



ANALYSING IP DATA TO SUPPORT EVIDENCE-BASED POLICY MAKING

EPO Patent Knowledge Week, Nov. 28 2023

Hélène Dernis, OECD Directorate for Science, Technology and Innovation



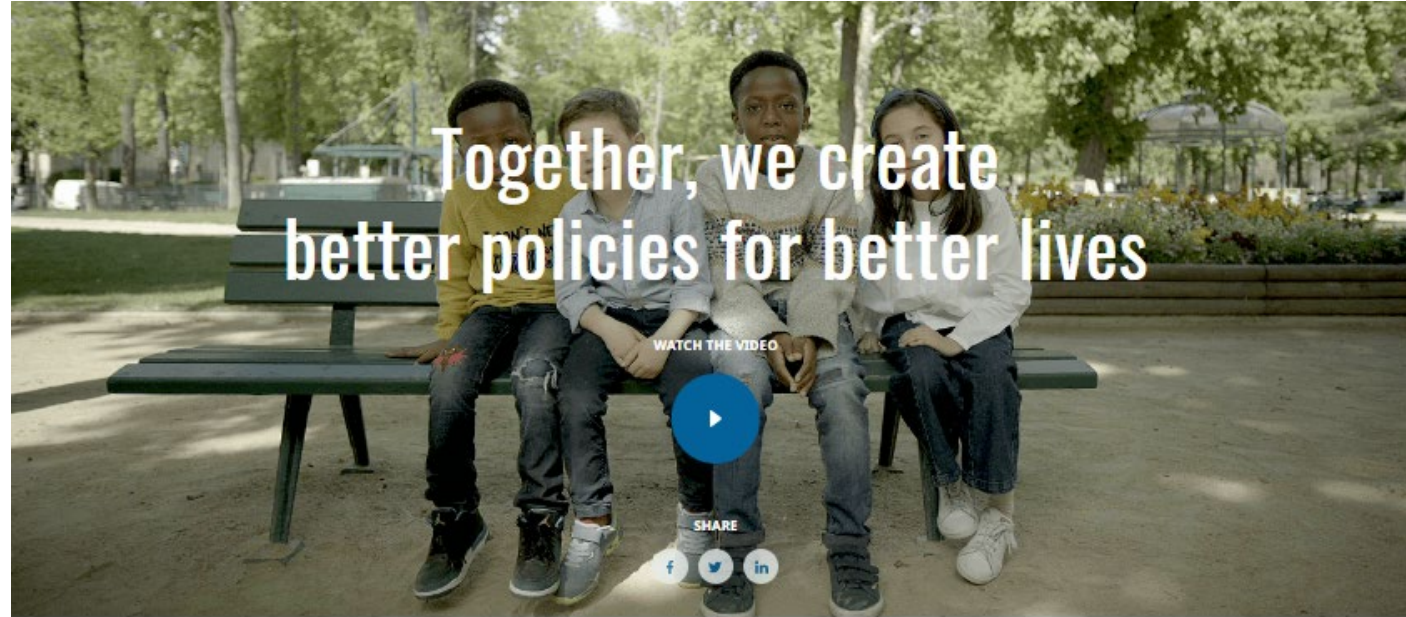
Organisation for Economic Co-operation and Development

Who we are

The Organisation for Economic Co-operation and Development (OECD) is an international organisation that works to **build better policies for better lives**.

Our goal is to shape policies that foster prosperity, equality, opportunity and well-being for all. We draw on **60 years of experience** and insights to better prepare the world of tomorrow.

Together with governments, policy makers and citizens, we work on establishing evidence-based international standards and finding solutions to a range of social, economic and environmental challenges. From improving economic performance and creating jobs to fostering strong education and fighting international tax evasion, we provide a **unique forum and knowledge hub for data and analysis, exchange of experiences, best-practice sharing, and advice on public policies and international standard-setting.**



