

OECD Activity on Patent Statistics

Economy Analysis and Statistics Division, Directorate for Science, Technology and Industry, OECD

Hélène Dernis

PROFILE

Hélène Dernis is a statistician in the Directorate for Science, Technology and Industry at the OECD. She has been responsible for the creation of a database on patents and contributed to methodological studies, within the framework of the OECD Patent Project. She is now responsible for the development of a micro-data infrastructure to which patent and trademark data are a central element.

Economy Analysis and Statistics Division, Directorate for Science, Technology and Industry, OECD

Ryo Funakoshi

Ryo Funakoshi is an administrator in the Directorate for Science, Technology and Industry at the OECD since 2010. Before the OECD, he was a patent examiner in the Japan Patent Office and was also in charge of research of technological trends in Patent Applications and planning policy for Industrial Property Digital Library (IPDL).

1 OECD work on Patent Statistics

The Organisation for Economic Co-operation and Development (OECD) launched a project on patent statistics in the late 1990s, aiming at developing a statistical infrastructure for patents. The objectives of the project are 1) to set up a database for patent statistics, 2) to establish methodologies for building patent-related statistics, and 3) to improve quality of patent statistics indicators.

In this framework, the OECD, in 2002, established the Patent Statistics Task Force whose members are European Commission (EC), European Patent Office (EPO), EUROSTAT, Japan Patent Office (JPO), Korean Intellectual Property Office (KIPO¹), National Science Foundation (NSF), United States Patent and Trademark Office (USPTO) and World Intellectual Property Organization (WIPO).

2 Database for patent statistics

In order to develop a common international database for building patent statistics, the EPO, one of the task force members, developed a new database suitable for statistical analysis, with the support of the OECD. The first edition of the Worldwide Patent Statistical Database, also known as PATSTAT, was released in April 2006. Since then, the OECD patent statistics mainly rely on PATSTAT data. However, before 2006, as there was no common patent database for statistical analysis, the OECD had built its own database which consisted of DOCDB and Epoline datasets distributed by the EPO.

The PATSTAT² database is released twice a year (in April and September) and includes bibliographic data on patent filed at more than 80 patent offices worldwide. The data includes information concerning:

- Application data

- Publication data
- Citation data
- Family data
- Applicant data
- Inventor data
- Classification data

As a complement to PATSTAT, the OECD has developed four additional databases for patent statistics.

1) Triadic Patent Families³

A triadic patent family is defined as a set of patents filed at the EPO and the JPO and granted by the USPTO, to protect a same invention.

2) REGPAT⁴

The REGPAT database provides regional codes based on the addresses of patent applicants and inventors.

3) Citations in EP & PCT patents⁵

The citations database provides consolidated patent citations from the EPO and international (PCT) patents, in addition to citations to non-patent literature (NPL).

4) HAN database

The HAN database proposes a set of harmonized patent applicant names, grouped together according to a common identifier.

Each of these databases can be linked to PATSTAT for more elaborate analysis.

3 Methodology for Patent Statistics

The OECD published a first manual “The Measurement of Scientific and Technological Activities: Using Patent Data as Science and Technology Indicators” in 1994, in the context of using patent data as a complementary science and technology indicators. Since the 1990s, the use of patent statistics for analytical purposes has increased markedly not only within OECD member countries but also in academia.

As patent statistics became more visible, concerns were raised about the comparability and the interpretation of patent statistics. Patent statistics were compiled in different ways by governments, IP offices and researchers,

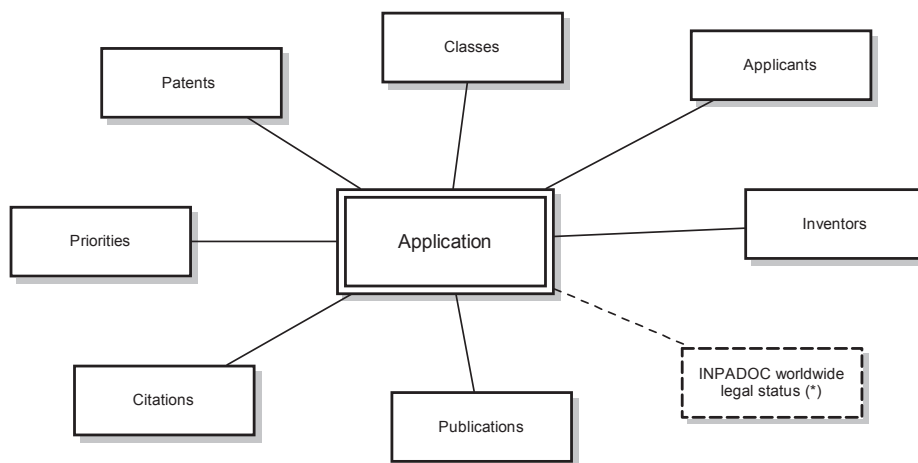


Fig1. Overview of PATSTAT data (Source: PATSTAT, April 2011)

