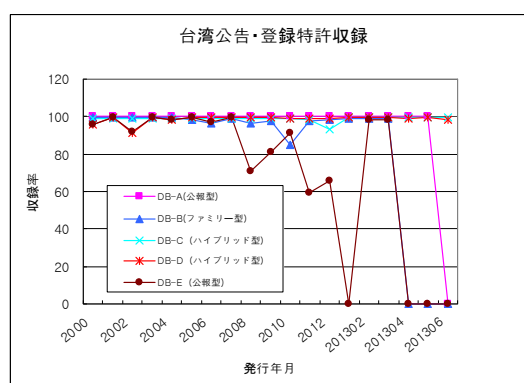


図5. 商用DB公告・登録特許収録確認

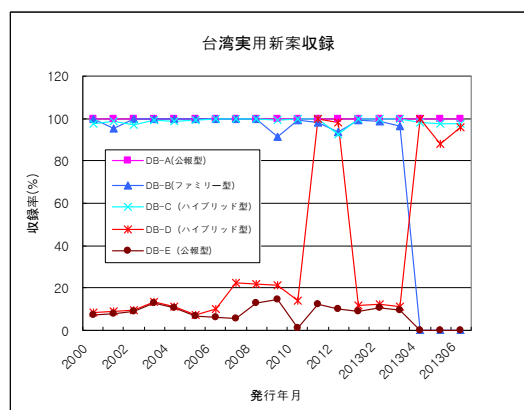


図6. 商用DB実案収録確認

### 3-2) TWPAT文字化け

発明の名称、要約における6表記(ナフタレン等)の正常表記、文字化けの結果を図7に示す。2009年頃から挙動が変化しているが、継続して文字化けがあり、文字化けも一種の異表記として考え、検索する必要がある。

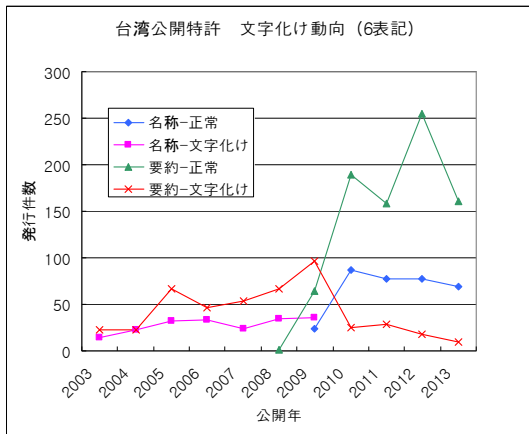


図 7. TWPAT 公開特許文字化け

### 3-3) 中国語繁体字、簡体字混用

中国語版 TWPAT では、中国本土で用いられる簡体字で検索しても、台湾地域で用いられる繁体字に自動変換され検索されているようで、DB により字体の扱いが異なる。(表.1)

表 1. 繁体字検索と簡体字検索

	手機 (繁体字)	手机 (簡体字)	備考
TWPAT	100%	*100%	自動変換
DB-A(公報型)	100%	0%	
DB-C (ハイブリッド型)	100%	*29%	中国ファミリー分
DB-D ハイブリッド型)	100%	*29%	中国ファミリー分

### 3-4) ハイブリッド DB 原語収録

ハイブリッド DB における台湾公開特許の原文クレーム収録率を図 8 に示す。

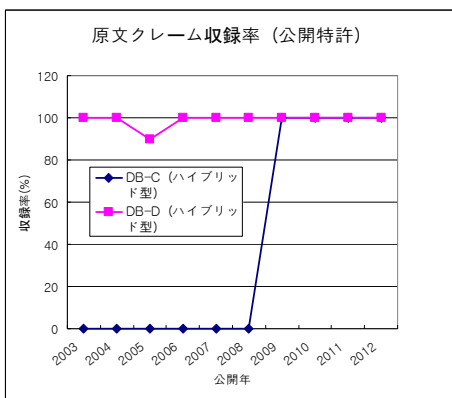


図 8. ハイブリッド DB 原文クレーム収録

3-5) 台湾特許中国語クレーム検索  
中国語検索可能な DB で、4 種のキーワードを用い検索し、TWPAT を基準に「再現性」をまとめ、図 9 に示す。

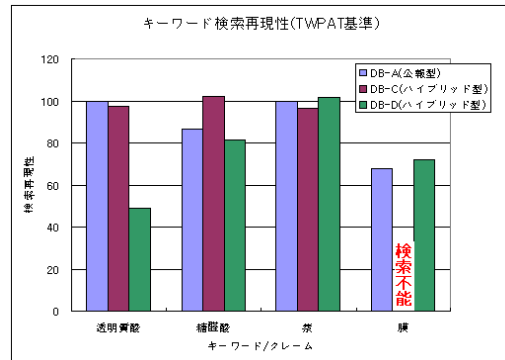


図 9. クレーム検索再現率(2013.7.18 確認)