www.oecd.org/about



OECD WORK ON IP DATA



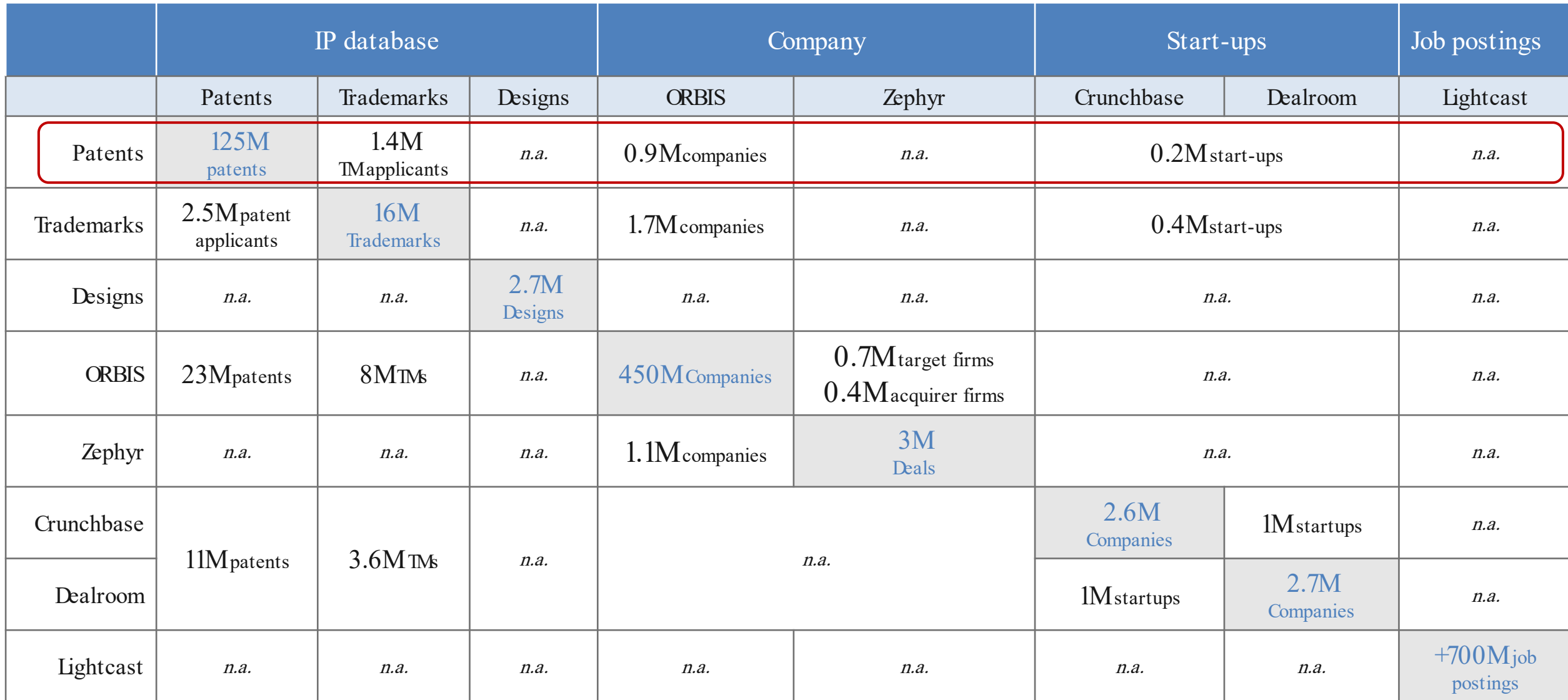
STI Micro-data Lab: An infrastructure for micro-level data

Collection of large scale micro-level datasets covering over 1 billion records

- Official administrative data (e.g. industrial property rights);
- Commercial data (e.g. firm-level data, start-ups, M&As, job postings)

Micro-data infrastructure to serve economic analysis related to innovation, entrepreneurship & productivity







IP database Patents, Trademarks, Design



Patent information

EPO PATSTAT Global (PATSTAT + EP Register),
Additional data for USPTO patents (PatentView)

Trademark & Design registrations

National/regional IP offices (EUIPO, JPO, USPTO)



Formatting
Curation





IP data curation and consolidation



- Name of applicants harmonised and **linked to business data**
- Name of inventors mapped to **gender**
- Addresses **regionalised**
- **Patents** allocated to different **technology fields**
- **Trademarks & design** decomposed in **goods and services**



Linkages
Taxonomies





Patent analysis What for?



Unique insights into the invention processes

Good proxy for technology-related & product innovations

Complement to existing data (R&D, innovation, trade, etc.).

Caveat: not all inventions are patented, skewed quality.

Patents indicators to investigate:

- *Technological performance ;*
- *Emerging technologies;*
- *Knowledge diffusion and dynamics of technical change;*
- *Geographical properties of inventive process;*
- *Innovation & firm dynamics;*
- *Economic value of inventions ;*
- *Role of universities in technological development ;*
- *Globalisation of R&D activities;*
- ...





