

UNIVERSITY OF WUPPERTAL
BERGISCHE UNIVERSITÄT WUPPERTAL

EUROPÄISCHE WIRTSCHAFT
UND
INTERNATIONALE MAKROÖKONOMIK



Zafir Mahmutovic

**Patentdatenbank:
Implementierung und Nutzung**

Diskussionsbeitrag 176
Discussion Paper 176

*Europäische Wirtschaft und Internationale Wirtschaftsbeziehungen
European Economy and International Economic Relations*

ISSN 1430-5445

Zafir Mahmutovic

**Patentdatenbank:
Implementierung und Nutzung**

Juli 2010

Herausgeber/Editor: Prof. Dr. Paul J.J. Welfens, Jean Monnet Chair in European Economic Integration

EUROPÄISCHES INSTITUT FÜR INTERNATIONALE WIRTSCHAFTSBEZIEHUNGEN (EIIW)/
EUROPEAN INSTITUTE FOR INTERNATIONAL ECONOMIC RELATIONS
Bergische Universität Wuppertal, Campus Freudenberg, Rainer-Gruenter-Straße 21,
D-42119 Wuppertal, Germany
Tel.: (0)202 – 439 13 71
Fax: (0)202 – 439 13 77
E-mail: welfens@eiiw.uni-wuppertal.de
www.eiiw.eu

JEL classification: C82, C88, O32, L62

Key words: MySQL, Datenbank, Patentzitation, Soziale Netzwerkanalyse

Summary: In the knowledge society, it is a fundamental interest to capture the technology-and market-relevant knowledge dynamics. This is why the analysis of patent statistics plays an important role. Thereby others regions, companies, sectors and countries as query elements in national and international databases come in question. The sorting of the raw data gives us the possibility to illustrate the particular knowledge networking, for example by using joint patent applications by several companies/persons in the form of patent citations, which reflect the historical context of knowledge creation. In addition, over time, the mobility of inventors and patent applicants between firms and sectors are covered, which is also an indicator for the diffusion of knowledge. The knowledge and economic dynamics of countries and regions can be analyzed. In this paper the main focus does not lie on the procedures and the interpretation of results, but only on the expiry of the provision of the required data. Main components are the development of data, the creation of the patent database (and implementation) and the possible uses of the database. It is the aim to provide data for the method of social network analysis and patent citation analysis. For the sake of complexity a more detailed explanation of complex facts and details from the database theory is omitted.

Zusammenfassung: In der Wissensgesellschaft ist es von elementarem Interesse, die technologie- bzw. marktrelevante Wissensdynamik zu erfassen. Hier spielt die Auswertung von Patentstatistiken eine wesentliche Rolle, wobei u.a. Regionen, Firmen, Sektoren und Länder als Abfrageelemente in nationalen und internationalen Datenbanken in Frage kommen. Durch eine Aufbereitung von Rohdaten kann man insbesondere auch die Wissensvernetzung verdeutlichen, etwa in Form gemeinsamer Patentanmeldungen durch mehrere Firmen/Personen oder unter Hinweis auf Patentzitierten, die den historischen Wissensentstehungskontext mit abbilden. Zudem kann im Zeitablauf die Mobilität von Erfindern bzw. Patentanmeldern zwischen Firmen und Sektoren erfasst werden, was auch ein Indikator für die Diffusion von Wissen ist. Die Wissens- und Wirtschaftsdynamik von Ländern und Regionen etwa kann hiermit abgebildet werden. In dieser Arbeit wird nicht auf die Verfahren und die Interpretation der Ergebnisse eingegangen, sondern nur auf den Ablauf der Bereitstellung der benötigten Daten. Es wird hierbei insbesondere auf die Gewinnung der Daten, der Erstellung einer Patentdatenbank (und Implementierung), der Nutzungsmöglichkeiten der Datenbank eingegangen. Dabei ist es das Ziel Daten für die Methode der Sozialen Netzwerkanalyse und der Patentzitiertenmethode zur Verfügung zu stellen. Aus Gründen der Komplexität wird auf eine tiefer gehende Erklärung von komplexen Sachverhalten und Details aus der Datenbank-Theorie verzichtet.

Zafir Mahmutovic, Research Assistant at European Institute for International Economic Relations (EIIW) at the University of Wuppertal, Rainer-Gruenter-Str. 21, D-42119 Wuppertal, Phone: +49-202-4391371, Fax: +49-202-4391377
welfens@eiiw.uni-wuppertal.de, www.eiiw.eu

Patentdatenbank: Implementierung und Nutzung

Discussion Paper 176

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	II
1. Einführung	1
2. Informationen zu PATSTAT	2
3. Informationen zum DBMS	5
4. Implementierung	6
5. Nutzung	7
5.1. Soziale Netzwerkanalyse	8
5.2. Patentzitimierungsmethode.....	9
6. Fazit	11
Literaturverzeichnis	12

Abbildungsverzeichnis

Figure 1: Konzeptuelles Modell	3
Figure 2: Logisches Modell, Quelle: EPO.....	4

1. Einführung

In der Wissensgesellschaft ist es von elementarem Interesse, die technologie- bzw. marktrelevante Wissensdynamik zu erfassen. Hier spielt die Auswertung von Patentstatistiken eine wesentliche Rolle, wobei u.a. Regionen, Firmen, Sektoren und Länder als Abfrageelemente in nationalen und internationalen Datenbanken in Frage kommen. Durch eine Aufbereitung von Rohdaten kann man insbesondere auch die Wissensvernetzung verdeutlichen, etwa in Form gemeinsamer Patentanmeldungen durch mehrere Firmen/Personen oder unter Hinweis auf Patenzitierungen, die den historischen Wissensentstehungskontext mit abbilden. Zudem kann im Zeitablauf die Mobilität von Erfindern bzw. Patentanmeldern zwischen Firmen und Sektoren erfasst werden, was auch ein Indikator für die Diffusion von Wissen ist. Die Wissens- und Wirtschaftsdynamik von Ländern und Regionen etwa kann hiermit abgebildet werden. Für die Unternehmen und Banken sowie auch für die Tarifvertragsparteien der betreffenden Wirtschaftsräume und die Wirtschaftspolitik ergeben sich damit wesentliche Informationen über die Entwicklung der Innovationsdynamik bzw. die Wissensspezialisierung.

