

2011年に日本でもリリースされた CPA GLOBAL DISCOVER。その特徴機能をシリーズでご紹介しています。今回は特許検索ツールに不慣れな方や検索条件が上手く設定できないような時にサポートしてくれる「Navigators」機能をご紹介しました。

今回は、「英語のキーワードとオリジナル言語のキーワード」を併用して検索することができる「Hybrid 検索」機能をご紹介させていただきます。なお、「Hybrid 検索」というのは正式な機能名ではありません。以前、社外の特許検索に詳しい方々と当該機能を検証した際に、エキスパートの方々がそのように呼称していらっしゃいました。それ以降、「Hybrid 検索」という呼称を利用させて頂いております。

Keywords	Publication Number	Similarity
Type or paste keywords below		
(飞机 OR 火车 OR 列车) AND 司机		

Keywords	Publication Number	Similarity
Type or paste keywords below		
ЦЕНТРОБЕЖНЫЙ		

上記サンプルは DISCOVER のマニュアルにも記載されている「中国語」と「ロシア語」を使った「non-latin 検索」の例です。DISCOVER ではラテン文字以外の言語でも検索をすることができます。この「non-latin 検索」と通常の英語による検索を組み合わせたのが「Hybrid 検索」です。

前回、ご案内させて頂いたように DISCOVER は 91 カ国の特許公報データを収録し、そのうち、主要国を中心にした 23 カ国(※Japan* と Korea*は抄録のみ)についてはフルテキストのデータを収録していますが、非英語圏のデータについてはオリジナル言語のテキストデータに加え、機械翻訳された英語フルテキストデータを収録しています。これら、ラテン文字以外のオリジナル・テキストデータを収録しているもの(例えば、中国語、ロシア語、スペイン語など)に対しては「Hybrid 検索」(英語 & オリジナル言語での検索)が可能となっています。

最近では中国特許公報の英語テキストデータ(機械翻訳されたものが大半のようです)が流通し、多くのツールでこれらのデータを利用できるようになりました。中国特許調査を行う際は先ず、英語テキストデータを対象に検索される方々も多いのではないのでしょうか。そして、よく議論になるのが機械翻訳された英語の精度です。機械翻訳されたテキストデータであるが故、確かに不安は残ります。そんな時に「英語と中国語」の両方で検索できれば少しでも検索の精度を向上させることができるのではないのでしょうか。今回は中国特許公報を対象にした「Hybrid 検索」の事例をご紹介致します。

今回、「英語のキーワードは分かるが、中国語のキーワードは分からない」という前提にします。よって、まずは判明している英語のキーワードで中国特許検索を行います。

検索式例 1: PC:("CN") and ("solar cell*" or "solar batter*" or "solar energy*")

検索式例 1 で検索実施後、前回ご紹介した「Navigator 機能」を利用して公開特許に絞りました。その結果、検索結果は「54,167 件」となりました。(2012/12/27 現在) 次に索出された特許公報の詳細を確認します。そして「Hybrid 検索」に利用すべき中国語のキーワードを探してみたいと思います。