according to their respective analytical needs (e.g., counting patents to measure R&D outputs, innovation, workload, etc.). The methodologies behind the construction of patent statistics are essential to a better understanding of the statistics and an improved quality of the data. How to define patent families, how to count co-inventions, how to count citations are examples of issues that can be addressed.

In order to clarify these issues, the Patent Statistics Task Force agreed to compile new guidelines regarding the methodology for building and interpreting patent statistics. The OECD, with the support of the Patent Statistics Task Force and researchers interested in the

field, revised extensively the previous manual. The “OECD Patent Statistics Manual” published in 2009 provides methodological guidelines regarding what patent statistics can or cannot be used for, and how patents shall be counted in order to maximize information on Science and Technology activities while minimizing noise and bias.

4 New patent indicators

The OECD has developed a variety of internationally comparable indicators on patents using different approaches – e.g. assessing the innovative performance of countries, of regions,

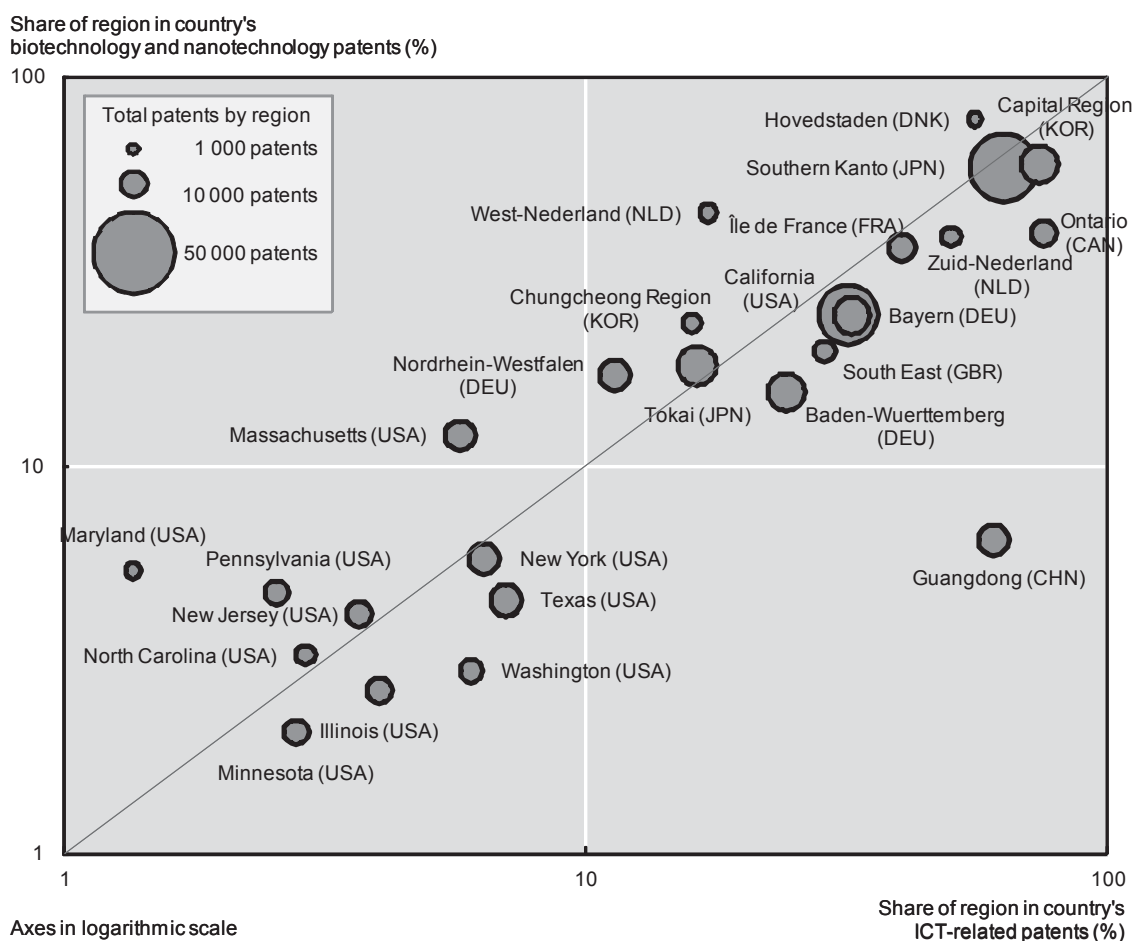


Fig2. Innovation hot spots in ICT, Biotechnologies and nanotechnologies, 2006-2008 (Source: OECD, STI Scoreboard 2011)

of firms; the innovative activity in selected technology domains; the co-operation in S&T activities, etc. Some examples are presented hereafter.

5 Patent Statistics Conference and Workshop

The OECD coordinates the organization of patent statistics conferences and workshops

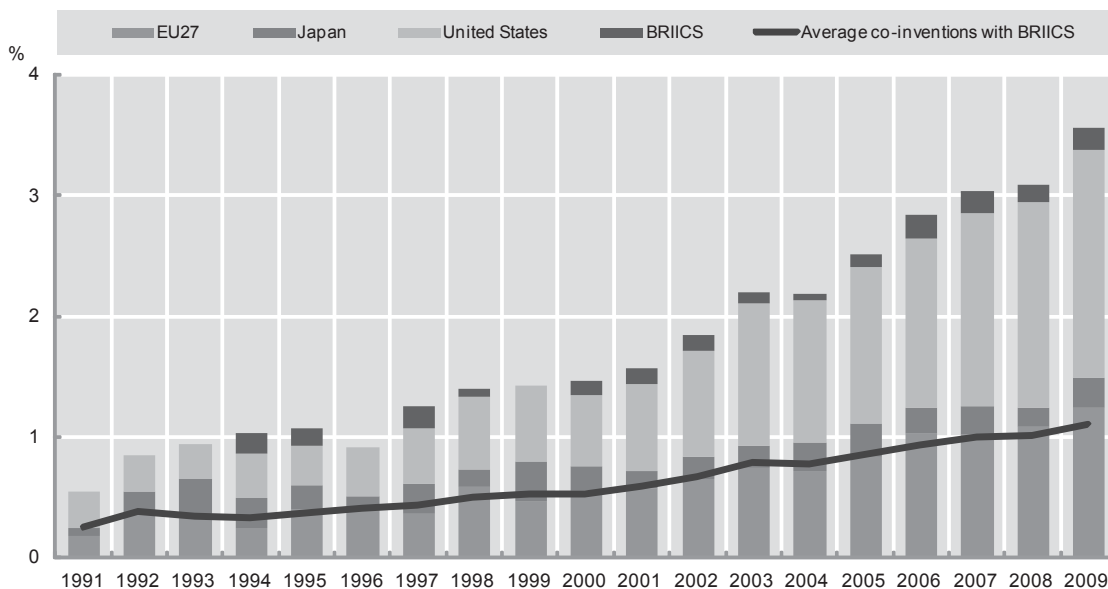


Fig3. Co-inventions with BRICS countries, 1991-2009 (Source: OECD, STI Scoreboard 2011)