Im Zuge der Globalisierung hat sich in den führenden Industrieländern eine verstärkte Spezialisierung auf wissens- und technologieintensive Produktionsbereiche ergeben, so dass man einerseits in der Industrie auch entsprechende Zuwachsraten an Patenten erwarten kann; zugleich ist andererseits zu vermuten, dass die Komplexität der Innovationsprozesse zunimmt, so dass die Entstehung von Patenten durch Zusammenarbeit im Bereich von Forschung und Entwicklung an Bedeutung gewinnt. Netzwerkstrukturen bei Innovationen bzw. Patentanmeldungen dürften von daher zunehmend wichtiger werden. Die Auswertung von Patentdatenbanken ist zunächst eine methodische Herausforderung und die weiterführende ökonomische der „Schumpeterschen Dynamik“ – also von Innovation und Diffusion – stellt eine oft erhebliche Herausforderung dar. Keineswegs wird hier dabei übersehen, dass es durchaus auch eine erhebliche Innovationsdynamik in vielen Ländern gibt, die sich nicht in Patentanmeldungen bzw. erteilten Patenten widerspiegelt. Im Dienstleistungssektor etwa sind Patente die Ausnahme, Copyrights und Markennamen spielen hingegen in diesem Bereich eine beträchtliche Rolle.

Patente werden auf Märkten wegen der bekannten Marktunvollkommenheiten von Informationsmärkten bzw. Märkten für technisches Wissen kaum gehandelt. Immerhin gibt es einen internationalen Lizenzmarkt, bei dem vorwiegend innerhalb von multinationalen Unternehmen Patente gehandelt werden. Auch ist ein gewisses Cross-Licensing zu beobachten, bei dem multinationale Unternehmen Lizenzen in verschiedenen Bereichen austauschen. Wenn man sich auf die Auswertung von Patentanmeldungen in der Industrie konzentriert, so wird man zunächst von den Informationen in den Patentdatenbanken ausgehen.

Im Zentrum der folgenden Überlegung steht eine Beschreibung der technischen Realisierung der Nutzung von Patentdaten, wie sie für die Analyse der modernen Wissensgesellschaft unerlässlich ist. Patente als Datenquelle haben einen sehr großen Wert, da sie zahlreiche Informationen beinhalten und in einer engen Verbindung mit innovativen Tätigkeiten stehen (OECD, 2001).

Ein Patent enthält unter anderem Informationen über die Erfindung selbst, über die technologische Klasse der Erfindung (IPC-Code), über den/die Inventor/en (inklusive der geographischen Lokalisierung), über die Organisation (inklusive der geographischen Lokalisierung), die die Schutzrechte an der neuen Erfindung erworben hat, sowie über die technologischen Vorgänger des Patents (in Form von Patentzitierten).

Im Rahmen dieses Projektes werden zwei methodologische Vorgehensweisen benutzt um Indikatoren für regionale Innovationsdynamik und Clusterbildung zu erhalten: die soziale Netzwerkanalyse und die Patentzitiertenmethode.

In dieser Arbeit wird nicht auf die Verfahren und die Interpretation der Ergebnisse eingegangen, sondern nur auf den Ablauf der Bereitstellung der benötigten Daten. Dieser unterscheidet sich erst bei den Abfrageprozeduren (Kapitel 5) für die beiden erwähnten Methoden.

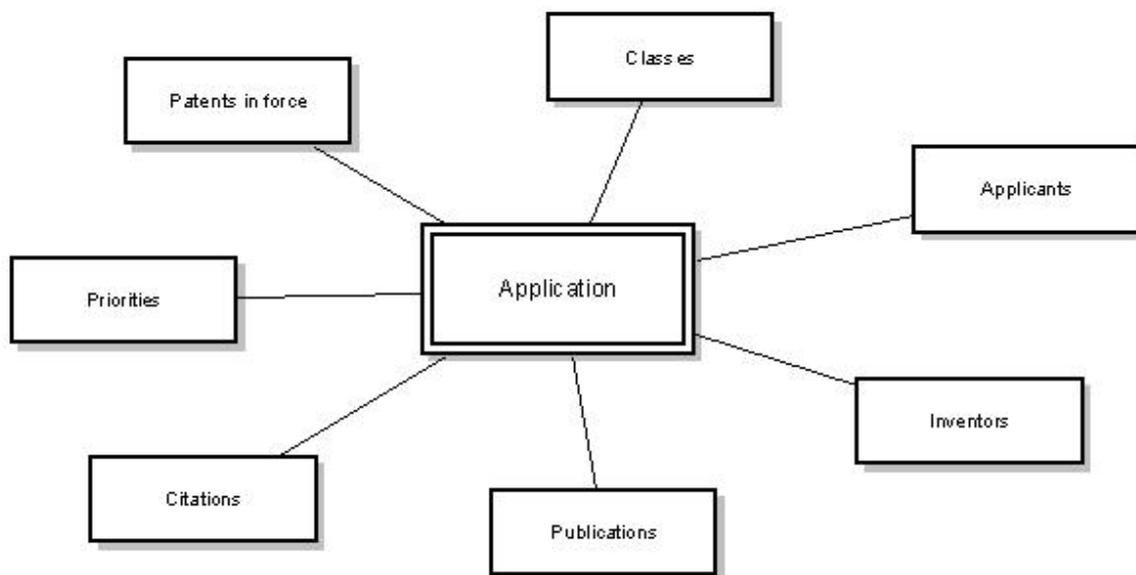
Zunächst wird auf die für Gewinnung der Daten eingegangen, indem die vom Europäischen Patentamt zur Verfügung gestellte, Datenbasis (PATSTAT) im zweiten Kapitel beschrieben wird. Dabei werden allgemeine Charakteristika beschrieben und die Struktur des Datenmodells näher erläutert. Diese Datenbasis bedurfte der Nutzung eines Datenbankmanagementsystems, das im dritten Kapitel erläutert wird. Die genauen Vorgehensweisen bei der Implementierung, Verbesserung und Erweiterung der Datenbank werden in Kapitel vier näher gebracht. Die Nutzung der Datenbank, d.h. die Programmierung der Abfrageprozeduren, die die benötigten Daten ausgeben, sowie die Aufbereitung dieser Daten, wird im abschließenden fünften Kapitel erklärt. Aus Gründen der Komplexität wird auf eine tiefer gehende Erklärung von komplexen Sachverhalten und Details aus der Datenbank-Theorie verzichtet.

2. Informationen zu PATSTAT

PATSTAT stellt die für die Untersuchungen benutzte Datenbasis dar. PATSTAT ist die Patentdatenbank des Europäischen Patentamts (EPO): "EPO Worldwide Patent Statistical Database". Sie beinhaltet eine sehr große Menge an Patenten (\approx 62 Millionen, Stand Oktober 2008) mit einer weltweiten Abdeckung der Patentinformationen. Zwei Mal im Jahr veröffentlicht die EPO an einem bestimmten Zeitpunkt eine Momentaufnahme der Quelldatenbank, die dann publiziert wird (im April und Oktober). Die in diesem Projekt benutzte Version ist vom Oktober 2008.