## 4. 考察

台湾特許の収録状況を見る限り、特許、実案の両方を問題なく収録できている DB は限られており、どの DB で調査するかにより、結果が大きく異なる可能性ある。しかし、数年前とは異なり、英語ベースの DB でも要約までは漏れ無く収録されている DB もあり、隔世の感がある。

また WPI を用いた検証から、台湾公開特許のうち約 30%は台湾のみにしか出願されておらず(図 10)、英語要約までの検索では限界がある。

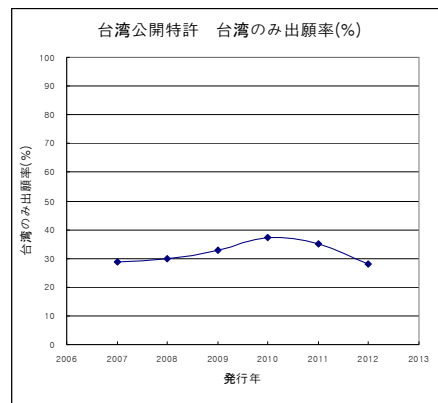


図 10. 台湾公開特許 台湾のみ出願率

ただし、中国語も併用して検索する場合は、注意が必要である。特に台湾では、TWPAT 上で文字化けがあり(図 3)、現在も DB 上で文字化け情報を含む特許が新規分についても存在し、総数として増加している。調査の際は、異表記として、文字化けがないかを確認し、もし、ある場合は、それも異表記の一つに加え検索する必要がある。

なお、TWPAT 上での文字化けは、現地代理人を通じて、台湾經濟部智慧財産局に確認した結果、中国語 OS では(修正パッチ)フォントをインストールすることにより、正常に見えるとのことであるが、日本の OS には非対応とのことであり、かつ文字化けしたものが異表記で存在する事実が変わりはない。

字体については、TWPAT では中国本土で使われる簡体字で入力しても、自動的に台湾で使われる繁体字に変換するが、他の DB では、ヒットしないので注意が必要である。(表 1)

また、台湾と、中国ではキーワードの表記も異なる場合があり、mobile phone の英語タイトルで検索した場合、表 2 のように同一ファミリーの特許でも、台湾、中国のキーワード表記が字体(繁体字、簡体字)の違いだけにとどまらず、表記そのものが違ってくる場合があるので注意が必要と考える。

表 2. 同一ファミリー mobile phone 表記例

中国簡体字表記	台湾繁体字表記	備考
手机	手機	字体変化
	行動電話	表記変化
移动通讯装置	行動通訊裝置	表記変化
手持通信	手持通訊	表記変化

また、台湾公開特許のクレーム検索で

の「再現性」を確認したが、中国語キーワードを用いて検索する場合は、その中国語部分の収録(図 8)にも気をつける必要があり、今回は収録そのものが問題ない範囲で、共通に比較できる中で検証した。(図 9)

図 9 の結果を見ると、単に収録だけではなく、検索ロジックや、DB の限界についても、考慮する必要があると考える。具体的にはキーワードによって、その「再現率」が大きく異なっており、これは部分一致という簡単な検索ロジックだけで、検索されるわけではなく、インデキシングされたものを検索する場合もあることを示していると考えられる。例えば、DB-A でのクレーム検索で、“糖醛酸”検索で漏れた特許を解析すると、“一聚半乳糖醛酸酶”や、“聚半乳糖醛酸”で検索するとヒットすることが判明している。

また、中国語では 1 文字でナフタレンを意味する“萘”の文字があるように、1 文字での検索が必要な場合があるが、その際も「再現性」が不十分または、検索不能な場合もあり、注意が必要である。