Extensions of PATSTAT: OECD patent datasets available on-line

5 datasets available for download
<http://oe.cd/ipstats>

+ 500 requests in 2023

Triadic patent families

Dernis, H. and M. Khan (2004), "Triadic Patent Families Methodology", OECD Science, Technology and Industry Working Papers, No. 2004/02, OECD Publishing, Paris, <https://doi.org/10.1787/443844125004>.

REGPAT – patents by regions

Maraut, S., et al. (2008), "The OECD REGPAT Database: A Presentation", OECD Science, Technology and Industry Working Papers, No. 2008/02, OECD Publishing, Paris, <https://doi.org/10.1787/241437144144>.

Harmonised applicant names – HAN

Squicciarini, M. and H. Dernis (2013), "A Cross-Country Characterisation of the Patenting Behaviour of Firms based on Matched Firm and Patent Data", OECD Science, Technology and Industry Working Papers, No. 2013/05, OECD Publishing, Paris, <https://doi.org/10.1787/5k40gxd4vh41-en>.

Curated citation data

Webb, C., et al. (2005), "Analysing European and International Patent Citations: A Set of EPO Patent Database Building Blocks", OECD Science, Technology and Industry Working Papers, No. 2005/09, OECD Publishing, Paris, <https://doi.org/10.1787/883002633010>.

Patent quality indicators

Squicciarini, M., H. Dernis and C. Criscuolo (2013), "Measuring Patent Quality: Indicators of Technological and Economic Value", OECD Science, Technology and Industry Working Papers, No. 2013/03, OECD Publishing, Paris, <https://doi.org/10.1787/5k4522wkw1r8-en>.



IP bundle of world top R&D investors

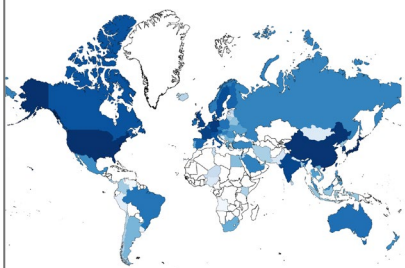
EC-JRC/ OECD COR&DIP database



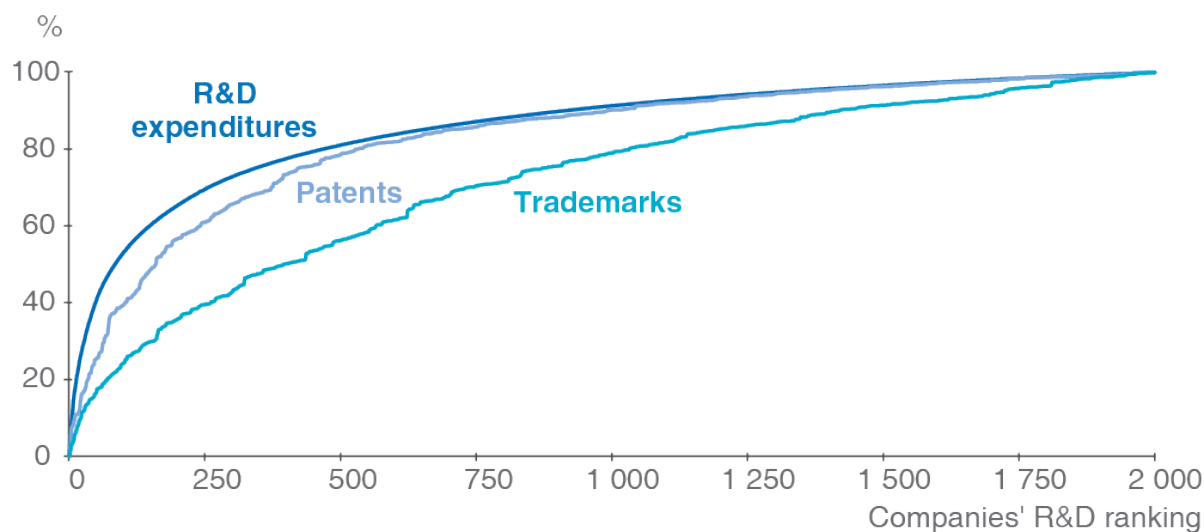
Top 2000 corporate R&D investors
Subsidiary structure linked to patent and trademark applicants
Latest version published in October 2023
Micro-level data available for download – oe.cd/ipstats