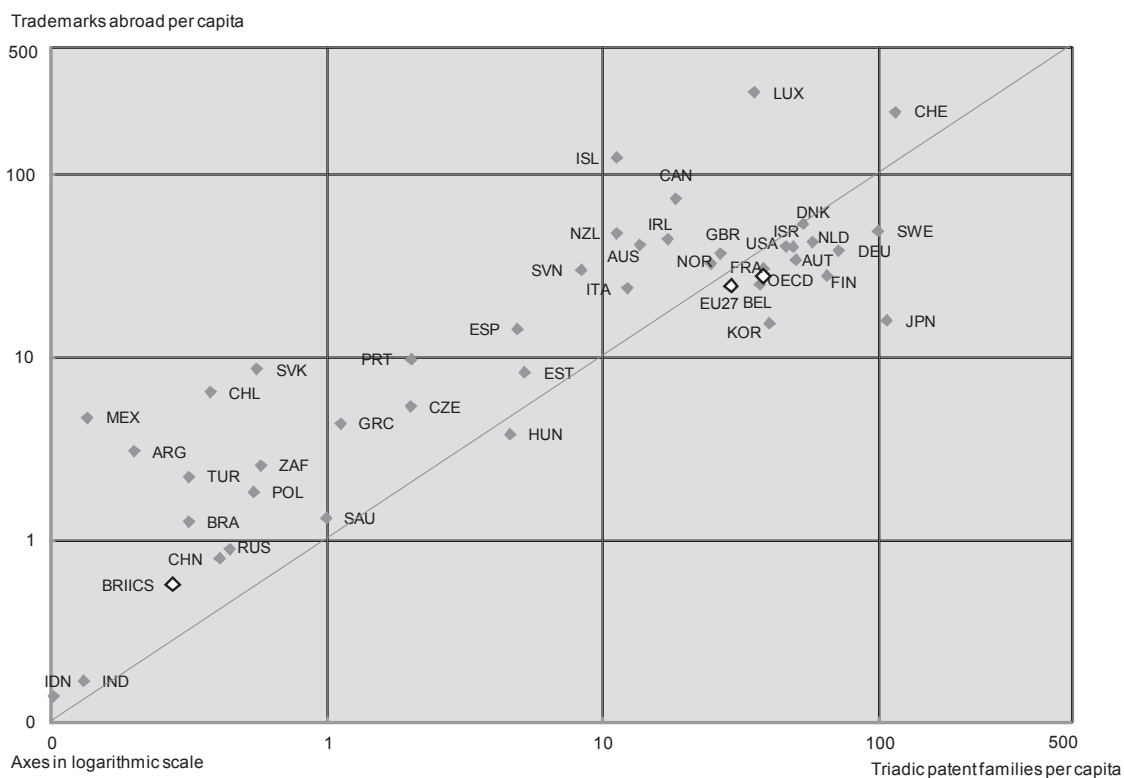


Fig4. Patents and Trademarks per capita, 2007-2009 (Source: OECD, STI Scoreboard 2011)



all over the world with other members of the Task Force. The first event was held in Geneva in 2003, with the support of the WIPO. From 2005 to 2010, the EPO organised the annual conference on patent statistics for decision makers in Vienna and Venice (2007). The 2011 event will be held in Alexandria (US) with the support of the USPTO. The annual patent statistics conference aims to gather researchers, industries and IP offices in order to discuss patent related issues, new data and new methodology and relevant policy issues related to patents.

In parallel, workshops have been organized by the OECD to discuss specific issues such as the typologies of patent families, the harmonization of patent applicant names, and, more recently, the development of trademarks statistics.

Conferences and workshops on patent statistics are typically the place where new methodological development, new statistics and new data emerge.

3. See "Hélène Dernis and Mosahid Khan, Triadic Patent Families Methodology, STI Working Paper 2004/2, Statistical Analysis of Science, Technology and Industry"
4. See "Stéphane Maraut, Hélène Dernis, Colin Webb, Vincenzo Spiezia and Dominique Guellec, THE OECD REGPAT DATABASE: A Presentation, STI Working Paper 2008/2, Statistical Analysis of Science, Technology and Industry"
5. See "Webb.C, H.Dernis, D.Harhoff and L.Hoisl, Analysing European and International Patent Citations: A set of EPO Patent Database Building Blocks" , STI Working Paper 2005/9"

6

Contact

For further information on the OECD patent-related analysis and access to the OECD patent statistics, please refer to the OECD web site at : www.oecd.org/sti/ipr-statistics

<Annotation>

1. KIPO joined in Patent Statistics Task Force in 2010.

2. See EPO' s web site for details.

<http://www.epo.org/searching/subscription/raw/product-14-24.html>

OECDの特許統計に対する取り組み（和訳）

経済協力開発機構 科学技術産業局経済分析統計課 H el ene Dernis

OECD 科学技術産業局の統計専門官。OECD 特許プロジェクトにおいて、特許に関するデータベースの構築および方法論の研究に従事している。現在は特許と商標を中心としたマイクロデータのインフラ開発に力を注いでいる。

