Semantic analysis of CSTP reports using the IPP vocabulary

OECD Directorate for Science, Technology and Innovation
Andrés Barreneche, Alina Deniau, Michael Keenan, Blandine Serve

OECD Library and Archives
Frédéric Abrazian, Mary-Ann Grosset, Jan-Anno Schuur, Thierry Vebr
This is the simplest of the approaches presented here, running the existing IPP vocabulary of 1200 terms against different corpora and presenting the results in simple bar charts and word clouds.

Analysis of the full CSTP corpus of substantive reports

782 full documents over 1994-2016
Includes a subset of 100 TIP substantive reports

Analysis of Research Policy articles

2527 title-abstracts over 1988-2017
50 Most frequent topics in CSTP Reports 1993 – 2017
50 Most frequent topics in ResPol 1988 – 2017

- higher education
- research
- technological development
- patent data
- diffusion
- inventor
- commercialisation
- biotechnology
- technological innovation
- innovation systems
- multinational corporation
- knowledge transfer
- innovation strategy
- knowledge flows
- telecommunication
- innovation impacts
- research projects
- firm strategies
- patent applications
- complementarity
- uncertainty
- complexity
- process innovation
- product innovation
- adoption
- engineering
- human capital
- scientific articles
- innovation policies
- research and development funding
- absorptive capacity
- innovation processes
- innovation survey
- product development
- innovation systems
- national innovation systems
- competitive advantage
CSTP: Topics becoming less frequent

Topic presence (% publications analysed)

- 1993-2000
- 2001-2008
- 2009-2017

- technological development
- biotechnology
- basic research
- diffusion
- technology diffusion
- intellectual property rights
- globalisation
- engineering
- agricultural level
- international innovation
- public sector research
- human resources in science
- brain circulation
- licensing
ResPol: Topics becoming less frequent

![Chart showing the topic presence over different years: diffusion, technological innovation, basic research, high-tech, applied research, technological trajectories, market failure, industrial structure, schumpeter, engineering, technological development. The x-axis represents the years 1988-1992, 1993-2000, 2001-2008, and 2009-2017. The y-axis represents the topic presence as a percentage of title-abstracts analysed. The chart indicates a decreasing trend in the presence of certain topics over time.]
CSTP: Topics becoming more frequent

Traditional core topics

Topic presence (% publications analysed)

- 1993-2000
- 2001-2008
- 2009-2017

- innovation systems
- scientific articles
- R&D funding
- PPPs
- energy
- R&D institutes
- venture capital
- scientific disciplines
- commercialisation
- patent applications
- open innovation
CSTP: Topics becoming more frequent

Emerging topics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>tax incentives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>social challenges</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>policy mix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>global challenges</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>open science</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>green growth</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>innovation evaluation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>innovation strategy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tertiary education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>climate change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>human capital</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>open access</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nanotechnology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>doctoral graduate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>government procurement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>business models</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ResPol: Topics becoming more frequent

![Graph showing the topic presence (%) of title-abstracts analysed over different periods: 1988-1992, 1993-2000, 2001-2008, and 2009-2017. The x-axis represents different topics such as innovation survey, uncertainty, inventor, knowledge transfer, absorptive capacity, innovation impacts, intellectual property rights, external knowledge spillovers, knowledge spillovers, human capital, externalities, R&D funding, open innovation, and performance of firms. The y-axis represents the topic presence in percentage (% title-abstracts analysed).]