The EC-JRC-OECD COR&DIP© Database v.4 - 2023

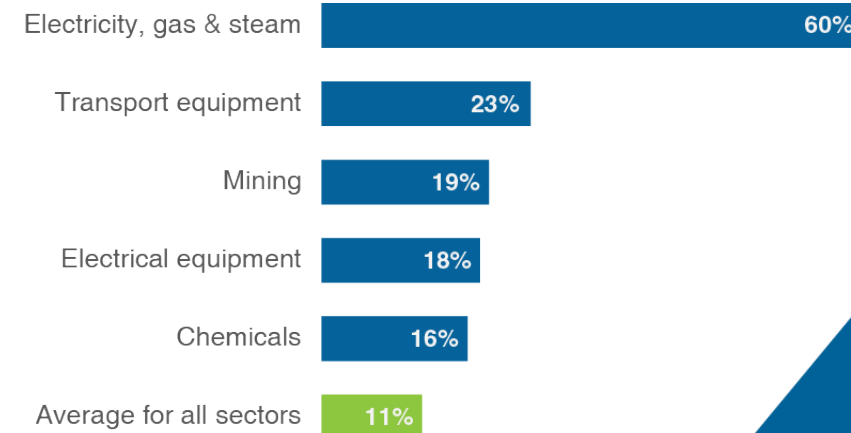
The Intellectual Property
Bundle of Top Corporate
R&D Investors



Cumulative distribution of R&D investments and the IP
bundle of the world's top R&D investors, 2018-20



Green patents in top R&D investors' portfolio, top five sector, 2018-20



* EC-JRC/OECD COR&DIP© v.4 data refer to the ranking of R&D investments for the year 2020, as reported in the 2021 EU industrial R&D investment scoreboard.