Die von der EPO erhaltene Datenbank ist das Ergebnis eines dreiteiligen Modellierungsprozesses, welcher die Zusammenhänge und Informationen aus der realen Welt, in ein für ein Datenbanksystem verwertbares Format bringt (EPO, 2008). Der erste Schritt in diesem Prozess stellt dabei das konzeptuelle Modell dar, welches das Ergebnis des Abstraktionsprozesses des Modellentwicklers im Rahmen des Modellentwurfs ist. Es ist rechnerunabhängig und stellt die interessanten Daten zusammengefasst unter Oberbegriffen dar, wobei auch grobe Verbindungen zwischen diesen Gruppen erfasst und mit aufgeführt werden. Die folgende Abbildung zeigt das konzeptuelle Modell, welches für die PATSTAT Datenbank benutzt wurde.

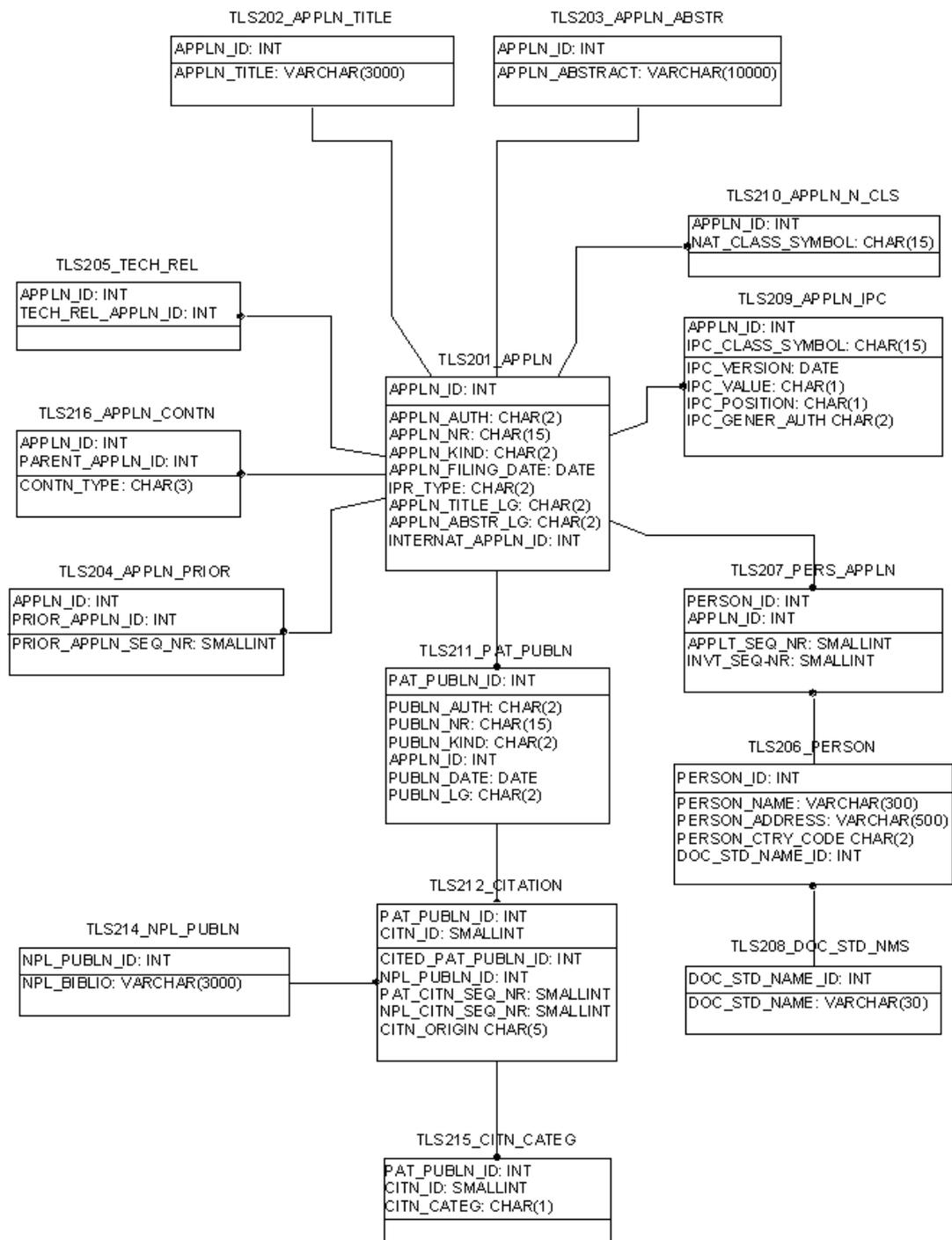
Figure 1: Konzeptuelles Modell



Quelle: EPO (2008)

Aus dem konzeptuellen Modell wird das logische Modell entwickelt, welches einer weitergehenden Konkretisierung der Daten zugrunde liegt. Für dieses logische Modell wurde das relationale Datenbankdesign verwendet. Eine relationale Datenbank kann man sich als Sammlung von Tabellen vorstellen, in welchen Datensätze abgespeichert sind. Jede Zeile in einer Tabelle ist ein Datensatz. Jede Tabelle besteht aus einer Reihe von Attributen, den Spalten der Tabelle. Weiterhin werden Verknüpfungen genutzt, um die Beziehungen zwischen Tabellen auszudrücken. Sowohl der Zusammenhang von Attributen innerhalb einer Tabelle, als auch eine Verknüpfung von Tabellen, stellen eine Relation dar, daher auch der Name. Abb. 2 zeigt das implementierte Design, wobei hier die Attribute der Übersicht in den Zeilen der Tabellen stehen.

Figure 2: Logisches Modell, Quelle: EPO



Quelle: HALL/JAFFE/TRAJTENBERG (2001)

Im Datenbanksystem selber ist das physische Modell implementiert: Unterschiede zum logischen Modell ergeben sich aus Gründen der Datenbanktechnik (Einhaltung von Normalformen und Performance). Somit erhält man am Ende 19 Tabellen mit Erfinderdaten, Bibliographischen Daten etc. die in der Datenbank erstellt werden und in die die Daten eingelesen werden. Mehr dazu im vierten Kapitel.

3. Informationen zum DBMS

Das für die Implementierung von PATSTAT benutzte Datenbankmanagementsystem (DBMS) ist MySQL vom Hersteller Sun Microsystems. Die Gründe hierfür sind:

- Eine sehr gute Performance. Wie es sich in verschiedenen Benchmarks gezeigt hat, ist MySQL eine der schnellsten Datenbanken auf dem Markt und muss sich in dem Punkt höchstens der sehr viel teureren Datenbank von Oracle geschlagen geben.
- Umfassende Möglichkeiten zur Anwendungsentwicklung. MySQL bietet umfassenden Support für die Entwicklung von verschiedensten Client-Anwendungen indem es Schnittstellen und Treiber (ODBC, JDBC etc.) zur Verfügung stellt, die es Programmiersprachen (C++, PHP, Java,...) und anderen Anwendungen erlaubt auf MySQL zuzugreifen.
- Einfache Administration. Die Installation und das Einrichten sind im Falle einer Standardkonfiguration, die für die meistens Anwendungsfälle ausreichend ist, sehr einfach und lassen sich schnell erledigen. Für PATSTAT aber mussten gewisse Änderungen vorgenommen werden, auf die ich gleich zu sprechen komme.
- Sehr niedrige Kosten, für nicht-kommerzielle Anwendungen kostenlos. MySQL wird sowohl unter einer kommerziellen Lizenz als auch als freie Software, unter der GPL, der General Public License, angeboten. Kommerzielle Nutzung ist hierbei ausdrücklich erlaubt, allerdings muss der Quellcode der Anwendung öffentlich zugänglich gemacht werden.