経済協力開発機構 科学技術産業局経済分析統計課 船越 亮

2001年に特許庁入庁。情報分野の特許審査、特許出願動向調査の担当、IPDL（特許電子図書館）の企画担当、ケンブリッジ大学客員研究員などを経て、2010年よりOECDに出向。

1 OECDの特許統計に対する取り組み

OECD（経済協力開発機構）は、1990年代後半に特許統計のインフラ発展を目指し、特許統計に関するプロジェクトを立ち上げた。そのプロジェクトの目的は、1) 特許統計用のデータベースを設けること、2) 特許統計の方法論を確立すること、3) 特許統計の指標の質を向上させること、である。

この枠組みの中で、OECDは2002年に、特許統計タスクフォースを設立した。タスクフォースメンバーは、欧州委員会（EC）、欧州特許庁（EPO）、欧州連合統計局（Eurostat）、日本特許庁（JPO）、韓国特許庁（KIPO¹）、米国国立科学財団（NSF）、米国特許商標庁（USPTO）、世界知的所有権機関（WIPO）である。

2 特許統計用データベース

共有で且つ国際的な特許統計用のデータベースを開発するため、特許統計タスクフォースメンバーの一員であるEPOはOECDの協力のもと、統計分析に適する新しいデータベースを開発した。そしてPATSTATとして知られるワールドワイドな特許統計データベースの第1版が2006年4月にリリースされた。それ以降、OECDの特許統計はPATSTATを用いることが

主流である。しかしながら、2006年以前は、共有の統計分析用の特許データベースは存在していなかった。そのため、OECDはEPOが提供しているDOCDBとEpolineのデータから成る自前のデータベースを構築し、それを利用していた。

PATSTATは年に2回（4月と9月）リリースされ、80以上の特許庁の書誌情報を含んで成る。これらの書誌情報は、

- －出願データ
 - －公報データ
 - －引用データ
 - －ファミリーデータ
 - －出願人データ
 - －発明者データ
 - －分類データ
- などから構成される。

OECDでは、PATSTATを補足するものとして、4つの特許統計用の付加的データベースを開発している。

1) Triadic Patent Families²

ここでの三極ファミリーの定義は、同じ発明を保護するために、EPOとJPOに出願されたものであって、USPTOで特許になった一連の特許出願と定義している。

2) REGPAT³

REGPATは出願人の住所と発明者の住所の各々に地

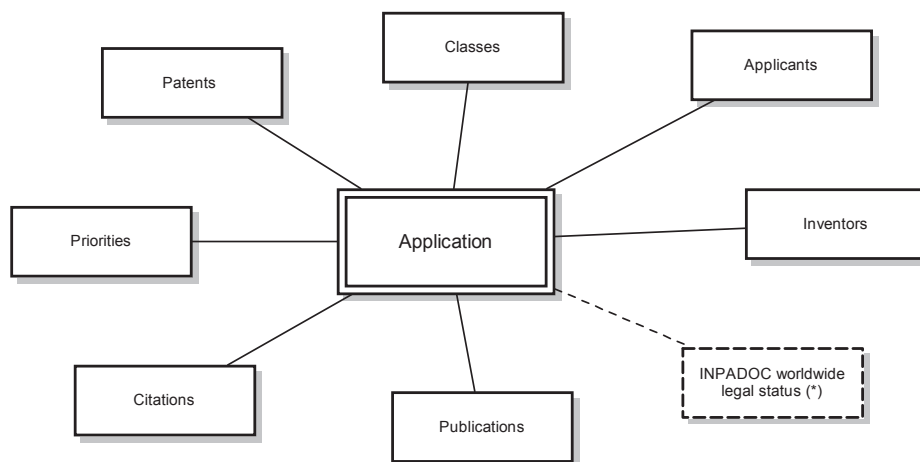


図1 PATSTAT の概要 (出典: PATSTAT April 2011)

域コードを付与したものである。

3) Citations in EP & PCT patents⁴

このデータベースは、EPO 出願および国際的出願 (PCT) における引用文献を収集しており、特許文献だけでなく非特許文献も含まれている。

4) HAN database

HAN データベースは名寄せされた一連の特許出願人名であり、共通の識別子によりグループ化されている。

いずれのデータベースもより詳細な分析のために、PATSTAT とリンクを図ることが可能である。

および解釈についての関心事が提起された。これは、特許統計が、各政府、各特許庁、そして研究者の分析軸 (R&D のアウトプットを測定するために特許をカウントするのか、イノベーションを測定する目的なのか、それとも、ワークロードを測定する目的なのか、等々) により異なる手法で収集されたことによる。特許統計を解釈する際の背景にある方法論には統計に対する理解とデータの質向上が不可欠である。パテントファミリーをどのように定義するのか、共同発明者の場合は発明者をどのようにカウントするのか、引用文献が複数ある場合は引用文献をどのようにカウントするのか、などは提起された例の一例である。