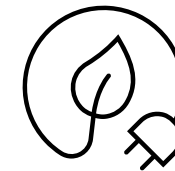


PATENTS TO UNCOVER INNOVATIVE ECO-SYSTEMS

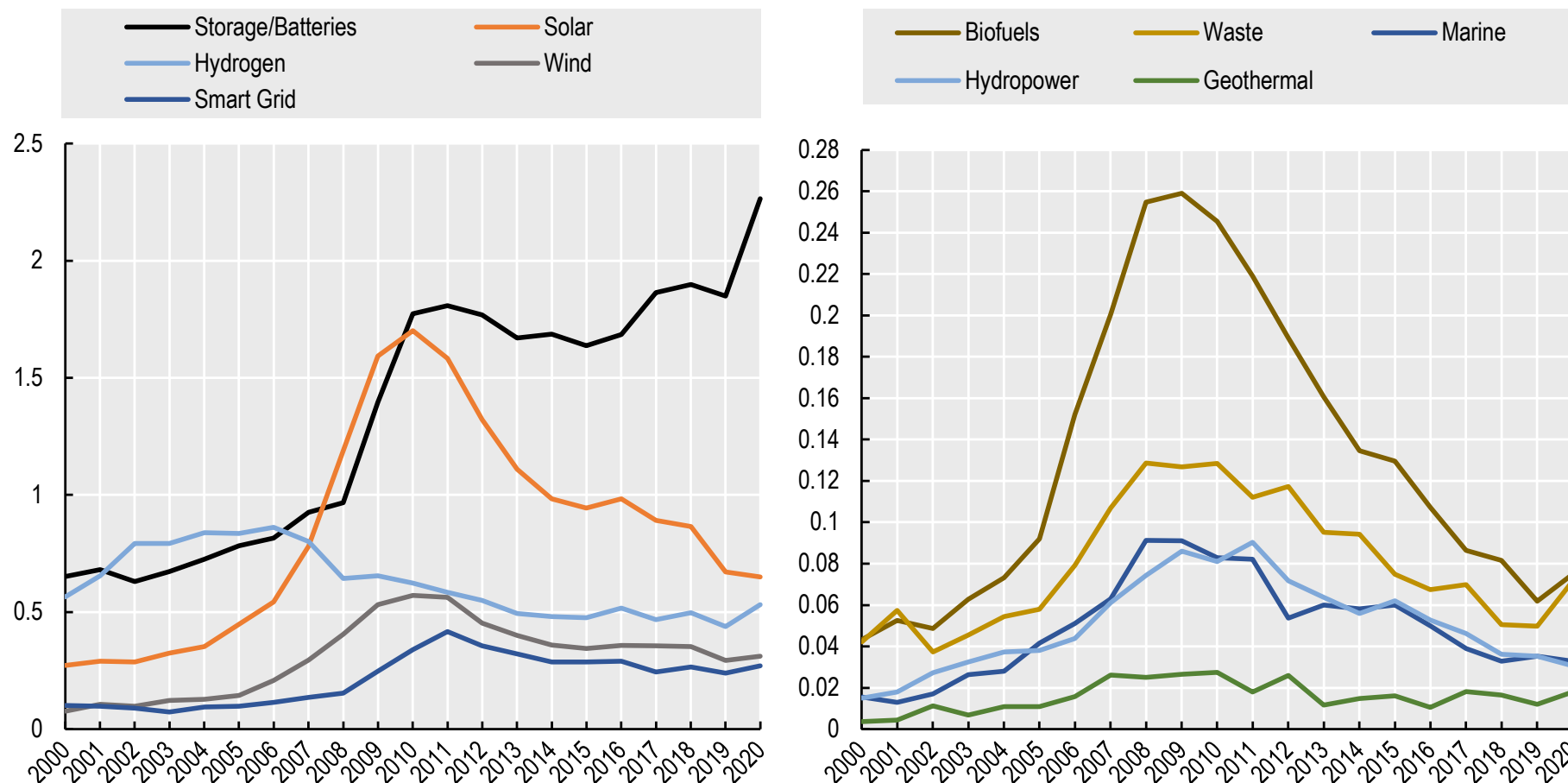




Patent-based measures to identify innovative eco-systems for renewable energy



Patents in renewable-energy related technologies, as a % of total patents



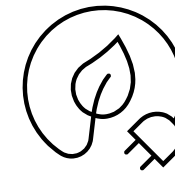
Patents in technologies related to renewable energy are identified using **IPC or CPC codes (Yo2 tag)**

Note: Data refer to families of patents filed in at least two patent offices worldwide, with at least one of the Five largest Intellectual Property (IP) offices worldwide, namely the EPO, JPO, KIPO, USPTO and CNIPA. Patent counts are provided according to the applicant country and earliest filing date, using fractional counts.

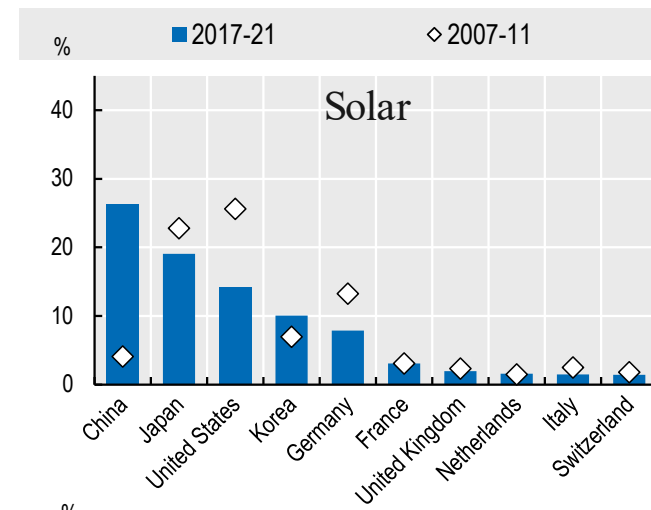
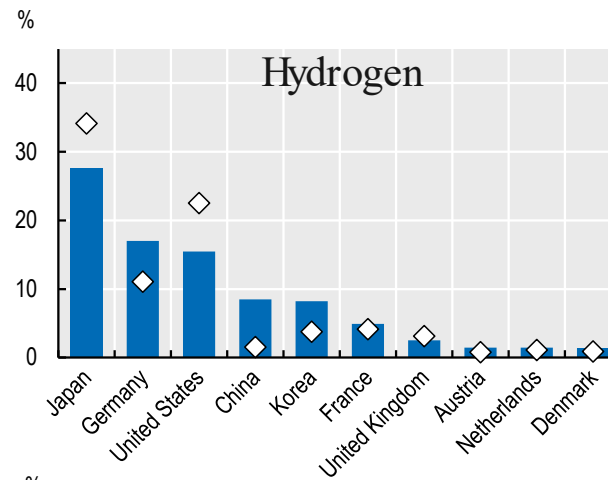
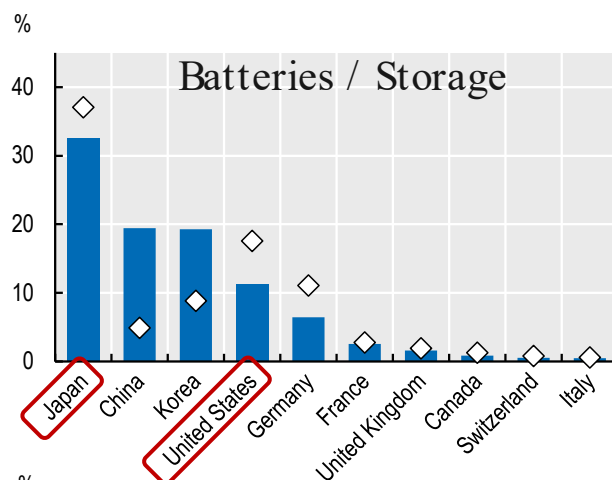
Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, November 2023.



