

中韓文献翻訳・検索システムの検証

2015年12月11日

アジア特許情報研究会

○佐野浩太郎(株式会社IHI)

前田佳治(東亜合成株式会社)

石田政司(株式会社神戸製鋼)

伊藤徹男(アジア特許情報研究会)

1. はじめに
2. 目的
3. JPO中韓システムの機能
 - 3-1 データの収録
 - 3-2 公報テキスト検索
 - 3-3 公報番号索引
4. 検証結果
 - 4-1 各事例紹介
 - 4-2 その他の翻訳不良事例
 - 4-3 網羅的検索のための検索式例
5. まとめ
 - 5-1 JPO中韓システムを利用した査読・検索
 - 5-2 システム改善に対する要望

1. 背景

中国



大量の特許情報

特許・実用新案の出願件数：70.4万件(2013年) → 92.8万件(2014年)

韓国



出願から登録までの審査が促進

審査期間(最初の通知まで)：14.8ヶ月(2012年) → 13.2ヶ月(2013年)

特許調査の向上

- ・ 収集した情報を効率的に、また素早く、内容を把握
- ・ 英語情報のみでは情報の漏れが発生するため、中国語やハンゲルによる検索で補完

中韓文献翻訳・検索システム (http://www.ckgs.jpo.go.jp/)



メニュー

サービスメニュー

[公報テキスト検索](#)

中国・韓国の翻訳された公報を、公報に表示されているIPCやキーワードで検索することができます。

[公報番号索引照会](#)

公報番号で中国・韓国の公報を照会することができます。

[検索可能範囲参照](#) ?

[中国特許和文抄録テキスト検索\(J-PlatPat\)](#)

J-PlatPatの特許・実用新案テキスト検索により、中国公開特許公報の和文抄録を検索することができます。

[中国特許和文抄録番号照会\(J-PlatPat\)](#)

J-PlatPatの外国公報DBにより、公報番号で中国公開特許公報の和文抄録を照会することができます。

日本語で文献を検索し、表示することが可能

調査担当者だけでなく研究者や技術者などが容易にアクセスでき、内容のスクリーニングが可能

2. 目的

(1) JPO中韓システムの利用方法の概略を紹介

(2) 開発者などが上記JPO中韓システムを利用して中国・韓国特許文献を査読する際の活用法を、多分野にわたる検証結果から提案

- ・異表記・誤表記の分野別特異性
- ・日本語で検索及び査読する際の現状及び留意点
- ・現状の機械翻訳精度におけるベストな日本語検索式

3. JPO中韓システムの機能

3-1 データの収録

中国特許・実用新案の収録状況 (2003年～2013年)

CN公開特許		2015/04/03			
発行年	CNIPR 書誌	中韓文献 翻訳・検索システム			
		書誌	要約	請求項	全文
2003	77472	100.0%	100.0%	100.0%	100.0%
2004	93944	100.0%	100.0%	100.0%	100.0%
2005	155446	99.9%	99.9%	99.9%	99.9%
2006	172424	100.0%	100.0%	99.9%	99.9%
2007	208345	99.3%	99.3%	99.3%	99.3%
2008	241182	99.4%	99.4%	99.4%	99.4%
2009	281006	99.9%	99.9%	99.9%	99.9%
2010	315836	99.5%	99.5%	99.5%	99.5%
2011	368434	98.3%	98.3%	98.3%	98.3%
2012	543296	98.5%	98.5%	98.5%	98.5%
2013	632585	99.3%	99.3%	99.3%	99.3%
2014	777336	95.8%	95.8%	95.8%	95.8%

各年の収録状況は98%～99%の収録率

韓国特許・実用新案の収録状況 (2003年～2013年)

KR公開特許		2015/04/03			
発行年	KIPRIS 書誌	中韓文献 翻訳・検索システム			
		書誌	要約	請求項	全文
2003	97909	100.0%	100.0%	100.0%	100.0%
2004	110582	100.0%	100.0%	100.0%	100.0%
2005	123492	99.2%	99.1%	99.1%	99.2%
2006	135946	96.2%	96.2%	96.2%	96.2%
2007	122581	99.9%	99.9%	99.9%	99.9%
2008	114829	100.0%	99.9%	99.9%	100.0%
2009	133121	100.0%	99.9%	99.9%	100.0%
2010	138946	95.7%	99.9%	100.0%	100.0%
2011	140131	99.9%	99.9%	99.9%	99.9%
2012	139463	99.4%	99.3%	99.3%	99.4%
2013	141194	99.3%	99.2%	99.2%	99.3%
2014	148110	98.8%	98.7%	98.7%	98.8%

