

# 日米の特許調査の高品質化・高効率化のための特許検索式の自動提案技術

IRD 国際特許事務所 所長／弁理士／有限会社アイ・アール・ディー **谷川 英和**

## PROFILE

1986年神戸大学工学部システム工学科卒業。同年、松下電器産業（株）に入社し、中央研究所等において、データベース管理システム等の研究開発に従事。1999年弁理士試験合格。2002年1月、IRD 国際特許事務所を開設。所長、弁理士。2003年～2007年3月京都大学 COE 研究員、2007年4月～京都大学非常勤講師、2009年4月～東京工業大学客員教授、2011年4月～大阪大学非常勤講師 博士（情報学）。弁理士会、情報処理学会各会員。

✉ htanigawa@ird-pat.com

☎ 06-6944-4530

IRD 国際特許事務所／有限会社アイ・アール・ディー **渡辺 俊規**

## PROFILE

2004年奈良工業高等専門学校専攻科電子情報工学専攻修了。2006年京都大学大学院情報学研究所修了。2004年ベンチャー企業設立を経て、2008年IRD 国際特許事務所入所。修士（情報学）。日本知財学会会員。

IRD 国際特許事務所／有限会社アイ・アール・ディー **増満 光**

## PROFILE

2008年関西大学大学院入学。2008年8月IPA：2008年度上期未踏IT人材発掘・育成事業（未踏コース）開始、特許請求項の参照関係可視化に関する研究開発に従事、2009年3月IPA：2008年度上期未踏IT人材発掘・育成事業（未踏コース）終了、2010年関西大学大学院卒業、同年2010年IRD 国際特許事務所入所。修士（情報学）。日本知財学会会員。

広島市立大学大学院情報科学研究科准教授 **難波 英嗣**

## PROFILE

1996年東京理科大学理工学部電気工学科卒業。2001年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了。同年、日本学術振興会特別研究員。2002年東京工業大学精密工学研究所助手。同年、広島市立大学情報科学部講師。2010年4月広島市立大学大学院情報科学研究科准教授、現在に至る。博士（情報科学）。言語処理学会、情報処理学会、人工知能学会、ACL、ACM各会員。

## 1 はじめに

我々は、2002年以降、発明の着想から権利化、権利行使に至る特許ライフサイクルにおける各作業について、工学的にアプローチを行う特許工学の研究を行ってきた<sup>1)</sup>。特許工学は、特許ライフサイクルにおける各種作業に対して、方法論を抽出し、ツール（以下、「特許工学ツール」という）と教育により、方法論の普及を図ることにより、各種作業の品質と効率の向上を目指すものである。

また、特許ライフサイクルの上流工程を構成する特許調査において、その品質や効率は、調査者のスキルに大きく依存する。例えば、特許調査の専門家は、適切な特許国際分類（IPC）、FI、Fターム、またはUSC（以下、これらを総称して「特許分類コード」という）を利用し、かつ、一のキーワードとその同義語、上位概念語、下位

概念語（以下、これらを「関連語」という）とを適切に利用し、漏れやノイズが少ない検索式を作成する。しかし、特許調査の専門家でも、対応する技術分野の範囲が広い場合、適切な関連語を抽出することは容易ではない。また、特許調査の非専門家にとって、適切に特許分類コードを決定することは困難である。さらに、日本人にとって、英語の関連語とUSCを含む検索式を作成することは、特に困難である。

そこで、我々は、ユーザが入力した日本語のキーワードに対応する関連語と特許分類コードを提案し、選択された関連語と特許分類コードから検索式を構築する特許調査支援システム（PatentSearchAssistant、以下、「PSA」という）を開発した。PSAにより、ユーザは、キーワードを入力するだけで、日本、米国を問わず、その関連語や特許分類コードを含む適切な検索式を容易に構築することができる。