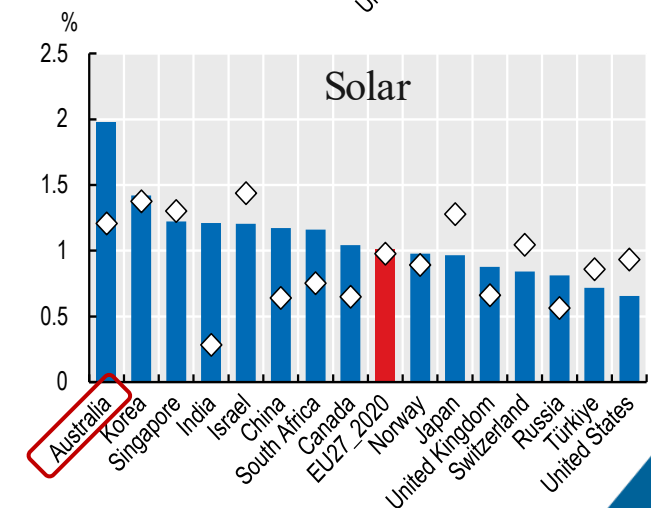
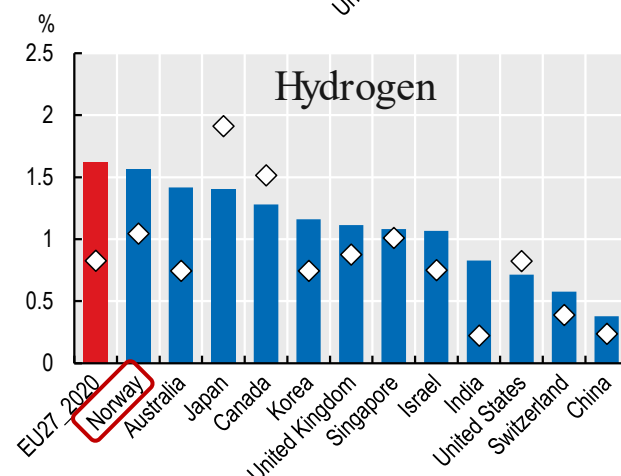
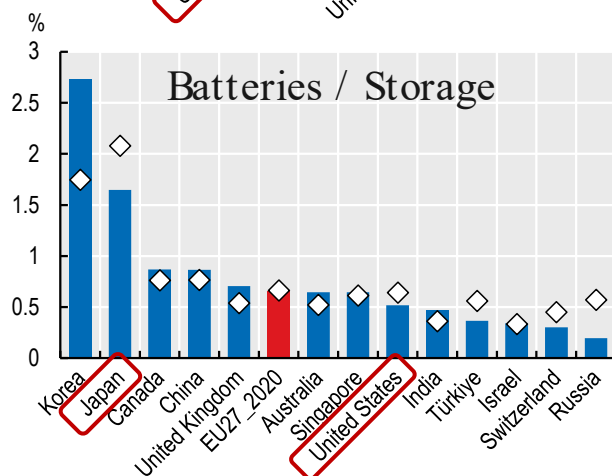
Top countries developing renewable technologies



Country's share
in total patents



Revealed
technology
advantage of
countries



Note: Data refer to PCT patents, by applicant country and earliest filing date, using fractional counts. For the revealed technology advantage, only economies with more than 1 000 patents in total and 10 patents in the given technology domain during each period considered are included.

Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, November 2023.