Die Änderungen, die an der Konfiguration des DBMS vorgenommen werden mussten, wurden aus Gründen der Steigerung der Performance gemacht, da PATSTAT aus sehr großen Tabellen besteht, bei denen die Abfragen bzgl. der benötigten Informationen verhältnismäßig viel Zeit in Anspruch nehmen, da viele Tabellen miteinander verknüpft werden müssen. Also wurden die Buffer-Werte in der Config-Datei von MySQL an eine solche Anwendung angepasst. Berücksichtigt wurde auch die Tatsache, dass, sobald die Daten eingelesen waren, später keine größeren Mengen mehr eingefügt werden, sondern nur noch ein lesender Zugriff auf die Datenbank stattfindet. Das war in diesem Fall von Vorteil, da also nur die Werte der Read-Buffer erhöht werden mussten und diese somit mehr Anteile am Arbeitsspeicher des Rechners bekommen konnten.

Das Datenbankmanagementsystem wurde als Single User Umgebung (d.h. MySQL Server und Client befinden sich auf einem Rechner) auf einer Workstation implementiert. Die wichtigsten Eigenschaften der Hardwarekonfiguration der Workstation sind bei einer Datenbankanwendung die Größe des Arbeitsspeichers, welcher großzügig bemessen sein sollte, und die Geschwindigkeit der Festplatte, die sich meistens als Flaschenhals erweist.

4. Implementierung

Bei der Implementierung der von der EPO zur Verfügung gestellten Datenbasis werden entsprechend dem physischen Modell (siehe Kapitel 2) die Tabellen, mit den jeweiligen Attributen und Verknüpfungen untereinander, in der Datenbank erstellt. In diese Tabellen werden dann die Daten eingelesen, die als txt-Dateien (tsv-Format) mit der Zeichensatzkodierung *utf8* vorliegen.

Im nächsten Schritt ist es erforderlich die Datenbank mit diversen Zusatzinformationen zu versehen, die in den Patentdaten selbst nicht enthalten sind. Dazu werden zusätzlichen Tabellen erstellt, die weiterführende Analysen ermöglichen. Um die Branchen, die von Interesse sind, von sonstigen Patenten abgrenzen zu können, ist es notwendig die IPC-Klassen den jeweiligen ISIC Klassen zuordnen zu können. Dieses wird mit einer Konkordanztafel, und zwar der „MERIT concordance table: IPC - ISIC (rev. 2)“, bewerkstelligt, die in der Datenbank als zusätzliche Tabelle erstellt wird (VERSPAGEN/VAN MOERGASTAL/SLABBERS, 1994). Für die Untersuchung von Wissensspillovers zwischen europäischen Regionen ist die Zuordnung der Patente zu den NUTS-Regionen des Untersuchungsraums erforderlich, da die Patente selbst zwar die Adresse der Anmelder und Erfinder, jedoch keine Information über die entsprechende NUTS-Region enthalten. NUTS („Nomenclature des Unites Territoriales Statistiques“) ist eine systematische, hierarchische regionale Gebietsgliederung der Europäischen Union. Die NUTS-Systematik der Gebietseinheiten wurde Anfang der 1980er Jahre von Eurostat im Auftrag der Europäischen Kommission entwickelt, für Zwecke der Regionalstatistik, Erfassung und Harmonisierung regionalstatistischer Daten.

Für die Zuordnung der Adressen der Anmelder und Erfinder zu den NUTS-Regionen wird die OECD REGPAT Datenbank benutzt, und zwar die Version von Januar 2009, mit der Abdeckung aller OECD Länder, China und Indien. Die OECD hat die Adressen zu den NUTS-Regionen mit Hilfe der Postleitzahlen in den Adressen zugeordnet. Wo die Zuordnung über Postleitzahlen nicht möglich war (Postleitzahl nicht eingetragen etc.), was hauptsächlich in China und Indien der Fall ist, wurden die Städtenamen verwendet. Deswegen gibt es auch Adressen die nicht eindeutig einer, sondern mehreren NUTS(2 oder 3) Regionen zugeordnet wurden (OECD, 2009). Aber auch bei über Postleitzahlen zugeordneten Adressen tritt dieses „Problem“ auf, da eine Postleitzahlenregion nicht unbedingt in nur einer NUTS3 Region liegt. Bei Analysen kann man dann z.B. die

anteilmäßig größte NUTS3 Region als Herkunftsregion nehmen. Es werden nun drei zusätzliche Tabellen angelegt und die Daten aus REGPAT in sie eingelesen. Eine Tabelle stellt die Zuordnung NUTS Code der Region – Adresse des Erfinders her, eine die Zuordnung NUTS Code der Region – Adresse des Anmelders und eine die Zuordnung NUTS Code der Region – Name der NUTS Region.

Mit diesem Schritt wären wir nun fertig mit der Implementierung. Jetzt können wir die benötigten Informationen für unsere Analysen festlegen und die Abfrageprozeduren schreiben.

5. Nutzung

Nun da die Datenbank implementiert worden ist, kann man damit beginnen, die für die Analysen relevanten Kriterien für die Daten festzulegen und die Abfrageprozeduren zu programmieren. Mit ihnen werden die benötigten Informationen für die beiden Methoden, die Soziale Netzwerkanalyse und die Patentzitimierungsmethode, aus der Datenbank ausgelesen.

