

統計分析用特許データベース の構築と意義

IIPデータベース

東京大学工学系研究科教授

元橋 一之

PROFILE

東京大学工学修士、コーネル大学MBA、慶応大学博士（商学）。経済産業省、OECDエコノミスト、一橋大学助教授などを経て2006年から現職。主な著書に『ITイノベーションの実証分析』（東洋経済新報社）、『日本経済競争力の構想』（日本経済新聞社）など。



財団法人知的財産研究所

山内 勇

PROFILE

2006年より現職。特許統計を用いた調査研究に従事。



1 はじめに

日本を始めとする多くの国において、特許出願の情報は18ヶ月以内に公開される。この公開特許情報は、新たに特許出願を行う際の先行技術調査に関する情報として有用だが、技術分野別のパテントマップの作成や研究開発戦略の策定への応用など、より高度な使い方も存在する。その中でイノベーション活動を測定するための統計情報としての活用方法が脚光を浴びている。

公開特許データを用いたデータベースとしては、米国の研究者が中心になって作成したNBERパテントデータベースが有名であるが、最近ではEPO（ヨーロッパ特許庁）がPATSTATという世界の特許出願に関するデータベースを公開し、研究者による利用が活発化している。また、OECDにおいては、特許情報の統計利用を促進するために日米欧の3極特許庁からなる特許統計に関するタスクフォースを結成し、毎年会議を行ってきた。

このような世界的な動きに対応して、日本においては「IIPデータベース」が開発された。このデータベースは、（財）知的財産研究所のHP上（<http://www.iip.or.jp/>）で公開されており、研究者向けに無料で提供されている。

2 IIPデータベースの構築

IIPデータベースは、後藤晃氏（東京大学名誉教授）や元橋一之氏（東京大学教授）らが中心となって、特許庁の整理標準化データから必要な項目を抽出し、実証分析用データベースとしてまとめたものである。このデータベースには、1964年1月以降の出願データが収録されており、現時点では、2004年1月までに公開された約902万件のデータが蓄積されている。さらに、各特許出願に対応した、審査・登録・引用データや出願人・権利人のデータなども含まれている。

データベース構築に当たって、延べ約5,600万件の整理標準化データから、イノベーション研究に有用な項目のみが選択・抽出されている。その上、更新時点ごとの出願人・権利人コードの違いが統一され、特許庁内部で使用される中間コードも変数としてまとめられている。さらに、データ形式は統計処理が行いやすい形式に変換されている。

こうした一連の作業により、実証分析に有益なデータセットが、(1) 特許出願ファイル、(2) 特許登録ファイル、(3) 出願人ファイル、(4) 権利者ファイル、(5) 引用情報ファイルという5つのカテゴリの下、コンパクトにまとめられ、容量も一般ユーザーにとって扱いやすい大きさに抑えられている（詳細は、後藤晃・元橋一之「特許データベースの開発とイノベーション」『知財研フ

オーラム』知的財産研究所 Vol.63、p.43-49、2005年を参照)。

3 IIPパテントデータベースの意義

出願件数や審査請求件数などのマクロ・データを見れば、一国の大まかな技術動向を把握することはできる。しかし、それだけでは、制度変更の効果や企業の知財戦略に影響を与えている要因を特定することはできない。

この点、IIPパテントデータベースを用いると、発明の性質や企業の属性までを考慮した分析が可能となり、より詳細かつ正確に特許制度や技術戦略に関する評価を行うことができるようになる。

IIPパテントデータベースの特許出願ファイルには、出願番号ごとに出願日、審査請求日、出願人名、請求項数、筆頭IPCなどのデータが含まれているが、例えば、審査請求される特許出願の特徴を知りたいければ、審査請求された出願と審査請求されなかった出願との間で、請求項数の違いや技術分野の違い、また、出願人の属性等の違いを比べればよい。これにより、請求項数やIPCなどの違いが、審査請求される確率にどの程度の効果を持っているかが分かる。

同じような考え方に基づいて、登録率や権利維持期間など分析者のニーズに応じた様々な対象を、年、技術分野、出願人のタイプといった諸々の属性により詳細に分析することができる。

さらに、出願人データを利用して財務データとの接続を図れば、企業の売上高や研究開発費などが持つ影響を分析することもできるし、出願日データを用いて特許制度が変更された日の前後でデータを区別すれば、そうした制度変更の効果も分析することもできる。

また、IIPパテントデータベースに含まれる引用・被引用データは、技術の発展プロセスや特許の価値を考慮した分析を行う上で極めて重要なデータである。

上述のような分析を厳密に行うには、ある程度の専門知識が必要とされ、また、扱うデータの量も膨大になりがちであるが、近年のコンピュータやソフトウェアの進歩に伴い、こうした研究は急激に進展してきている。

海外における特許データベースに目を向けると、前述の通り、NBER（全米経済研究所）のデータベース、

EPO（欧州特許庁）のPATSTATなどがあるが、特許の質の偏りやデータの精度を考慮すると、日本の企業活動や政策効果を分析する際には、日本の特許データベースを用いる方が望ましいと言える。

ただし、近年の企業活動はグローバル化しており、こうした海外のデータベースとIIPパテントデータベースを上手く組み合わせれば、外国出願されているかどうかといった種々の属性を考慮した、より質の高い研究を行うことも可能である。

4 今後のアップデート

今後の知識社会においては、知識の創出・活用を含めた戦略的な知財管理がますます重要になってくると考えられる。それに対応して我が国でも様々な特許制度改革が行われてきている。こうした中で、直近の制度変更の効果や技術動向を的確に把握するためには、最新のデータを使った分析が必要となる。

そこで、IIPパテントデータベースは整理標化データの平成19年度第5回配布版（2007年5月時点までの情報を収録）までを取り込んだデータのアップデートを行っているところである。データのアップデート作業は、本年度末をめどとして行っており、完成後すみやかに公開される予定である。

なお、今回のアップデートでは、発明者データも追加する予定である。これにより、出願人レベルよりもさらに詳細な発明者レベルでの分析が可能となる。例えば、報奨制度が発明者の意欲に与える影響や、産学連携の効果など、様々な研究において、より質の高い成果が得られることが期待される。

我が国が今後も競争力を維持していく上では、イノベーションのプロセスを効率化していく努力が必要となるが、IIPパテントデータベースは、そうした活動を評価・分析する上での基礎データを提供するものである。

このデータベースが広く研究に利用され、我が国のイノベーション研究や特許制度・技術政策に関する議論が、より高度なものへと発展することを切に望んでいる。