各年の収録状況は95%以上の収録率

IPC別の収録（中国）

2015/08/17現在									
CN	A61K			B32B			C08F		
公開年	CNIPR	JPO中韓	収録率	CNIPR	JPO中韓	収録率	CNIPR	JPO中韓	収録率
2003	7473	7473	100.00%	654	654	100.00%	797	797	100.00%
2004	8127	8125	100.00%	687	687	100.00%	827	827	100.00%
2005	15142	15126	99.90%	1177	1176	99.90%	1337	1334	99.80%
2006	15743	15731	99.90%	1280	1279	99.90%	1290	1290	100.00%
2007	18986	18913	99.60%	1613	1606	99.60%	1653	1632	98.70%
2008	18397	18291	99.40%	1927	1918	99.50%	1784	1770	99.20%
2009	19703	19686	99.90%	2432	2431	100.00%	2095	2095	100.00%
2010	20153	19571	97.10%	2394	2391	99.90%	2266	2209	97.50%
2011	21445	19390	90.40%	3031	3010	99.30%	2555	2276	89.10%
2012	28240	25984	92.00%	5590	5535	99.00%	4450	4064	91.30%
2013	33379	33006	98.90%	6568	6517	99.20%	4608	4564	99.00%
2014	44287	42356	95.60%	7865	7536	95.80%	5462	5154	94.40%

全体的には98%以上の収録率であるが、一部低い部分あり

3-2 公報テキスト検索

公報テキスト検索画面上部

公報テキスト検索

発行国 中国 韓国 発行種別 特許 実用 文献種別 公開 登録 公知日 [] ~ [] 出願日 [] ~ []

フリー検索条件 ?					NOT検索条件 ?				
項番	項目接続 ?	検索項目 ?	検索キーワード ?	項目内接続 ?	項番	項目接続 ?	検索項目 ?	検索キーワード ?	項目内接続 ?
<input checked="" type="checkbox"/>	1	-	発明の名称	ディーゼルエンジン ディーゼル機関	OR	<input checked="" type="checkbox"/>	-	要約+請求の範囲	OR
<input checked="" type="checkbox"/>	2	AND	公開日	20030101:20131231	OR	<input checked="" type="checkbox"/>	AND	公報全文(書誌を除く)	OR
<input checked="" type="checkbox"/>	3	AND	国際特許分類		OR	<input checked="" type="checkbox"/>	AND	国際特許分類	OR
<input checked="" type="checkbox"/>	4	AND	出願人/権利者		OR	<input checked="" type="checkbox"/>	AND	出願人/権利者	OR

検索条件クリア 行追加 最終行削除

- ・ キーワード：1項目100文字以内、1用語の長さは20文字以内
- ・ 国際特許分類：サブクラス以下まで入力が必要
- ・ 出願人：原語(中国語又はハングル)による検索

公報テキスト検索画面下部

項目間接続 ?		検索条件		
<input checked="" type="checkbox"/>	AND	近傍検索 ?	未設定	設定
<input checked="" type="checkbox"/>	AND	拡張NOT検索 ?	未設定	設定
<input type="checkbox"/>	AND	検索結果集台式 ?	1	

近傍検索条件設定

検索項目

検索キーワード1

文字間隔 語順指定 検索キーワード2

文字間隔 語順指定 検索キーワード3

優先順序指定 前半優先 後半優先

*検索キーワードには複数のキーワードを半角プラスまたはスペース区切りで入力することが可能です

異表記展開オプション

- カタカナ ? ひらがな ? アルファベット ?

検索結果

項番	ヒット件数		発行種別	文献種別	公知日		出願日		フリー検索条件	NOT検索条件
	中国	韓国			From	To	From	To		
1	2099	1440	特 実 公 登	-	-	-	-	発明の名称：ディーゼルエンジン ディーゼル機	-	

近傍検索などの機能あり

3-3 公報番号索引

番号入力画面

公報番号

CNA-101981105
 CNA-102432913
 CNA-1867606
 CNA-1471558
 CNA-101945935
 CNA-101175789

入力番号フォーマット

中国	特許公開	CNA-nnnnnnnn
		CNA-nnnnnnnnnn
	特許登録	CNB-nnnnnnnn
		CNB-1nnnnnnnnn
		CNC-nnnnnnnn CNC-nnnnnnnnnn
	実用登録	CNU-2nnnnnnnnn
CNY-nnnnnnnn		
CNY-nnnnnnnnnn		
韓国	特許公開	KRA-yyyynnnnnnnn
	特許登録	KRB1-10nnnnnnnn
	実用公開	KRU-yyyynnnnnnnn

