Linking Patent Data with other Data Sources

Overview and practical applications
Other data sources?

IP data sources

- Bibliographical data
- Legal & procedural data
- Trade mark & design data
- Plant variety data
- Sequences of genes
Other data sources?

IP data sources
Other data sources?

Non-IP data sources

- Economics data
- Stock market data
- Standards data (DIN, IEEE)
- Scientific data / publications
- Business data
- Regional data
Other data sources?

Non-IP data sources
Other data sources?

Non-IP data sources

Social science
How do we link IP with non-IP data sources?
How do we build bridges?

**Attributes**
- Applicant names
- Inventor names
- Addresses
- Classification codes
- Patent numbers
- Non-patent literature
- (Key)words
- Trade marks
- Logos
- Authors
- Pictures

**Attributes**
- National accounts
- Industry sectors
- Stock symbol (ticker)
- SEPs - DIN...
- Researchers (authors)
- Universities
- Scientific publications
- (Key)words
- Authors
- Annual reports
- Regions (NUTS)
How do we build bridges?

- **Attributes**
  - Applicant names
  - Inventor names
  - Addresses
  - Classification codes
  - Patent numbers
  - Non-patent literature
  - (Key)words
  - Trade marks
  - Logos
  - Authors
  - Pictures

- **Common attributes**
  - SEAGATE
    - IEEE implantable body sensor networks
  - UNIV SINGAPORE
    - Bioelectrical signal
  - LT HUIHUI
    - About 300 IEEE NPL documents

- **Attributes**
  - National accounts
  - Industry sectors
  - Stock symbol (ticker)
  - SEPs - DIN...
  - Researchers (authors)
  - Universities
  - Scientific publications
  - (Key)words
  - Authors
  - Annual reports
  - Regions (NUTS)
Where do we stand now?

PATSTAT Data: open data

Existing
- OECD – triadic patents
- OECD – citations
- TM, DESIGN, PVR
- Companies
- Disclosed SEPs
- Eurostat NUTS codes
- Eurostat Nace codes

Future
- European Case Law Identifier (ECLI)
- OpenCorporates
Summarised

Linking data requires bridges

Bridges require common attributes
IPR-intensive industries and economic performance in the European Union

- Joint updated study by EUIPO and EPO
- How do industries that use IPR intensively contribute to the EU economy?
- The full report is available for download at www.euipo.europa.eu/ipcontribution
  www.epo.org/ipr-intensive-industries
Background of the study

- **Main objectives**
  - Identify IPR-intensive sectors in the EU
  - Quantify their contribution to employment, GDP, remuneration and trade

- **Broad scope of the study**
  - 28 EU member states
  - 6 different IP rights
General approach

Step 1
Match IP register with firm register

Step 2
Add firm/industry level business information

Step 3
Identify IPR-intensive industries and calculate economic contribution
Step 1: Combining data sets

PATSTAT
CPVO
TMView
Design View
ORBIS

Data cleaning and harmonization

IPR_owner
PK: owner_key
FK1: Harmonized name seat (NUTS3 region) legal form

Orbis_companies
PK: BvD_id
FK1: Harmonized name other relevant vars

Matched IPR-ORBIS table
PK: Harmonized_name
IPR variables
ORBIS variables

Final result
Set of concordance tables making it possible to link company financial and IPR data
Step 1: Name cleaning

- Matching data from so many sources requires a lot of data cleaning work.
- Disambiguation of multiple matches more problematic than finding matching records across multiple data sources.
Steps 1&2: Adding data from firms’ register (ORBIS)

ABC NET GmbH
ABC NET Gesellschaft mbH
78467 Konstanz
NACE code 7112
Engineering activities and related technical consultancy
Small company with 26 employees
Set up 16/03/1990
Subsidiary of UK based company manufacturing electronic components (semiconductors) listed on London Stock Exchange

Operating revenue (th Euro)

<table>
<thead>
<tr>
<th>Year</th>
<th>Patent</th>
<th>Trademark</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>25000</td>
<td>1</td>
</tr>
<tr>
<td>2006</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2011</td>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>10000</td>
<td></td>
</tr>
</tbody>
</table>
Results: IPR-intensive industries in the EU

- 342 industries are intensive in at least one IP right
- Most industries are intensive in more than one type of IPR

![Venn diagram showing the distribution of industries across different types of IPR rights: Trade mark (TM), Design, Patent (P), Copyright (C), Geographical indication (GI), Plant variety rights (PV).]
Results: Contribution of IPR-intensive industries to EU employment

IPR-intensive industries:
82 million jobs
38% of EU employment

28% (60 m.) directly generated
10% (22 m.) indirectly, by industries that supply goods and services to the IPR-intensive industries
Results: Contribution of IPR-intensive industries to EU GDP

IPR-intensive industries:
€ 5.7 trillion of value added
42% of EU GDP
Results: Contribution of IPR-intensive industries to remuneration – Wage premium
**Results:** Contribution of IPR-intensive industries to trade

- **Share of EU imports:** 86%
- **Share of EU exports:** 93%

**Trade surplus (in billion €)**
- TM: €14
- Design: €244
- Patent: €73
- Copyright: €18
- GI: €12
- PVR: €0
- All IPR: €97
Main requirements

1. Data
   → open data (but free?), open format

2. Bridges
   → many times through names → need for harmonisation and/or standardisation

3. Multidisciplinary expertise
   patents, economics, publications, IT, technological area → team?

Question: Will linked data offer usable solutions?
Need more information?

Ilja Rudyk, irudyk@epo.org
Geert Boedt, patstat@epo.org

epo.org/searching-for-patents

Follow us on

facebook.com/europeanpatentoffice   youtube.com/EPOfilms
twitter.com/EPOorg                  Linkedin.com/company/european-patent-office
Patents filed by EU applicants in climate change mitigation technologies (CCMT) at the EPO
Performance of CCMT-intensive industries in the EU

<table>
<thead>
<tr>
<th>Economic indicator</th>
<th>Share</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EU employment (direct)</td>
<td>1.2%</td>
<td>2.6 million</td>
</tr>
<tr>
<td>EU GDP</td>
<td>2.1%</td>
<td>282 billion €</td>
</tr>
<tr>
<td>EU wage premium (to non-IPR)</td>
<td>+ 90%</td>
<td>+ 463 €/week</td>
</tr>
<tr>
<td>EU trade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- % total EU imports</td>
<td>11.1%</td>
<td></td>
</tr>
<tr>
<td>- % total EU exports</td>
<td>17.4%</td>
<td></td>
</tr>
<tr>
<td>Trade surplus</td>
<td></td>
<td>+ 103 billion €</td>
</tr>
</tbody>
</table>
Methodology

- **Patents, Trade marks, Designs and PVRs**
  - IPRs $\rightarrow$ firms (Orbis) $\rightarrow$ sectors (NACE) $\rightarrow$ employment
  - *IPR-intensive* industry, if above-average number of IPRs per employee

- **Copyrights**
  - Based on WIPO classification

- **Geographic indications**
  - Based on DG AGRI (2012), at national level
Careful interpretation of results required

- No causal relationships between IPR and economic variables
- No value of IPR for firms or for industries
- Assumption of no variation in the use of IPR between firms within an industry
- Assumption of no variation in the industry use of IPR across countries: an industry is IPR intensive or not, regardless of where it is located