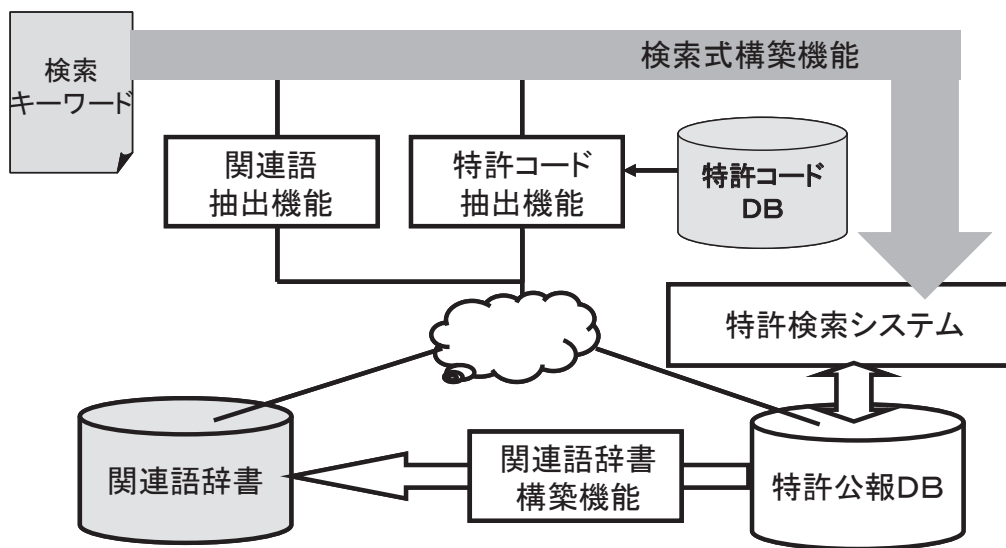


図1 PatentSearchAssistant の概要

## 2 関連業績

HYPAT-i、PATOLIS、NRI サイバーパテントなどの特許調査ツールの多くは、概念検索という機能を有する。概念検索は、検索対象の技術を説明する文を入力すると、入力した文が示す技術的概念に関連の深い特許公報を検索する機能であり、特許調査の非専門家が利用しやすい検索機能である。しかし、概念検索は、システム内部の処理が外からは見えず、調査方法の良し悪しの判断、調査結果の評価、および調査結果の再利用が困難である。そこで、一般的に、特許調査の専門家は、調査対象に関連したキーワードと、その関連語や特許分類コードで構成される検索式を用いる。しかし、上述したように、検索式の作成は、調査者の経験とスキルに大きく依存し、特許調査の専門家でも容易ではない。

## 3 特許調査支援システム (PSA) の概要

PSA は、キーワードを入力するだけで、特許公報に記載されている関連語と IPC などの特許分類コードを用いた検索式を提案する。この検索式は、調査専門家が作成

する検索式と同様に品質が高い。

PSA は、関連語辞書構築機能と、検索式構築機能の 2 つの機能で構成される。PSA の概要を図 1 に示す。

また、PSA における関連語の提案、および特許分類コードの提案の画面例を、図 2 に示す。

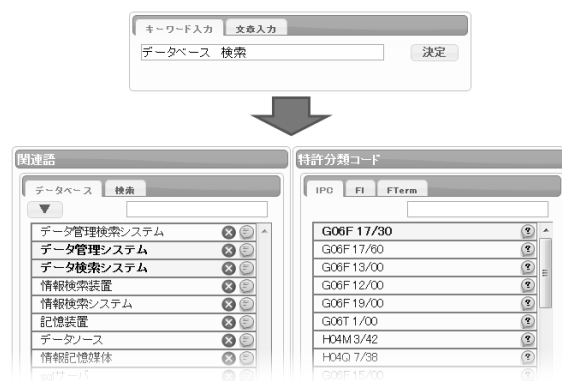


図2 関連語・特許分類コードの提案

図 2 は、ユーザが入力したキーワード「データベース」および「検索」に対し、それぞれの関連語、および、これらのキーワードとユーザが選択した関連語（以下、これらを「検索語」という）との関連性が高い特許分類コードを提示している画面である。

## 4 PSAの各種機能

### 4.1 関連語辞書構築機能

関連語辞書構築機能は、まず、日本の公開特許公報から手がかり句を用いて用語間の関係性を取得し、関連語辞書を構築する<sup>2)</sup>。本機能では、例えば、「などの」と「等の」という表現の前後に記載されている用語は上位概念語、下位概念語の関係にあることに着目する。例えば、「HDDなどの記録媒体」などである。この例では、「記録媒体」が「HDD」の上位概念語の関係にある。しかし、「などの」と「等の」だけでは、「パソコンなどのキーボード」のように上位概念語、下位概念語関係にない場合もある。そのため、「パソコンのキーボード」のように「などの」と「等の」を「の」に言い換えられる場合は、上位概念語、下位概念語の関係にないとした。また、「メモリやHDDなどの記録媒体」のように共通の上位概念語、もしくは下位概念語を持つ用語が列記されている場合も上位概念語、下位概念語の関係にないとした。そして、共通の上位概念語と下位概念語を持つ用語は、同義語であるとした。