検索結果表示画面

2015/8/27

公報番号	有無	発行国	公知日	出願日	発明の名称	出願人名
CNA-102432913	有	中国	20120502	20100818	保護した膨張性ポリテレフタル酸アルキレングリコールエステルを受けること	阿玛赛尔企业有限公司
CNA-101945935	有	中国	20090702	20071220	難燃化の耐衝撃性変性のポリアルキレン基がフタル酸エステル/ポリカーボネート組成物に対すること	拜尔材料科学股份公司
CNA-101175789	有	中国	20061116	20050512	酸化チタンゾルおよびそれを用いたポリテレフタル酸アルキレイルエステルの製造方法	富士钛工业公司
CNA-1867606	有	中国	20050421	20031010	生産がベンゼンジカルボン酸アルキル基2アルコールエステルに対しての方法に集め、ベンゼンジカルボン酸アルキル基2アルコールエステル成形体に対する方法及びポリテレフタル酸アルキル基2アルコールエステル成形体に集めることを生産すること	旭化成化学株式会社
CNA-1471558	有	中国	20020502	20001025	亜リン酸エステル安定したポリでベンゼンジカルボン酸アルカンジオールエステル組成物に対すること	拜尔公司
CNA-10981105	無					

- ・ 公報番号 (完全一致) を一度に1000件まで入力可
- ・ 未収録データもあるが随時追加

4. 検証結果

4-1 各事例の紹介

【 検証概要 】

- ・ 検索性数:「発明の名称」中の用語
- ・ 期間:2003 ~ 2013年公開特許
- ・ 対象:中国・韓国語翻訳の日本語表記ゆれ
- ・ 検索 : フィールド内は or 演算、フィールド間には and 演算

カーボンナノチューブの事例 : 中国

CNIPR	TI	JPO中韓システム	TI
碳納米管	2918	カーボンナノチューブ	2694
		炭素ナノチューブ	118
		炭ナノチューブ	57

「炭ナノチューブ」といった予測しがたい訳語あり

要約：本發明披露一種處理**碳納米管**的方法，包括如下的步驟：將**碳納米管**分散在分散介質中，製各分散體；將所述分散體與吸附劑混合，以使所述分散體中特定類型或特性的**碳納米管**吸附在所述吸附劑上，其中所述吸附劑經化學/生物改性劑改性而對不同類型或特性的**碳納米管**具有不同的選擇吸附性；以及將吸附劑與分散體分離，由此使得吸附在所述吸附劑上的特定類型或特性的**碳納米管**與富集在分散體中其它類型或特性的**碳納米管**分離。本發明還披露了經該方法處理的**碳納米管**以及**碳納米管**器件。

(54) 【發明名稱】 碳納米管處理方法、碳納米管及碳納米管器件

(57) 【要約】

本發明是處理**碳納米管**的方法，包括如下步驟：將**碳納米管**分散在分散介質中，製各分散體；將所述分散體與吸附劑混合，以使所述分散體中特定類型或特性的**碳納米管**吸附在所述吸附劑上，其中所述吸附劑經化學/生物改性劑改性而對不同類型或特性的**碳納米管**具有不同的選擇吸附性；以及將吸附劑與分散體分離，由此使得吸附在所述吸附劑上的特定類型或特性的**碳納米管**與富集在分散體中其它類型或特性的**碳納米管**分離。本發明還披露了經該方法處理的**碳納米管**以及**碳納米管**器件。

原語は同一の中国語
表記なのに何故？

同一案件中の訳文異表記 ⇒ 査読時に留意すべき

クラッシュボックスの事例：中国

CNIPR	TI	JPO中韓システム	TI
吸能盒	19	クラッシュボックス	1
		含ませは容器	7
		含ませ・・・ 4種	4
		ことの容器を吸入すること	4
		ことの容器・・吸入すること	3