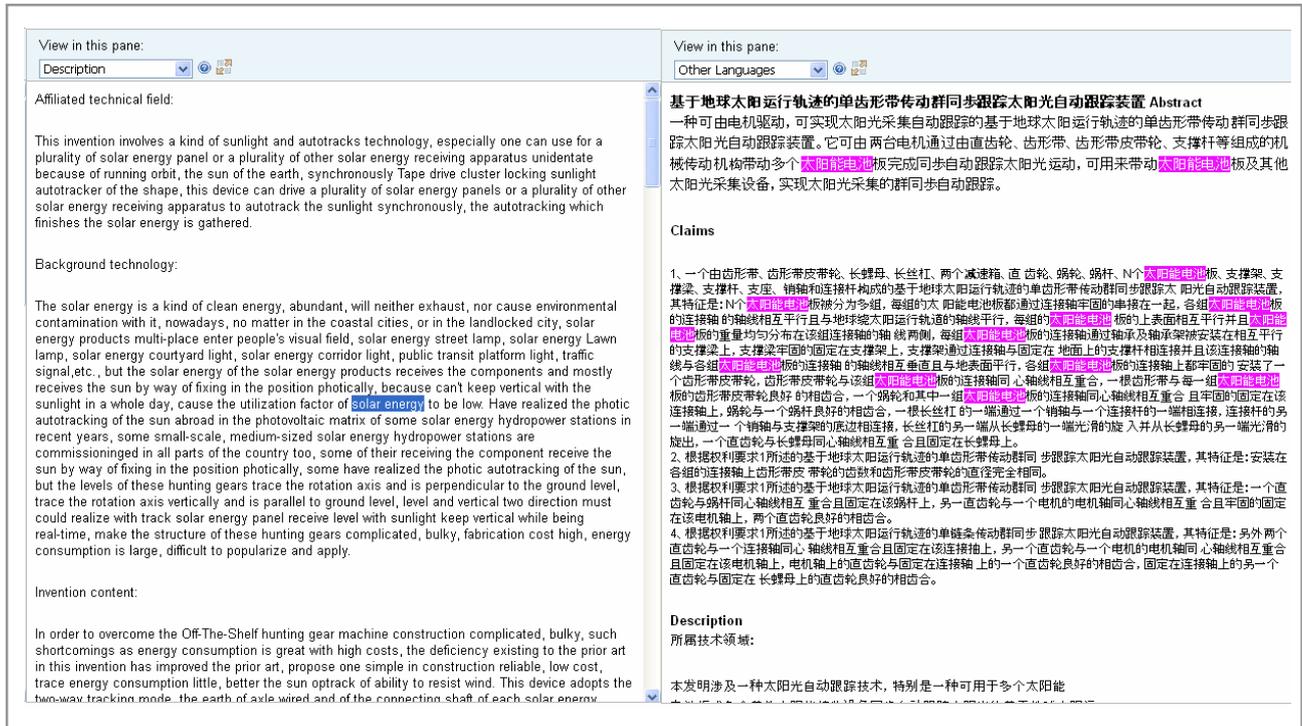
The screenshot shows a patent search result page with two columns. The left column is titled 'View in this pane: Description' and contains English text under 'Technical Field' and 'Background Art'. The right column is titled 'View in this pane: Other Languages' and contains Chinese text under 'Abstract' and 'Claims'. The Chinese text includes the title '对封装材料具有改进的粘结性的太阳能电池背板 Abstract' and several numbered claims. The Chinese text has several words highlighted in purple, including '太阳能电池'.

上記画面は DISCOVER の詳細データ表示画面の一部です。機械翻訳された英語のテキストとオリジナルの中国語テキストデータを並列表示しています。画面右のパープルでハイライトされているのは今回の検索式に含まれていた英語のキーワード「solar battery」で、画面左にブルーでハイライトさせている「太阳能电池」が「solar battery」を示す中国語だと思われます。他の公報も確認し、今回は「太阳能电池」と「太阳电池」という中国語キーワードを見つけることができました。

検索式例 2: PC:("CN") and ("solar cell*" or "solar batter*" or "solar energy*" or "太阳能电池" or "太阳电池")

予備検索で利用した検索式例 1 に「太阳能电池」と「太阳电池」という中国語のキーワードを追加して検索した結果は「54,203 件」でした。差は「36 件」となりました。この差分を検証してみます。

詳細データ表示画面でオリジナルの中国語を確認すると間違いなく「太阳能电池」というキーワードが存在します。次に機械翻訳された英語のテキストを確認したところ、これら 36 件の特許公報では「太阳能电池」は「solar panel」、「solar array」、「storage battery」、「sun battery」などと翻訳されていました。その為、オリジナルの中国特許公報中に「太阳能电池」というキーワードが含まれていても索出することができませんでした。



このように、英語キーワードに加え、オリジナル言語のキーワードを掛け合わせることで、翻訳されたデータを対象に検索する場合の弱点を補うことができます。

この手法は中国語だけではなく、ロシア語などを対象に利用することもできます。しかし、中国特許公報の漢字表記とは異なりロシア語などの場合は予備検索後、オリジナルのテキストデータを見て、対応するキーワードを見つけるのは難しいかもしれません。このような場合は、Web 翻訳ツールを利用して対応すると思われる技術用語を予め翻訳しておくべきかもしれません。