関連語辞書構築機能は、次に、取得した日本語の関連語辞書用語に対して、日英機械翻訳機を用いて、日本語の関連語辞書に登録されている全ての用語を英語に翻訳し、取得した関連語のうち、実際に米国特許公報で使用されている用語を取得して、英語の関連語辞書を構築する。

### 4.2 検索式構築のための機能

PSAは、調査者の入力したキーワードから関連語を抽出する関連語抽出機能と、検索語から適切な特許分類コードを抽出する特許分類コード抽出機能で構成される。

関連語抽出機能の出力例を図3に示す。図3において、ユーザはキーワード「データベース」と「検索」を入力した場合の例である。



図3 関連語抽出機能の出力例

関連語抽出機能は、キーワードに対応した関連語を関連語辞書から抽出する。抽出する関連語は、入力されたキーワードと、関連語抽出機能が出力した関連語のうちからユーザが選択した関連語である。

次に、特許分類コード抽出機能の出力例を図4に示す。



図4：特許分類コード抽出機能の出力例

図4において、特許分類コード抽出機能は、関連語抽出機能で、ユーザが選択した検索語を用いて、検索に適したIPC、FI、Fタームを抽出する。抽出する特許分類コードは、すべての特許公報に対して、各特許公報に記載されている全用語数に対する検索語の出現割合を算出し、その特許公報が保持する各特許分類コードにその割合を重みとして与え、特許分類コードごとに重みの総和（以下、「コード重要度」という）を算出し、コード重要度が高いものとした。そして、最終的に提案する検索式は、取得した特許分類コードの組み合わせと検索語（ユーザ入力のキーワードと、ユーザ選択の関連語）から構成される。この時、ユーザは使用する特許分類コードを任意に選択し、検索式を組み替えることができる。

関連語抽出機能、および特許分類コード抽出機能は、米国の関連語と、特許分類コードをも出力する。

米国版の出力例を図5に示す。ユーザは、「database」と「search」を入力したとする。



図5 米国版の出力例

図5において、関連語抽出機能、および特許分類コード抽出機能は、日本版と同様にして関連語、およびコードを出力する。なお、米国版 PSA は、USC を提案する。

### 4.3 特許分類コードの選択支援機能

ユーザが選択した特許分類コードが正しいものであるか否かを判断するために、PSA は、抽出した特許分類コードの意味を出力する特許分類コード説明出力機能、同階層コード説明出力機能、および選択した特許分類コードが付されている特許公報を提示する関連公報提示機能を有する。

特許分類コード説明機能の出力例を図4に示す。図4において、ユーザは、「G06F 17/30」の横にある「？」アイコンを選択したとする。特許分類コード説明出力機能では、ユーザが選択した各特許分類コードの説明を階層的に出力する。

同階層コード説明出力機能の出力例を図6に示す。



図6 同階層コード説明出力機能の出力例

図6において、同階層コード説明出力機能では、ユーザが選択した特許分類コードと同じ階層にある特許分類コードを出力する。

なお、特許分類コード説明出力機能、および同階層コード説明出力機能で出力された特許分類コードは、ユーザが選択することで、検索式に追加することができる。例えば、特許分類コード説明出力機能で出力された上位階層の特許分類コード（具体的には、「G06F 17/30」の上位階層の「G06F」等）を検索式に追加すると、選択した特許分類コードを前方一致のための検索キーとして用いることができる。

また、米国版の各コード説明機能の出力例を図7に示す。

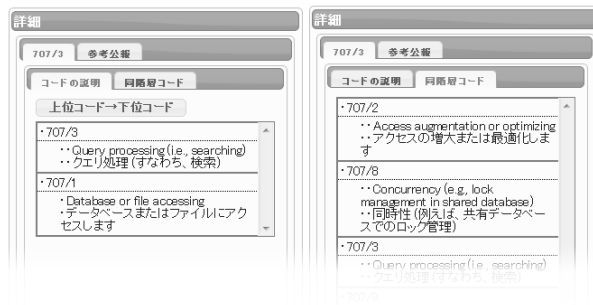


図7 米国版の各コード説明機能の出力例

米国版の特許分類コード説明出力機能は、ユーザが選択した USC、または IPC の説明を階層に出力する。USC の説明は、英文のみであるが、特許分類コード説明出力機能は、英文の説明に対して英日機械翻訳器を用いて、英語と日本語を併記して出力する。また、米国版の同階層コード説明出力機能も、特許分類コード説明出力機能と同様に、USC、または IPC の説明を英語と日本語を併記して出力する。