※ クラッシュボックス：自動車のフレーム先端に配置される衝突吸収部材

- 19件中、正しい日本語に翻訳されたのはわずかに1件
- その他の訳語からはクラッシュボックスの想定が困難

シェールガスの事例：中国

shale gas		2003-2013公開特許、発明の名称	
CNIPR	TI	JPO中韓システム	TI
頁岩気(shale gas)	64	シェールガス	59
		シェール気	1
		頁岩ガス	1
		頁岩気	1
		炭質けつ炭気	2
頁岩含気	1	シェール気	1
頁岩瓦斯	2	炭質けつ炭ガス	2
頁岩油(shale oil)	54	頁岩油	46
頁岩油		シェール油	1
油頁岩油		炭質けつ炭油	1
頁岩油藏		シェール埋蔵石油	1
油頁岩油气		炭質けつ炭天然ガス	1
頁岩油气		シェール天然ガス	2
頁岩油气		頁岩天然ガス	2
油頁岩	206	炭質けつ炭	190
		オイルシェール	2
		油頁岩	3
油母頁岩	22	炭質けつ炭	19
		未収録	3

赤字は、原語のまま

全く訳されずに日本語漢字を充てられている公報あり

ターボチャージャの事例：韓国

KIPRIS	TI	JPO 中韓システム	TI
터보차저	283	ターボチャージャー	280
		翻訳不良	3
터보 차저	61	ターボチャージャー	61
터보차저	93	ターボチャージャー	93
터보 차저	28	ターボチャージャー	28
터보차저	21	ターボチャージャー	21
터보 차저	12	ターボチャージャー	4
		ターボチャゾ	8
터보차저	7	ターボチャージャー	7

- ・ 異表記のいずれもほぼ「ターボチャージャー」と翻訳
- ・ 査読時に判別し難い検索用語が少数あり

カーボンナノチューブの事例：韓国

KIPRIS	TI	JPO中韓システム	TI
탄소나노튜브	1471	炭素ナノチューブ	1471
"탄소 나노튜브"	324	炭素ナノチューブ	324
"탄소나노 튜브"	3	炭素ナノチューブ	3
"탄소 나노 튜브"	62	炭素ナノチューブ	61
		탄소나노튜브	1
"탄소 나노관"	1	炭素ナノ管	1
카본나노튜브	187	カーボンナノチューブ	187
"카본 나노튜브"	79	カーボンナノチューブ	79
"카본나노 튜브"	0		
"카본 나노 튜브"	32	カーボンナノチューブ	32

検索用語としても、査読用語としても、ほぼ問題なし

4-2 その他翻訳不良事例

化学分野：数字、括弧を含む化合物で誤訳

- ・ 聚(对苯二甲酸乙二醇酯) ⇒ ポリ(正しいポリエチレンーテレフタレート)
(poly(ethylene terephthalate))
- ・ 四羧二酰亚胺 ⇒ 4のカルボキシルジイミド
(tetracarboxylic acid diimide)
- ・ 氮杂二环(3.1.0)己烷 ⇒ アザビシクロ(3. 1. 0)もうジン
(azabicyclo (3.1.0) hexane)

語順の反転

問題なし

- ・ 高密度積層印制板 ⇒ 高密度の積層プリント基板
- ・ 樹脂組合物及層積体 ⇒ 樹脂組成物及び積層体

問題あり

- ・ 皮膚护理设备 ⇒ スキンケア装置
- ・ 护理皮肤的美容方法 ⇒ 介護皮膚の美容方法

長いキーワードの場合 (1)

公報番号	CNIPR		JPO中韓
CN104563755A	铝型材	铝(n) + 型材(n)	アルミニウムプロファイル
CN104482390A	铝型材	铝(n) + 型材(n)	アルミニウムプロファイル
CN102701064A	铝型材吊篮	铝(n) + 型材(n) + 吊篮(n)	アルミニウムプロファイルゴンドラ
CN103573112A	铝型材门扇	铝(n) + 型材(n) + 门扇(n)	アルミニウムプロファイルドアリーフ
CN103591435A	铝型材盖板	铝(n) + 型材(n) + 盖板(n)	アルミニウムプロファイルカバー
CN103527032A	铝型材扣板	铝(n) + 型材(n) + 扣板(n)	アルミニウムプロファイルピンチプレート
CN102699086A	铝型材挤压机	铝(n) + 型材(n) + 挤压机(n)	アルミニウムプロファイル押出機
CN103537502A	铝型材挤压机工艺	铝(n) + 型材(n) + 工艺(n)	アルミニウムプロファイル押出工程
CN103537508A	铝型材挤压机装置	铝(n) + 型材(n) + 挤压机装置 (n)	アルミニウムプロファイルプレス装置
CN203330168U	铝型材挤压机模具 碱洗系统	铝(n) + 型材(n) + 挤压机模具 (n) + 碱洗(n) + 系统 (n)	アルミニウムプロファイル押出ダイスアルカリ洗浄システム
CN203254003U	铝型材挤压机生产 线牵引机	铝(n) + 型材(n) + 挤压机生产线 (n) + 牵引机 (n)	アルミニウムプロファイル押圧生産ライン牽引機