3 特許統計の方法論

OECD は最初のマニュアル “The Measurement of Scientific and Technological Activities: Using Patent Data as Science and Technology Indicators” を 1994 年に発行した。このマニュアルでは、科学技術指標を補足するものとして、特許データの利用が取りあげられていた。そして 1990 年代以降、分析を目的とした特許統計の利用は、OECD 加盟国だけではなく、大学においても飛躍的に増加した。

特許統計が普及する一方で、特許統計データの比較性

これらの論点を明らかにするため、特許統計タスクフォースは特許統計の構築と解釈に向けた方法論に関する新しいガイドラインを編集することに同意した。OECD は、特許統計タスクフォース及びその分野の研究者の協力のもと、前述したマニュアルを大々的に改訂した。そして、2009 年に “OECD Patent Statistics Manual” が発行された。このマニュアルは、何に対して特許統計を利用することが可能であり、何に対しては利用されないのか、そして、ノイズやバイアスを最小限にする一方で、科学技術に関する情報を最大化するために、どのように特許をカウントするべきか等について、方法論的な指針を示している。

4 特許に関する新しい指標

OECDは様々なアプローチに基づいて、国際的に比較可能な特許指標を幅広く創出している。たとえば、各国、各地域、各企業のイノベティブなパフォーマンスを評価するための指標、特定の科学技術領域でのイノベティブなパフォーマンスを評価するための指標、科学技術領域における共同活動を評価するための指標、等である。以下にその例を示す。

5 特許統計カンファレンスとワークショップ

OECDはタスクフォースメンバーと共催で、世界中で特許統計カンファレンスとワークショップを開催している。最初はWIPOの協力のもと、2003年にジュネーブで開催した。2005年から2010年は毎年、EPOがウィーンとベネチア（2007年）において、政策立案者を対象としたカンファレンスを開催した。2011年はUSPTOの協力のもと、アレキサンドリア（米国）で開催される予定である。特許統計カンファレンスの目的は大学、産業界、各特許庁を一同に集め、特許に関連した問題、新しいデータや方法論、そして特許に関連した政策について議論をすることである。またカンファレンスと並行して行われるOECD開催のワークショップ