Top sectors for renewable energy technologies

- **Linked IP-company data** to characterise innovative eco-systems
- Combine patent data with business register data using the ORBIS© database (Bureau van Dijk / Moody's) using **name similarity**

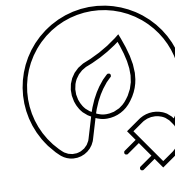


Note: Data refer to PCT patents, by applicant country and earliest filing date, using fractional counts. Patent applicant names were linked to the company names as provided in ORBIS© database. Only countries with matching rates above 60% are included in the figure, namely: Australia, Austria, Belgium, China, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Saudi Arabia, Sweden, Switzerland, United Kingdom and United States. The firm age is observed at the date of first filing of the patent and is calculated as the difference between the date of patent filing and the date of incorporation provided in ORBIS©.

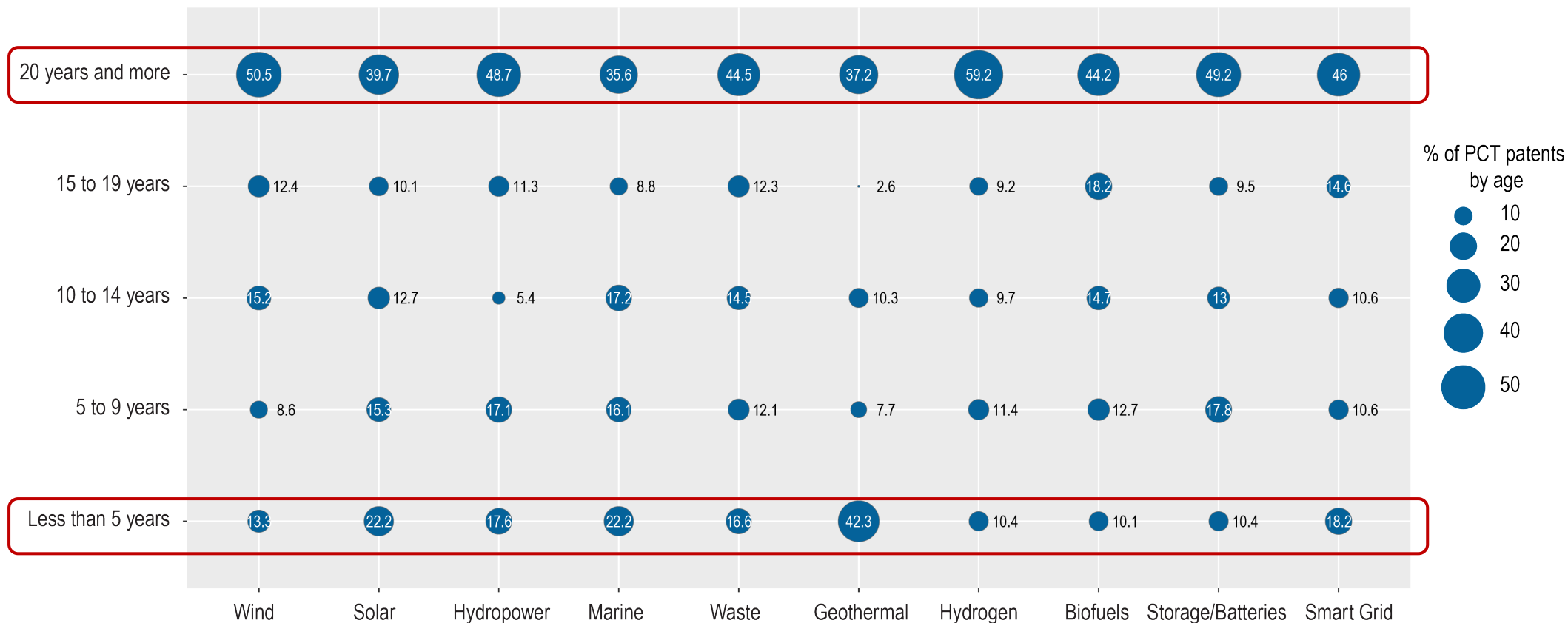
Sources: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, and ORBIS, version 2022.1, Bureau van Dijk, November 2023.



Companies developing renewable energy technologies by age



Renewable energy patents by firm age bands



Note: Data refer to PCT patents, by applicant country and earliest filing date, using fractional counts. Patent applicant names were linked to the company names as provided in ORBIS® database. Only countries with matching rates above 60% are included in the figure, namely: Australia, Austria, Belgium, China, Denmark, Finland, France, Germany, Iceland, Ireland, Italy, Japan, Luxembourg, Netherlands, Norway, Portugal, Saudi Arabia, Sweden, Switzerland, United Kingdom and United States. The firm age is observed at the date of first filing of the patent and is calculated as the difference between the date of patent filing and the date of incorporation provided in ORBIS®.