Für beide Verfahren gleich ist das Kriterium, dass im Rahmen dieser Studie der Automotive und IKT Bereich betrachtet werden. Den Automotive Bereich identifizieren wir in der Datenbank über die oben erwähnte Konkordanztabelle und nehmen die ISIC-Klasse 3843 (SCHERNGELL, 2007). Der IKT Bereich kann leider nicht eindeutig mit Hilfe dieser Tabelle einer oder mehreren ISIC-Klassen zugeordnet werden, deswegen erfolgt die Erfassung direkt über die IPC-Klassen. Aufgrund der Zuordnung jedes Patents zu einem technologischen Feld durch Vergabe eines IPC-Codes sind auch andere Analysen, z.B. detaillierte Branchenanalysen möglich. So kann beispielsweise die Inventionstätigkeit von Unternehmen eines bestimmten Sektors in bestimmten technologischen Feldern untersucht werden. Als nächstes Kriterium gilt es den Untersuchungszeitraum abzugrenzen. Die Abgrenzung erfolgt über das Attribut priority date der Patente. Das ist das Datum der Erstanmeldung des Patents. Das priority date ist zwar nicht das aktuellste Datum, das ein Patent enthält, und das Zulassungsdatum mag interessanter, da aktueller, aussehen, aber es reflektiert nicht das Datum der Erfindung. Dem kommt das priority date am nächsten. Deswegen eignet es sich am besten zur Erfassung des Zeitpunktes der Wissensproduktion (OECD, 2001). Die nächste Abgrenzung die es vorzunehmen gilt, ist die Zuordnung der Patente zu den Regionen die im Rahmen dieser Studie betrachtet werden. Da sowohl für den/die Erfinder als auch für den Patentanmelder die geographische Lokalisierung angegeben wird, sind prinzipiell zwei Möglichkeiten der Zuordnung von Patenten zu Raumeinheiten gegeben, nämlich einerseits auf Basis der NUTS Region des/der Erfinder/s, andererseits auf Basis der NUTS Region des Patentanmelders. Wir entscheiden uns für die Zuordnung der Patente auf Basis des Wohnsitzes des/der Erfinder/s. Diese ist einer Zuordnung auf Grundlage der Adresse des

Patentanmelders vorzuziehen, da Unternehmen ein Patent oft dem Firmensitz zuweisen, der jedoch weit vom Ort der Wissensproduktion entfernt sein kann. Mit der Zuordnung auf Basis des Wohnsitzes der Erfinder sollte eine möglichst genaue Erfassung des Ortes der Wissensproduktion erreicht werden (SCHERNGELL, 2007).

Nachdem diese Abgrenzungen vorgenommen wurden, gilt es nun, die für die Analysemethoden spezifischen Kriterien in die Abfrageprozeduren zu implementieren.

5.1. Soziale Netzwerkanalyse

Im Rahmen der sozialen Netzwerkanalyse haben wir uns für zwei gleich große Zeiträume entschieden, die betrachtet werden, und zwar von 1992 bis 1999 und von 2000 bis 2007.

Wir führen in jedem der Zeiträume Analysen jeweils für den Automotive und IKT Bereich in den festgelegten Regionen durch. Die Methodik der sozialen Netzwerkanalyse sieht eine Unterteilung der zu analysierenden Netzwerke vor, in Kooperations- und Mobilitätsnetzwerke (SCHERNGELL, 2007). Dadurch ergeben sich folgende Unterschiede in den jeweiligen Abfrageprozeduren:

Bei *Kooperationsnetzwerken* werden nur Patente mit mehr als einem Anmelder betrachtet

Bei *Mobilitätsnetzwerken* werden nur Patente mit genau einem Anmelder betrachtet

Als nächstes kommen wir zu einem Beispiel für eine Abfrageprozedur:

```
select                                                                                               distinct
  invt_name,applt_name,regpat_regions.name,regpat_invt.reg_code,applt_filing_date,ROUND(1/applt_share),ROUND(1/invt_share),regpat_applt.appln_id,regpat_applt.reg_share,regpat_invt.reg_share
from regpat_applt,regpat_invt,isic201_209,epo_ipc,regpat_regions
where prio_year between 2000 and 2007
and regpat_invt.reg_code in ('DEA18','DEA19','DEA1A')
and ISIC in(3843)
and applt_share > 0.99
and regpat_invt.appln_id=isic201_209.appln_id
and regpat_invt.appln_id=regpat_applt.appln_id
and regpat_invt.appln_id=epo_ipc.appln_id
and regpat_applt.reg_code=regpat_regions.code;
```

Sie gibt uns alle Patente, welche folgende Kriterien erfüllen, aus:

- *Priority date* im Untersuchungszeitraum (2000 bis 2007)
- Der Erfinder des Patentes kommt aus dem Bergischen Städtedreieck (NUTS3 Regionen: DEA18, DEA19, DEA1A)
- Bereich Automotive (ISIC-Klasse 3843)
- Genau ein Anmelder pro Patent, d.h. wir erhalten dann ein Mobilitätsnetzwerk

Wenn wir dann die Daten haben erfolgt eine Nachbearbeitung, die z.B. in Microsoft Excel vorgenommen werden kann. Dabei müssen erstens die Erfindernamen und Anmeldernamen bearbeitet werden, da in den Originaldaten nicht immer eine einheitliche Schreibweise gegeben ist. So wird z.B. der Name „Dr. Hans Becker“, manchmal aber „Dipl. Ing. Hans Becker“ verwendet, wobei jedoch ein und dieselbe Person gemeint ist. Zweitens wird dann eine Matrix mit der Dimension (Anzahl der Erfinder) x (Anmelder der Anmelder) erstellt, die Transponierte dieser Matrix berechnet, und dann das Produkt dieser beiden Matrizen. Damit wäre die Nachbearbeitung abgeschlossen, und man kann dann die berechnete Endmatrix in ein *Netzwerk-Visualisierungsprogramm* (z.B. UCInet) einlesen und mit Netzwerkanalysen fortfahren.

5.2. Patentzitimierungsmethode

Die Patentzitimierungsmethode besteht aus zwei verfahrenstechnischen Teilen: Zum einen wird die Anzahl der Patentpaare bestimmt, bei der die Patente, die andere Patente zitieren, einen Erfinder aus der betrachteten Region vorweisen müssen. Zum anderen bestimmt man die Anzahl der Patentpaare bei denen die Patente die zitiert werden einen Erfinder aus einer jener Regionen enthalten. Das Problem, dass sich bei diesem Verfahren ergibt, wird hervorgerufen durch eine grundlegende Eigenschaft von Patentzitimierungen: Eine Patentzitimierung kann erst erfasst werden, wenn ein zitierendes Patent angemeldet worden ist. Dieser Zeitraum, von der Anmeldung des zitierten bis zur Anmeldung eines zitierenden Patent, wird als Citation Lag bezeichnet. Daher vergeht auch eine gewisse Zeitspanne, bis ein Patent überhaupt eine vernünftige Zitierungsintensität erreicht (HALL/JAFFE/TRAJTENBERG, 2001). Das berücksichtigend (ausgehend von einem durchschnittlichen Citation Lag von 10 Jahren) wurden die Zeiträume für die beiden einzelnen Teile der Patentzitimierungsmethode folgendermaßen gewählt:

- Betrachtung der zitierenden Patente: 1997 bis 2007
- Betrachtung der zitierten Patente: 1987 bis 1997