ただし、中国語版 TWPAT では、コマンド検索でも、検索フィールドごとに個別にキーワード設定する必要があり、図 11 のように検索式が複雑になり易く、その場合、タイムアウトエラーが発生し、検索できないことも多いので、注意が必要である。またダウンロードも 1 回 300 件までの制限があり注意が必要である。

ID=19930719:20130719 and (IC="G02B-005/20\*" or IC="G02F-001/133\*" or ((color filter) or (colour filter) or 彩色濾光 or 彩色分色 or 彩色過濾 or 彩色濾波 or 色彩濾 or 色片 or 濾光片 or 濾色板)@TI or (彩色濾光 or 彩色分色 or 彩色過濾 or 彩色濾波 or 色彩濾 or 色片 or 濾光片 or 濾色板)@AB or (彩色濾光 or 彩色分色 or 彩色過濾 or 彩色濾波 or 色彩濾 or 色片 or 濾光片 or 濾色板)@CL or (彩色濾光 or 彩色分色 or 彩色過濾 or 彩色濾波 or 色彩濾 or 色片 or 濾光片 or 濾色板)@DE)

↓ 検索後、タイムアウトエラー発生

## 504 Gateway Timeout

図 11. TWPAT での検索エラー

これらの結果から、台湾特許を網羅的に、調べるのは非常に困難が伴うが、特許出願件数自体が 5 万件/年程度であり、力技的に、広く取るような手法も可能と考える。なお、登録特許の年金支払状況も、TWPAT より、書誌情報と一緒にダウンロード可能で、登録特許の生死判別も可能である。

なお、TWPAT は 2013 年 7 月に大幅なリニューアルがあり、調査分析に役立つ機能も登場した。例としては、特許分析機能がある。これは、無料ユーザー登録をした上で、ログインした際にのみ使用できる機能で、Max3000 件のデータが解析でき、各項目上位 30 位までが表示可能である。解析できる項目は、筆頭出願人、IPC、発行年、発明者である。図 12 に、筆頭出願人について分析した結果を示す。

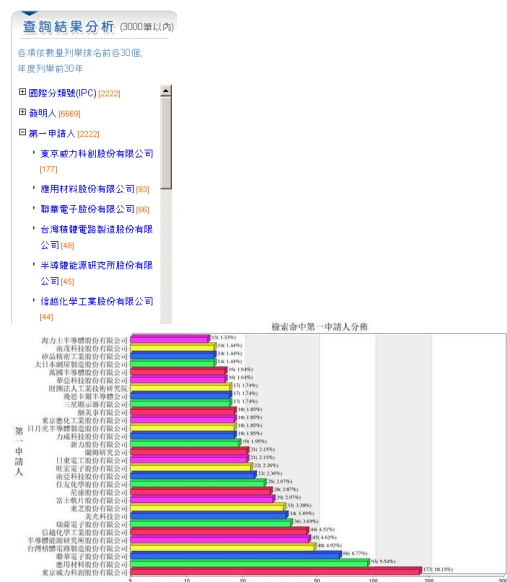


図 12. 筆頭出願人分析例 (中国語版 TWPAT ログイン)

このように、中国語出願人名で統計解析すると、ローカル出願人についても、表記割れ少なく解析でき、リスト表示及びグラフ表示可能で、非常に有用である

と考える。

その他にも、TWPAT には、中英辞書、技術辞書、出願人辞書、同義語辞書など、調査に役立つツールもあり、10 月の本発表の際に時間のある限り紹介したいと考えている。

なお、本研究は、アジア特許情報研究会 2013 年度ワーキングによるもので、記載されたメンバー以外にも多くの方々にアドバイスいただいたこと深く感謝します。

## 5. 参考文献

- [1] 吉居未来 他 : 早期公開・登録および遅延公開をめぐる東アジア各国の諸問題 (第 8 回情報プロフェッショナルシンポジウム、2011/10)
- [2] 伊藤徹男 : アジア・新興国知財情報収集の現状 (日本知財学会 第 10 回年次学術研究発表会、2012/12)
- [3] 田畑文也 他 : 英語・原語によるハイブリッド検索 (第 8 回情報プロフェッショナルシンポジウム、2011/10)
- [4] 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 (第 6 回情報プロフェッショナルシンポジウム、2009/11)