また、米国版の関連公報提示機能の出力例を図8に示す。米国版の関連公報提示機能も、日本版と同様にして関連公報を提示する。

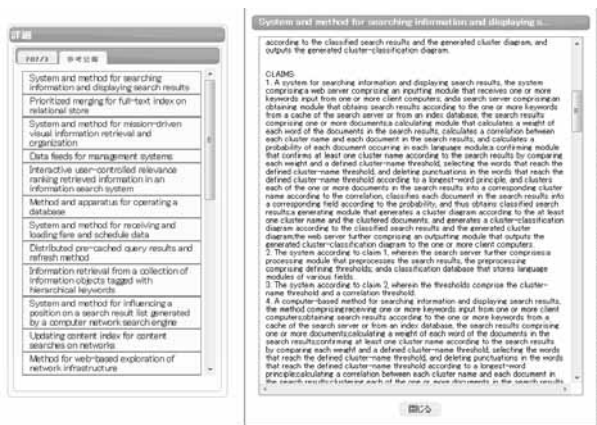


図8 米国版の関連公報提示機能の出力例

存した検索式の削除やダウンロードを行うことができる。ダウンロードは、特許検索サービスに移行する場合に行う。また、社内または部内などで共有し、組織的な知の蓄積に用いることもできる。

また、米国版の検索式保存機能の出力例を図10に示す。

### 4.4 検索式保存機能

検索式保存機能は、ユーザが作成した検索式を保存する機能である。検索式保存機能の使用例を図9に示す。



図9 検索式保存機能の使用例

図9において、ユーザは、「要約+請求項」に「(データベース + データ管理システム) \* (検索)」、 「IPC」に「G06F 17/30」を指定して検索式を保存している。図9の上段の領域は、検索式作成領域である。検索式作成領域では、ユーザの入力したキーワード、ユーザが選択した関連語、およびユーザが選択した特許分類コードを含めた検索式が作成され、表示される。なお、検索式作成領域では、ユーザが手動で検索式を編集することもできる。

ユーザが、検索式を作成し終えて検索式保存ボタンを押すと、検索式が保存される。図9の下段の領域は保存検索式表示領域である。保存検索式表示領域では、保



図10 米国版検索式保存機能の出力例

図10において、米国版の検索式保存機能も、日本版と同様である

## 5 PSAの検証

「OS」、「タスク」、「速度」をキーワードとする調査を例に、特許電子図書館 (IPDL) を用いて、本ツールの検証を行った。

検索式「要約+請求の範囲：OS \* タスク \* 速度」による検索結果は41件であり、そのうち適合特許は13件であった。適合特許とは、調査対象として適切な公報であり、人手で公開特許公報を読んで判断した。

一方、PSAを用いて作成した検索式「要約+請求の範囲：(OS+オペレーティングシステム+Operating System+基本ソフト+基本ソフトウェア) \* (タスク+プロセス+ジョブ+スレッド+task) \* (速度+スピード) \* (G06F 9/46 + G06F 9/45 + G06F



9/50)」による検索結果は 39 件であり、そのうち適合特許は 35 件であった。PSA により、調査対象数が 41 件から 39 件に減少し、適合数が 13 件から 35 件に増加しており、再現率、適合率がともに上昇し、極めて有効な検索式が提案できたことが分かる。

## 6 おわりに

以上、非専門家でも、専門家並みの特許調査を可能にする日本語版および米国版の PSA と、その効果について説明した。今後、中国版 PSA、EU 版 PSA の開発も行いたい。

### 謝辞

本論文は、独立行政法人情報通信研究機構（NICT）の民間基盤技術研究促進制度に基づく委託研究「知的財産（特許・商標）構築・活用のための情報通信基盤技術の研究開発」の研究成果に基づきます。

### 参考文献

- 1) 谷川英和他, 特許工学入門, p1 ~ p7(2003), 中央経済社
- 2) 難波英嗣他, 「特許、論文データベースを統合した検索環境および動向分析ツールの構築」, 2006 年度 JAPIO 「特許情報活用の時代の検索と機械翻訳技術」, pp116-119, 2006.