Als nächstes kommen wir zu einem Beispiel für eine Abfrageprozedur:

```
select
tls212_citation.pat_publn_id,tls212_citation.cited_pat_publn_id,regpat_applt
.applt_name,regpat_applt.address,person_name,person_address,applt_seq_n
r,invnt_seq_nr,invnt_code.reg_code
```

```

from regpat_invt, regpat_applt, isic201_209, epo_ipc, tls206_person,
tls207_pers_appln, tls211_pat_publn, tls212_citation, cited211plus212,
invt_code
where prio_year between 1997 and 2007
and tls212_citation.cited_pat_publn_id <> 0
and pat_citn_seq_nr <> 0
and tls211_pat_publn.pat_publn_id = tls212_citation.pat_publn_id
and tls211_pat_publn.appln_id = regpat_invt.appln_id
and tls211_pat_publn.appln_id = regpat_applt.appln_id
and regpat_invt.appln_id = epo_ipc.appln_id
and regpat_invt.appln_id = isic201_209.appln_id
and          tls212_citation.cited_pat_publn_id          =
cited211plus212.cited_pat_publn_id
and cited211plus212.appln_id = tls207_pers_appln.appln_id
and tls206_person.person_id = tls207_pers_appln.person_id
and cited211plus212.appln_id = invt_code.appln_id
and regpat_invt.reg_code in ('DEA18','DEA19','DEA1A')
and ISIC in(3843);

```

Sie gibt uns alle Patentpaare, welche folgende Kriterien erfüllen, aus:

- Es werden zitierende Patente betrachtet
- Das legt auch den Untersuchungszeitraum fest (1997 bis 2007)
- Der Erfinder des zitierenden Patentes kommt aus dem Bergischen Städtedreieck (NUTS3 Regionen: DEA18, DEA19, DEA1A)
- Das zitierende Patent ist aus dem Bereich Automotive (ISIC-Klasse 3843)

Wenn wir dann die Daten haben, erfolgt (ähnlich zu der sozialen Netzwerkanalyse) eine Nachbearbeitung, die z.B. in Microsoft Excel vorgenommen werden kann. Es müssen die Anmeldernamen vereinheitlicht werden, da sie ja, wie aus dem vorherigen Kapitel bekannt, nicht einheitlich geschrieben sind. Dieses ist notwendig um sogenannte Selbstzitationen erkennen zu können. Das sind die Zitierungen bei denen der Anmelder von zitierendem und zitiertem Patent identisch ist. Sie stellen keine Wissensspillover dar, und müssen daher entfernt werden.

6. Fazit

Die hier vorgestellte Methodik ist in der Forschungspraxis durch eine Reihe von praktischen Problemen – wie exemplarisch geschildert – bisweilen erschwert. Dennoch kann man feststellen, dass die Auswertung umfangreicher Patentdatenbestände ein besonderes Wissen über die Innovationsdynamik von Unternehmen, Regionen und Ländern bzw. Sektoren ermöglicht. Die moderne Datenbankanalyse erlaubt es in hervorragender Weise, die nationalen und internationalen Wissensentstehungs- und Wissensdiffusionsprozesse abzubilden. Die Analyse von Kooperationsnetzwerken und Mobilitätsnetzen sowie die Nutzung von Patentzitierten zur Ausleuchtung etwa der intersektoralen Wissensverbindungen und der historischen Innovationsdynamik in der Industrie können von großem Nutzen für Wirtschaft und Wirtschaftspolitik sein. Auch die Finanzierungsakteure im Markt für Risikokapital können hier wesentliche Informationen entnehmen, die eine Eingrenzung von Risiken und eine Erhöhung der erwarteten Ertragsraten erlauben.

Literaturverzeichnis

- EPO (2008), EPO Worldwide Patent Statistical Database, European Patent Office.
- HALL, B.; JAFFE, A.; TRAJTENBERG, M. (2001), The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools, NBER Working Paper No. 8498.
- OECD (2001), Using Patent Counts for Cross-Country Comparisons of Technology Output.
- OECD (2009), The OECD REGPAT Database, EPO & PCT Patent Applications at Regional Level.
- SCHERNGELL, T. (2007), Interregionale Wissensspillovers in der europäischen High Tech Industrie, Deutscher Universitäts-Verlag
- VERSPAGEN, B.; VAN MOERGASTAL, T.; SLABBERS, M. (1994), MERIT Concordance Table: IPC - ISIC (rev. 2). Maastricht, MERIT.

EIIW Discussion Papers

ISSN 1430-5445:

Standing orders (usually 13 issues or more p.a.): academic rate 95 Euro p.a.; normal rate 250 Euro p.a.

Single orders: academic rate 10 Euro per copy; normal rate 20 Euro per copy.

Die Zusammenfassungen der Beiträge finden Sie im Internet unter:

The abstracts of the publications can be found in the internet under:

<http://www.eiiw.eu>

- No. 100 **Gavrilenkov, E.**: Macroeconomic Situation in Russia - Growth, Investment and Capital Flows, October 2002
- No. 101 **Agata, K.**: Internet, Economic Growth and Globalization, November 2002
- No. 102 **Blind, K.; Jungmittag, A.**: Ausländische Direktinvestitionen, Importe und Innovationen im Dienstleistungsgewerbe, February 2003
- No. 103 **Welfens, P.J.J.; Kirn, T.**: Mittelstandsentwicklung, BASEL-II-Kreditmarktprobleme und Kapitalmarktperspektiven, Juli 2003
- No. 104 **Standke, K.-H.**: The Impact of International Organisations on National Science and Technology Policy and on Good Governance, March 2003
- No. 105 **Welfens, P.J.J.**: Exchange Rate Dynamics and Structural Adjustment in Europe, May 2003
- No. 106 **Welfens, P.J.J.; Jungmittag, A.; Kauffmann, A.; Schumann, Ch.**: EU Eastern Enlargement and Structural Change: Specialization Patterns in Accession Countries and Economic Dynamics in the Single Market, May 2003
- No. 107 **Welfens, P.J.J.**: Überwindung der Wirtschaftskrise in der Eurozone: Stabilitäts-, Wachstums- und Strukturpolitik, September 2003
- No. 108 **Welfens, P.J.J.**: Risk Pricing, Investment and Prudential Supervision: A Critical Evaluation of Basel II Rules, September 2003
- No. 109 **Welfens, P.J.J.; Ponder, J.K.**: Digital EU Eastern Enlargement, October 2003
- No. 110 **Addison, J.T.; Teixeira, P.**: What Have We Learned About The Employment Effects of Severance Pay? Further Iterations of Lazear et al., October 2003
- No. 111 **Gavrilenkov, E.**: Diversification of the Russian Economy and Growth, October 2003
- No. 112 **Wiegert, R.**: Russia's Banking System, the Central Bank and the Exchange Rate Regime, November 2003
- No. 113 **Shi, S.**: China's Accession to WTO and its Impacts on Foreign Direct Investment, November 2003