英語 日本語 中国語 言語を検出する

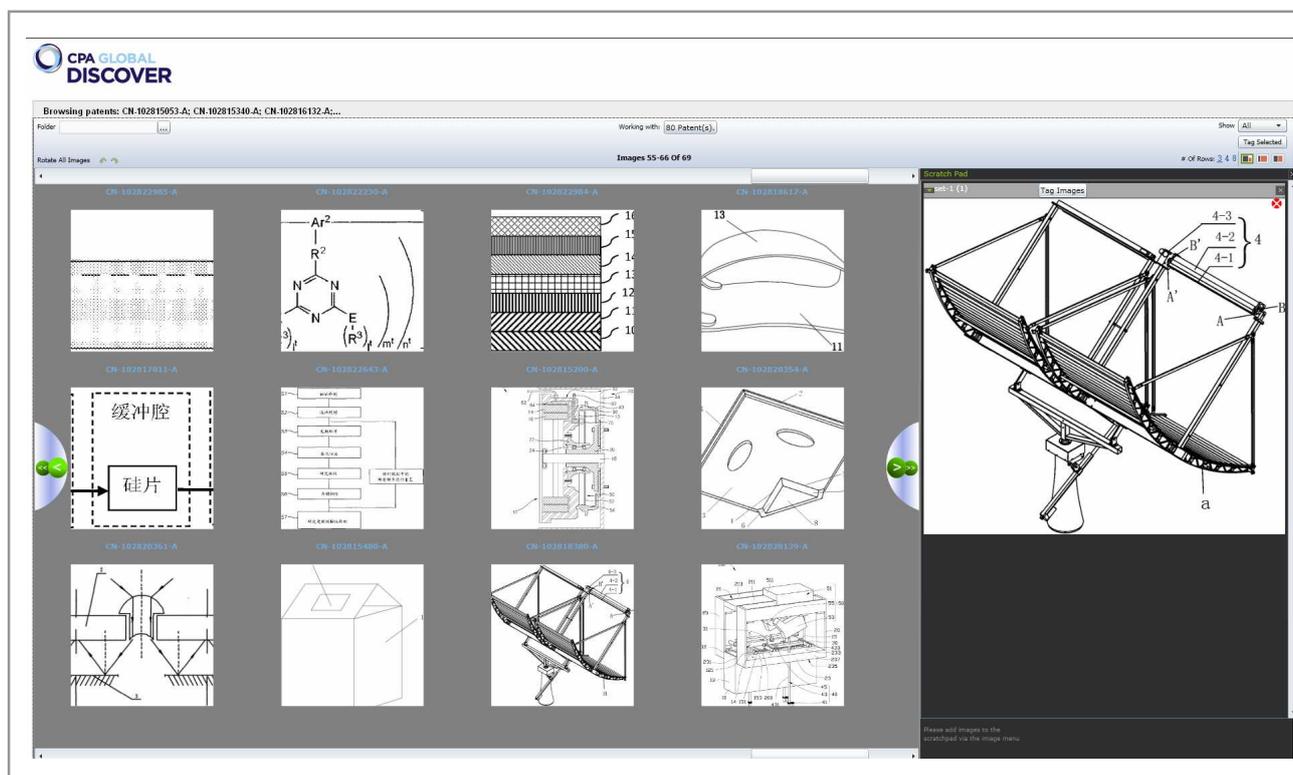
solar battery
solar cell
solar energy

日本語 スペイン語 ロシア語

солнечная батарея
солнечная батарея
солнечной энергии

前ご紹介した「Navigators 機能」、今回の「Hybrid 検索」以外にも DISCOVER には数多くの支援機能が
ございます。

次回は第 3 弾として、検索した結果を図面でスクリーニングする際に便利な「Image Browser 機能」をご紹
介させて頂く予定です。



上の画面は「Image Browser 機能」を使って複数の特許公報に含まれている全図面を連続表示させていると
ころです。スクリーニングの際、気になった図面は右側の「Scratch Pad」にコピーして残しておくこともできます。
「Image Browser 機能」の詳細は次回、ご案内いたします。

CPA GLOBAL DISCOVER に関するお問い合わせは下記担当者までお願い致します。

IP 総研 長谷川 雅則、白岡 剛

IP 総研代表メール ip-soken@ngb.co.jp