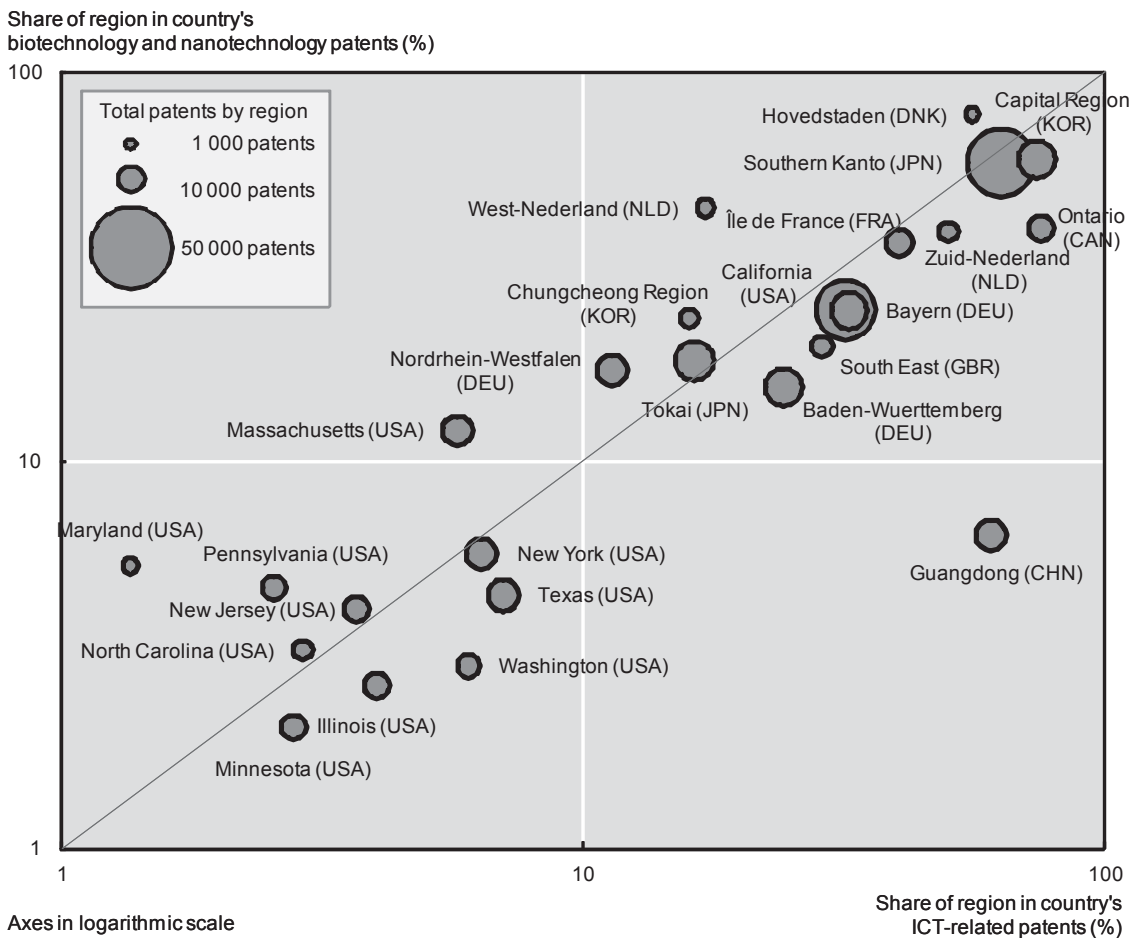


図2 ICT、バイオテクノロジー、ナノテクノロジー分野におけるイノベーションホットスポット、2006-2008 (出典：OECD STI Scoreboard2011)