- ・ 名詞が並んだ複合キーワードについては、比較的正確に翻訳
- ・ 上記キーワードでは“型材(プロファイル)”の部分のみが誤訳

長いキーワードの場合 (2)

公報番号	CNIPR		JPO中韓
CN103573115A	铝型材固定横框	铝(n) + 型材(n) + 固定 (v) + 横框(n)	アルミニウムプロファイル固定した横枠
CN103272955A	铝型材多层型 角持刀装置	铝(n) + 型材(n) + 多层型(n) + 角持刀(n) + 装置(n)	アルミニウムプロファイル多層タイプ角が刃装 置を持つこと
CN203395563U	挤压铝型材边封 侧板	挤压(v) + 铝型材(n) + 边封(v) + 侧板(n)	アルミニウムプロファイル辺に側板を封じること を押し出すこと
CN201676911U	铝型材挤压轴一 体化结构	铝(n) + 型材(n) + 挤压模具 (n) + 碱洗(n) + 系统 (n)	アルミニウムプロファイル押し出し軸一体化構造
CN103978371A	铝型材挤压切割 一体机	铝(n) + 型材(n) + 挤压 (v) + 切割(v) + 一体机 (n)	アルミニウムプロファイル押圧ダイシング一体 型パソコン

複合キーワードの一部に動詞等が入り込んでくると誤訳が生じやすい
(1つのキーワードではなく文として処理)

4-3 網羅的検索のための検索式例

JPO中韓システムの日本語表記を網羅的に抽出し、原語検索の検索結果とを比較

CNIPR	碳纳米管 or 纳米碳管 or 炭纳米管 or 纳米炭管 or 碳奈米管 or 奈米碳管 or 碳毫微管 or 碳纳米细管 or 碳基纳米管 or 纳米级碳管	3084件
中韓	カーボンナノチューブ / 炭素ナノチューブ / 炭ナノチューブ / ナノカーボンチューブ / カーボンナノ管 / ナノカーボン管 / 炭素ナノ管 / 炭ナノ管 / ナノ炭管 / 炭素基ナノチューブ / ナノ木炭管	3034件
KIPRIS	탄소나노튜브 + “탄소 나노튜브” + “탄소나노 튜브” + “탄소 나노 튜브” + “탄소 나노관” + 카본나노튜브 + “카본 나노튜브” + “카본나노 튜브” + “카본 나노 튜브” + “carbon nanotube”	1733件
中韓	カーボンナノチューブ カーボンナノ管 ナノカーボンチューブ ナノカーボン管 炭素ナノチューブ 炭素ナノ管 炭ナノチューブ 炭ナノ管 ナノ炭素チューブ ナノ炭素管 ナノ炭チューブ ナノ炭管	1677件

検索時、各用語に充てられる日本語訳を確認することは不可能ではない
(原語検索と同等な検索可能)

キーワードの抽出に相当な手間とスキルを要するため実務的ではない

5. まとめ

5-1 JPO中韓システムを利用した査読・検索

本システムで有効活用できる態様

(1) 査読

- ・ 技術内容は充分理解している研究者などが査読を行う際、効率が格段に向上
- ・ 翻訳不良部分については翻訳ツールを利用することでカバー可能
- ・ 理解しがたい請求項などについてのみ、適宜、人の手による翻訳

(2) 検索

- ・ 網羅的な検索をすることが不可能ではないが煩雑
- ・ 原語データベースで網羅的検索を実施し、抽出した公報をJPO中韓システムの番号検索で抽出し、内容を査読するのが最良

5-2 システム改善に対する要望

本システムのみで網羅的検索が可能となるために要するシステム改善

(1) 翻訳不良、異表記及び表記ゆれ

- ・ 翻訳辞書登録の充実
- ・ 日本語と原語で同時検索できる(ハイブリッド検索)システム機能

(2) 同一の原語が複数の日本語に翻訳されている

同一の用語も文章の前後の位置、文法的な観点からの翻訳(品詞の種類も加味した翻訳)も考慮

おわりに

本報告は、2015年度の「アジア特許情報研究会」のワーキングの一環として報告するものであり、研究会の皆様には情報の提供及び数々のアドバイスをいただきました。ここに改めてお礼申し上げます。