- No. 114 **Welfens, P.J.J.:** The End of the Stability Pact: Arguments for a New Treaty, December 2003
- No. 115 **Addison, J.T.; Teixeira, P.:** The effect of worker representation on employment behaviour in Germany: another case of -2.5%, January 2004
- No. 116 **Borbèly, D.:** EU Export Specialization Patterns in Selected Accession Countries, March 2004
- No. 117 **Welfens, P.J.J.:** Auf dem Weg in eine europäische Informations- und Wissensgesellschaft: Probleme, Weichenstellungen, Politikoptionen, Januar 2004
- No. 118 **Markova, E.:** Liberalisation of Telecommunications in Russia, December 2003
- No. 119 **Welfens, P.J.J.; Markova, E.:** Private and Public Financing of Infrastructure: Theory, International Experience and Policy Implications for Russia, February 2004
- No. 120 **Welfens, P.J.J.:** EU Innovation Policy: Analysis and Critique, March 2004
- No. 121 **Jungmittag, A.; Welfens, P.J.J.:** Politikberatung und empirische Wirtschaftsforschung: Entwicklungen, Probleme, Optionen für mehr Rationalität in der Wirtschaftspolitik, März 2004
- No. 122 **Borbèly, D.:** Competition among Cohesion and Accession Countries: Comparative Analysis of Specialization within the EU Market, June 2004
- No. 123 **Welfens, P.J.J.:** Digitale Soziale Marktwirtschaft: Probleme und Reformoptionen im Kontext der Expansion der Informations- und Kommunikationstechnologie, Mai 2004
- No. 124 **Welfens, P.J.J.; Kauffmann, A.; Keim, M.:** Liberalization of Electricity Markets in Selected European Countries, July 2004
- No. 125 **Bartelmus, P.:** SEEA Revision: Accounting for Sustainability?, August 2004
- No. 126 **Welfens, P.J.J.; Borbèly, D.:** Exchange Rate Developments and Stock Market Dynamics in Transition Countries: Theory and Empirical Analysis, November 2004
- No. 127 **Welfens, P.J.J.:** Innovations in the Digital Economy: Promotion of R&D and Growth in Open Economies, January 2005
- No. 128 **Welfens, P.J.J.:** Savings, Investment and Growth: New Approaches for Macroeconomic Modelling, February 2005
- No. 129 **Pospiezna, P.:** The application of EU Common Trade Policy in new Memberstates after Enlargement – Consequences on Russia’s Trade with Poland, March 2005
- No. 130 **Pospiezna, P.; Welfens, P.J.J.:** Economic Opening up of Russia: Establishment of new EU-RF Trade Relations in View of EU Eastern Enlargement, April 2005
- No. 131 **Welfens, P.J.J.:** Significant Market Power in Telecommunications: Theoretical and Practical Aspects, May 2005
- No. 132 **Welfens, P.J.J.:** A Quasi-Cobb Douglas Production Function with Sectoral Progress: Theory and Application to the New Economy, May 2005
- No. 133 **Jungmittag, A.; Welfens, P.J.J.:** Institutions, Telecommunications Dynamics and Policy Challenges: Theory and Empirical Analysis for Germany, May 2005

- No. 134 **Libman, A.:** Russia's Integration into the World Economy: An Interjurisdictional Competition View, June 2005
- No. 135 **Feiguine, G.:** Beitritt Russlands zur WTO – Probleme und Perspektiven, September 2005
- No. 136 **Welfens, P.J.J.:** Rational Regulatory Policy for the Digital Economy: Theory and EU Policy Options, October 2005
- No. 137 **Welfens, P.J.J.:** Schattenregulierung in der Telekommunikationswirtschaft, November 2005
- No. 138 **Borbély, D.:** Determinants of Trade Specialization in the New EU Member States, November 2005
- No. 139 **Welfens, P.J.J.:** Interdependency of Real Exchange Rate, Trade, Innovation, Structural Change and Growth, December 2005
- No. 140 **Borbély D., Welfens, P.J.J.:** Structural Change, Innovation and Growth in the Context of EU Eastern Enlargement, January 2006
- No. 141 **Schumann, Ch.:** Financing Studies: Financial Support schemes for students in selected countries, January 2006
- No. 142 **Welfens, P.J.J.:** Digitale Innovationen, Neue Märkte und Telekomregulierung, März 2006
- No. 143 **Welfens, P.J.J.:** Information and Communication Technology: Dynamics, Integration and Economic Stability, July 2006
- No. 144 **Welfens, P.J.J.:** Grundlagen rationaler Transportpolitik bei Integration, August 2006
- No. 145 **Jungmittag, A.:** Technological Specialization as a driving Force of Production Specialization, October 2006
- No. 146 **Welfens, P.J.J.:** Rational Regulatory Policy for the Digital Economy: Theory and EU-Policy Options, October 2006
- No. 147 **Welfens, P.J.J.:** Internationalization of EU ICT Industries: The Case of SAP, December 2006
- No. 148 **Welfens, P.J.J.:** Marktwirtschaftliche Perspektiven der Energiepolitik in der EU: Ziele Probleme, Politikoptionen, Dezember 2006
- No. 149 **Vogelsang, M.:** Trade of IT Services in a Macroeconomic General Equilibrium Model, December 2006
- No. 150 **Cassel, D., Welfens, P.J.J.:** Regional Integration, Institutional Dynamics and International Competitiveness, Dezember 2006
- No. 151 **Welfens, P.J.J., Keim, M.:** Finanzmarktintegration und Wirtschaftsentwicklung im Kontext der EU-Osterweiterung, März 2007
- No. 152 **Kutlina, Z.:** Realwirtschaftliche und monetäre Entwicklungen im Transformationsprozess ausgewählter mittel- und osteuropäischer Länder, April 2007
- No. 153 **Welfens, P.J.J.; Borbély, D.:** Structural Change, Growth and Bazaar Effects in the Single EU Market, September 2008
- No. 154 **Feiguine, G.:** Die Beziehungen zwischen Russland und der EU nach der EU-Osterweiterung: Stand und Entwicklungsperspektiven, Oktober 2008