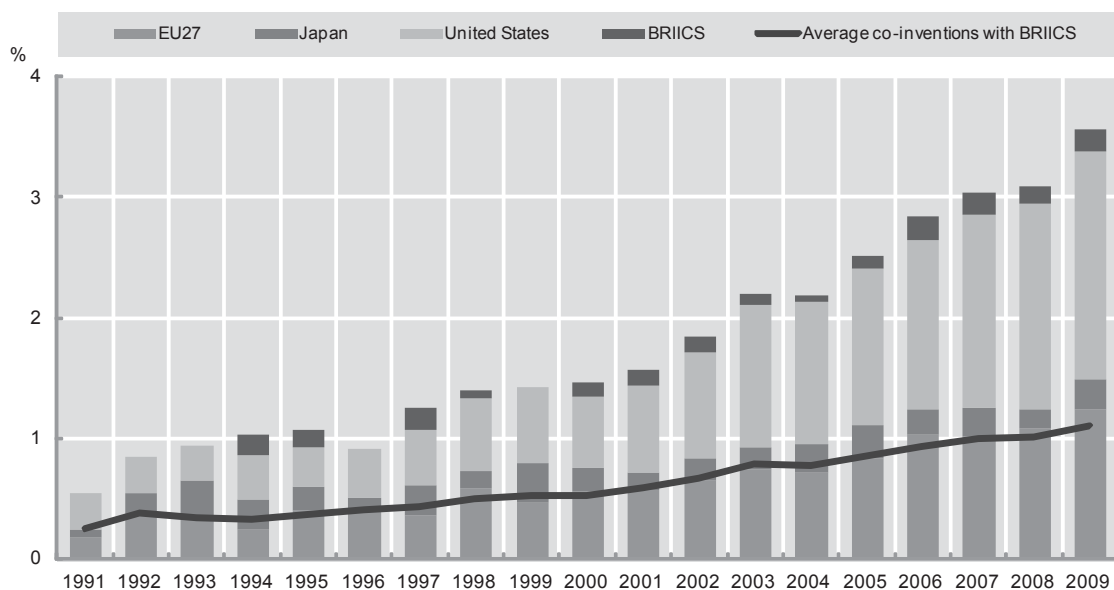


図3 Bricsとの共同発明、1991-2009（出典：OECD STI Scoreboard2011）

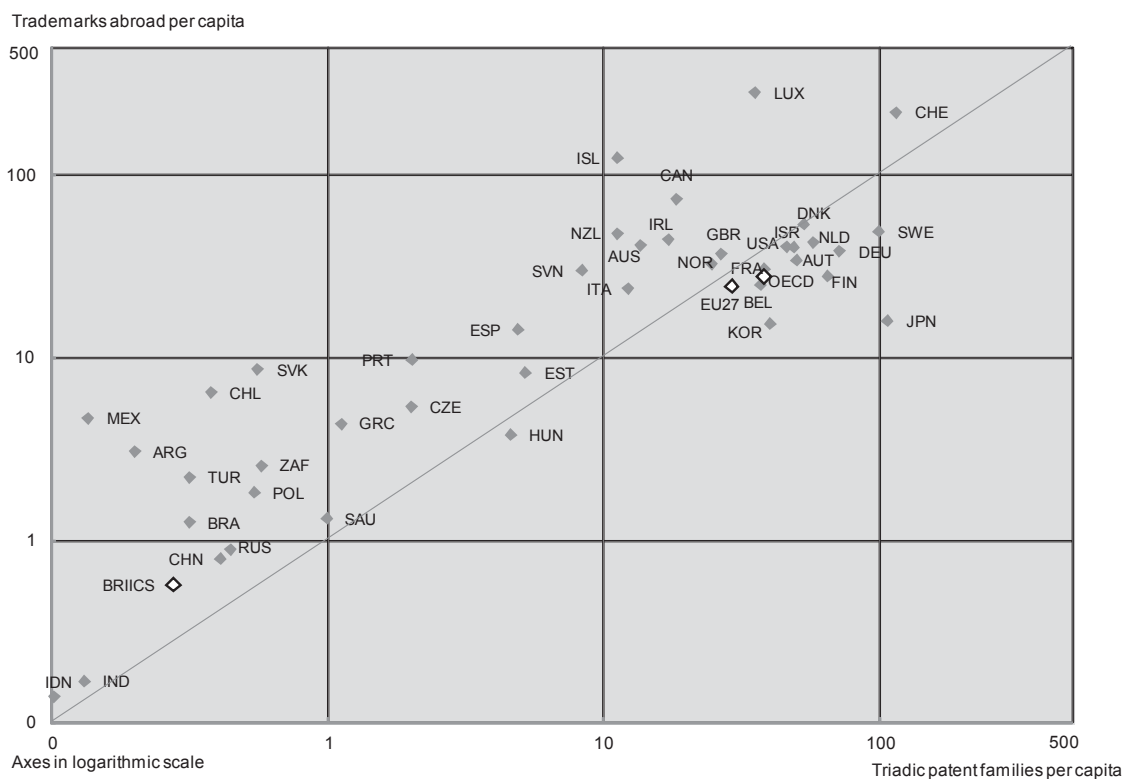


図4 各国一人当たりの特許出願と商標出願、2007-2009（出典：OECD STI Scoreboard2011）

では、特許ファミリーの類型、出願人名の名寄せ、そして最近では商標統計の開発についての議論を行っている。

特許統計に関するカンファレンスとワークショップは、新しい方法論の開発、新しい統計、新しいデータが創出される場所となっている。

6 連絡先

OECDの特許分析についての詳細な情報、OECDの特許統計にアクセスするためには、OECDのウェブサイトを訪ねてください。

www.oecd.org/sti/ipr-statistics

<注釈>

1. 2010年にKIPOが特許統計タスクフォースとして参加した。
2. 詳細はEPOのウェブサイトを参照。
<http://www.epo.org/searching/subscription/raw/product-14-24.html>
3. “Hélène Dernis and Mosahid Khan, Triadic Patent Families Methodology, STI Working Paper 2004/2, Statistical Analysis of Science, Technology and Industry” を参照。
4. “Stéphane Maraut, Hélène Dernis, Colin Webb, Vincenzo Spiezia and Dominique Guellec, THE OECD REGPAT DATABASE: A Presentation, STI Working Paper 2008/2, Statistical Analysis of Science, Technology and Industry” を参照。
5. “Webb.C, H.Dernis, D.Harhoff and L.Hoisl, Analysing European and International Patent Citations: A set of EPO Patent Database Building Blocks” , STI Working Paper 2005/9” を参照。