

産業における研究開発・イノベーションに関するデータ整備について  
(2013年 11月 18日)

1. 全体構想

文部科学省 科学技術・学術政策研究所(NISTEP)では、平成 23 年度(2011 年度)から、文部科学省の「科学技術イノベーションにおける“政策のための科学”推進事業」の一環として「産業における研究開発・イノベーションに関するデータ整備」を進めています。

これは、産業セクターにおける研究開発とイノベーションの実態およびイノベーション・プロセスについての理解の深化、産業イノベーションにおいて政府が果たしている／果たすべき役割の解明、産業別のイノベーションの特性の解明、さらに、技術知識と知財の性質の解明など、産業における研究開発・イノベーションについての理解を深めるために民間企業の知財、研究開発、事業等に関するデータを体系的に整備し関係付けした情報基盤の構築を行うものです。

特に、企業情報検索の索引ともいえる企業名辞書を中心に研究開発インプットとアウトプットのデータをマイクロレベルでリンクさせ、名称や形態の変遷が日常的ともいえる企業の情報を容易に追跡でき、また、定量的・構造的に分析できるようにすることを目指しています。

その全体的な構想は、図 1 に示すようなモデルで示されます。前述のような分析が可能になるよう、特許データ等の企業名寄せや各データを連結するための共通基盤的なデータを中心に、データ整備を進めています。

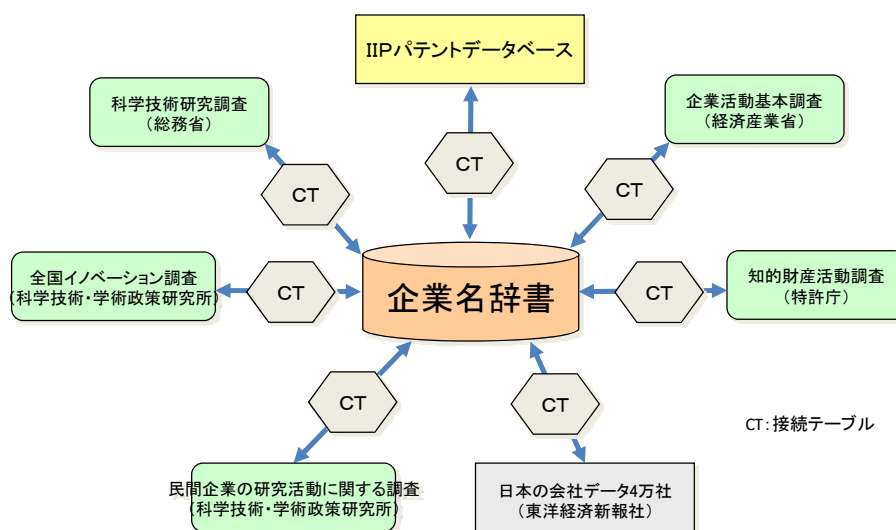


図 1 産業における研究開発・イノベーションに関するデータ整備の概念モデル

## 2. 具体的なデータ整備内容

この事業を開始した 2011 年度より、以下のデータの整備を行っており、データの調整等が済み次第、順次、公開します。(統計法の規定により公的統計調査に関するデータや接続情報は除く。)

### (1) NISTEP 企業名辞書 (以下、「企業名辞書」と略)

企業名辞書には、特許出願実績を 100 件以上有する企業、および、株式上場企業を掲載しています。

この企業名辞書は、変遷を含めた企業単位、企業グループ、業種・規模単位等の分析のための索引、および、基礎情報源として使用することができます。

民間企業では、名称の変更、合併等による企業形態の変化、清算・倒産などが日常的であり、データ分析において、旧名称時代から現在までの一貫通貫させた取り扱いの困難さがあります。企業名辞書では変遷企業名を収録するとともに、それら個別、または、グループ化して関係データを取り扱えるよう配慮した構造になっております。

### (2) 企業名の名寄せ

特許データ、各種統計調査等のデータベースの名寄せ・同定を行い、企業名辞書と対応付けを行いました。

特に、特許データの企業出願人の名寄せにより、1970 年から 2010 年までの企業の特許出願総数 (約 1 千万件) の約 90% が 100 件以上の実績を持つ企業による出願で占められていること確認でき、これら企業を企業名辞書に掲載いたしました。

なお、特許データは、IIP パテントデータベース (一般財団法人知的財産研究所より公開される特許庁の整理標準化データをもとに特許統計分析用に開発されたデータベース) を対象に名寄せ処理しています。

### (3) 各種情報源の企業情報のリンク

企業名の名寄せにより、企業名辞書と特許データ、各種統計調査等のデータベース間の同一企業情報をリンクする接続テーブルを生成いたしました。

これにより、イノベーション研究に必要なデータセット間に跨がるイン・アウト指標を適切に関連付けた、分析に供せるデータの取り出しを可能とすることを実現しました。

## 3. データの公開について

「産業における研究開発・イノベーションに関するデータ整備」事業で得られた成果は、できる限り公開し、利用者の研究・調査等の目的に活用していただくとともに、ご意見・ご要望をいただいてデータや事業の改善を図りたいと考えています。

2 で述べたデータセットのうち、現在、以下の 3 つのデータを web サイトにおいて公

開しています。

(a)NISTEP 企業名辞書データ：2 (1)で説明したとおりです。

(b)外部データ接続テーブルデータ：NISTEP 企業名辞書と以下の外部データとの接続テーブルデータで、2 (2)/(3)で述べた整備により得られたものです。

- ・ IIP パテントデータベース
- ・ 日本の会社データ 4 万社（東洋経済新報社）

これらの内容及び使用方法については、それぞれ、「NISTEP 企業名辞書利用マニュアル」をお読み下さい。