- No. 155 **Welfens, P.J.J.:** Ungelöste Probleme der Bankenaufsicht, Oktober 2008
- No. 156 **Addison J.T.:** The Performance Effects of Unions. Codetermination, and Employee Involvement: Comparing the United States and Germany (With an Addendum on the United Kingdom), November 2008
- No. 157 **Welfens, P.J.J.:** Portfoliomodell und langfristiges Wachstum: Neue Makroperspektiven, Portfolio Modelling and Growth, November 2008
- No. 158 **Welfens, P.J.J.:** Growth, Structural Dynamics and EU Integration in the Context of the Lisbon Agenda, November 2008
- No. 159 **Welfens, P.J.J.:** Portfolio Modelling and Growth, December 2008
- No. 160 **Islami, M.:** Interdependence Between Foreign Exchange Markets and Stock Markets in Selected European Countries, December 2008
- No. 161 **Welfens, P.J.J.:** Portfolio Modelling and Growth, January 2009
- No. 162 **Bartelmus, P.:** Sustainable Development – Has It Run Its Course?, January 2009
- No. 163 **Welfens, P.J.J.:** Intégration Européenne et Mondialisation: Défis, Débats, Options, February 2009
- No. 164 **Welfens, P.J.J.:** ЭКОНОМИЧЕСКИЙ РОСТ, ИННОВАЦИИ И ПРИРОДНЫЕ РЕСУРСЫ, February 2009
- No. 165 **Welfens, P.J.J.; Vogelsang, M.:** Regulierung und Innovationsdynamik in der EU-Telekommunikationswirtschaft, February 2009
- No. 166 **Welfens, P.J.J.:** The International Banking Crisis: Lessons and EU Reforms, February 2009
- No. 167 **Schröder, C.:** Financial System and Innovations: Determinants of Early Stage Venture Capital in Europe, March 2009
- No. 168 **Welfens, P.J.J.:** Marshall-Lerner Condition and Economic Globalization, April 2009
- No. 169 **Welfens, P.J.J.:** Explaining Oil Price Dynamics, May 2009
- No. 170 **Welfens, P.J.J.; Borbély, D.:** Structural Change, Innovation and Growth in the Single EU Market, August 2009
- No. 171 **Welfens, P.J.J.:** Innovationen und Transatlantische Bankenkrise: Eine ordnungspolitische Analyse, August 2009
- No. 172 **Erdem, D.; Meyer, K.:** Natural Gas Import Dynamics and Russia's Role in the Security of Germany's Supply Strategy, December 2009
- No. 173 **Welfens P.J.J; Perret K.J.:** Structural Change, Specialization and Growth in EU 25, January 2010
- No. 174 **Welfens, P.J.J.; Perret K.J.; Erdem D.:** Global Economic Sustainability Indicator: Analysis and Policy Options for the Copenhagen Process, February 2010
- No. 175 **Welfens, P.J.J.:** Rating, Kapitalmarktsignale und Risikomanagement: Reformansätze nach der Transatlantischen Bankenkrise, Februar 2010

EIIW Economic Policy Analysis:

- No. 1 **Welfens, P.J.J.:** Globalisierung der Wirtschaft und Krise des Sozialstaats: Ist die Wirtschaftswissenschaft am Ende?, April 1997
- No. 2 **Welfens, P.J.J.:** Nach der D-Mark kommt die E-Mark: Auf dem Weg zur EU-Währungsunion, Juli 1997
- No. 3 **Welfens, P.J.J.:** Beschäftigungsförderliche Steuerreform in Deutschland zum Euro-Start: Für eine wachstumsorientierte Doppelsteuerreform, Oktober 1998

Fordern Sie den EIIW Newsletter an: www.eiiw.eu

Please subscribe to EIIW Newsletter: www.eiiw.eu

Weitere Beiträge von Interesse: Titels of related interest:

Most recent books also see the last page.

VOGELSANG, M. (2010), Digitalization in Open Economies, Heidelberg: Springer.

WELFENS, P.J.J.; et al. (2009), A Europe of achievements in a Changing World, European Commission.

WELFENS, P.J.J.; BORBÉLY, D. (2009), Europäische Integration und Digitale Weltwirtschaft, Band 4: EU-Ostererweiterung, IKT und Strukturwandel, Stuttgart: Lucius & Lucius.

BLEISCHWITZ, R.; WELFENS, P.J.J.; ZHANG, Z. (2009), Sustainable Growth and Resource Productivity, Sheffield: Greanleaf.

WELFENS, P.J.J.; ADDISON, J.T. (2009), Innovation, Employment and Growth Policy Issues in the EU and the US, Heidelberg: Springer.

WELFENS, P.J.J.; RYAN, C.; CHIRATHIVAT, S.; KNIPPING, F. (2009), EU-ASEAN, Facing Economic Globalisation, Heidelberg: Springer.

WELFENS, P.J.J. (2009), Transatlantische Bankenkrise, Stuttgart: Lucius & Lucius.

WELFENS, P.J.J.; WOLF, H.C.; WOLTERS, J. (eds., 2008), International Economics and Economic Policy, Heidelberg: Springer.

WELFENS, P.J.J.; WALTHER-KLAUS, E. (eds., 2008), Digital Excelence, University Meets Economy, Heidelberg: Springer.

WELFENS, P.J.J. (2008), *Digital Integration, Growth and Rational Regulation*, Heidelberg: Springer.

WELFENS, P.J.J. (2007), *Innovation in Macroeconomics*, Heidelberg: Springer

WELFENS, P.J.J.; WESKE, M. (eds., 2007), *Digital Economic Dynamics, Innovations, Networks and Regulations*, Heidelberg: Springer.

WELFENS, P.J.J., WESKE, M. (eds., 2006): *Innovations, Digital Economic Dynamics and Regulatory Policy*, Heidelberg: Springer.

WELFENS, P.J.J., KNIPPING, F., CHIRATHIVAT, S., RYAN, C. (eds., 2006): *Integration in Asia and Europe: Historical Dynamics, Political Issues and Economic Perspectives*, Heidelberg: Springer.

BROADMAN, H.G., PAAS, T., WELFENS, P.J.J. (eds., 2006): *Economic Liberalization and Integration Policy Options for Eastern Europe and Russia*, Heidelberg: Springer

BORBÉLY, D. (2006): *Trade Specialization in the Enlarged European Union*, Heidelberg/Berlin: Springer.

JUNGMITTAG, A. (2006): *Internationale Innovationsdynamik, Spezialisierung und Wirtschaftswachstum in der EU*, Heidelberg: Physica.

WELFENS, P.J.J., WZIATEK-KUBIAK, (eds., 2005): *Structural Change and Exchange Rate Dynamics – The Economics of EU Eastern Enlargement*; Heidelberg: Springer.

WELFENS, P.J.J., ZOCHE, P., JUNGMITTAG, A. (et al. 2005): *Internetwirtschaft 2010* (final Report for the German Federal Government; joint study EIIW and Fraunhofer Institute for System Dynamics and Innovation, Karlsruhe), Heidelberg: Physica.

GRAHAM, E., ODING, N., WELFENS, P.J.J., (2005): *Internationalization and Economic Policy Reforms in Transition Countries*, Heidelberg: Springer.

GAVRILENKOW, E., WELFENS, P.J.J., (2005): *Infrastructure, Investments and Economic Integration: Perspectives for Eastern Europe and Russia*, Moscow: HSE.

APOLTE, T.; CASPERS, R.; WELFENS, P.J.J. (2004), *Ordnungsökonomische Grundlagen nationaler und internationaler Wirtschaftspolitik*, Stuttgart: Lucius & Lucius.

GAVRILENKOV, E.; WELFENS, P.J.J.; WIEGERT, R. (2004), *Economic Opening Up and Growth in Russia*, Heidelberg and New York: Springer.