Sources: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, and ORBIS, version 2022.1, Bureau van Dijk, November 2023.



USE PATENT DATA TO UNCOVER WHAT'S AT THE CORE OF AI DEVELOPMENTS





Identify AI-related patents



OECD Definition for patents in AI-related technologies

- Classified in **IPC codes** closely related to AI
- Classified in **IPC codes** semi-related to AI and featuring in their English abstract or claims at least one **AI keyword**
- Classified in **CPC codes** semi-related to AI and featuring in their English abstract or claims at least one **AI keyword**
- Featuring **at least three AI keywords** in their English abstract or claims

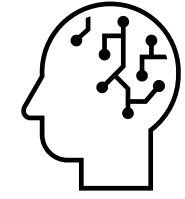


Baruffaldi, S., et al. (2020), "Identifying and measuring developments in artificial intelligence: Making the impossible possible", *OECD Science, Technology and Industry Working Papers*, No. 2020/05, OECD Publishing, Paris, <https://doi.org/10.1787/5f65ff7e-en> .

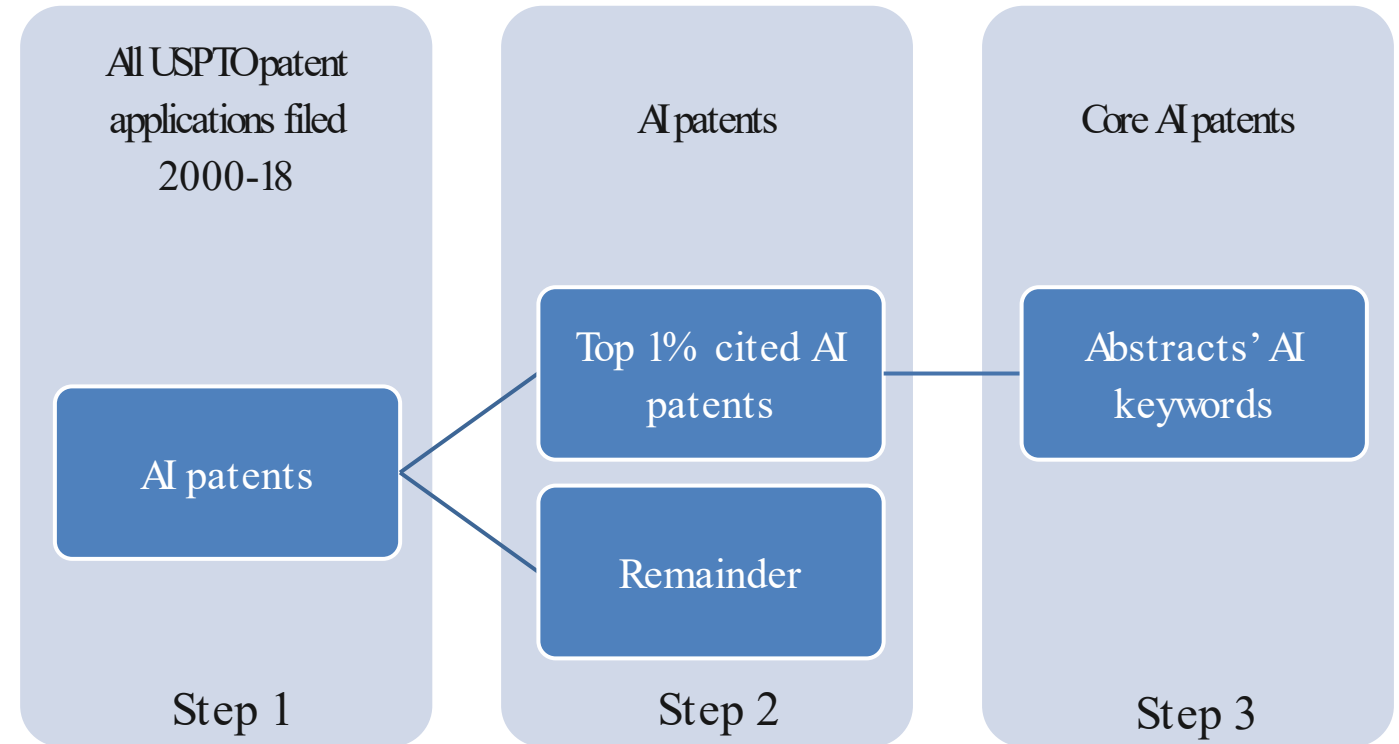


Dive into AI-related technologies

Uncover core AI patents



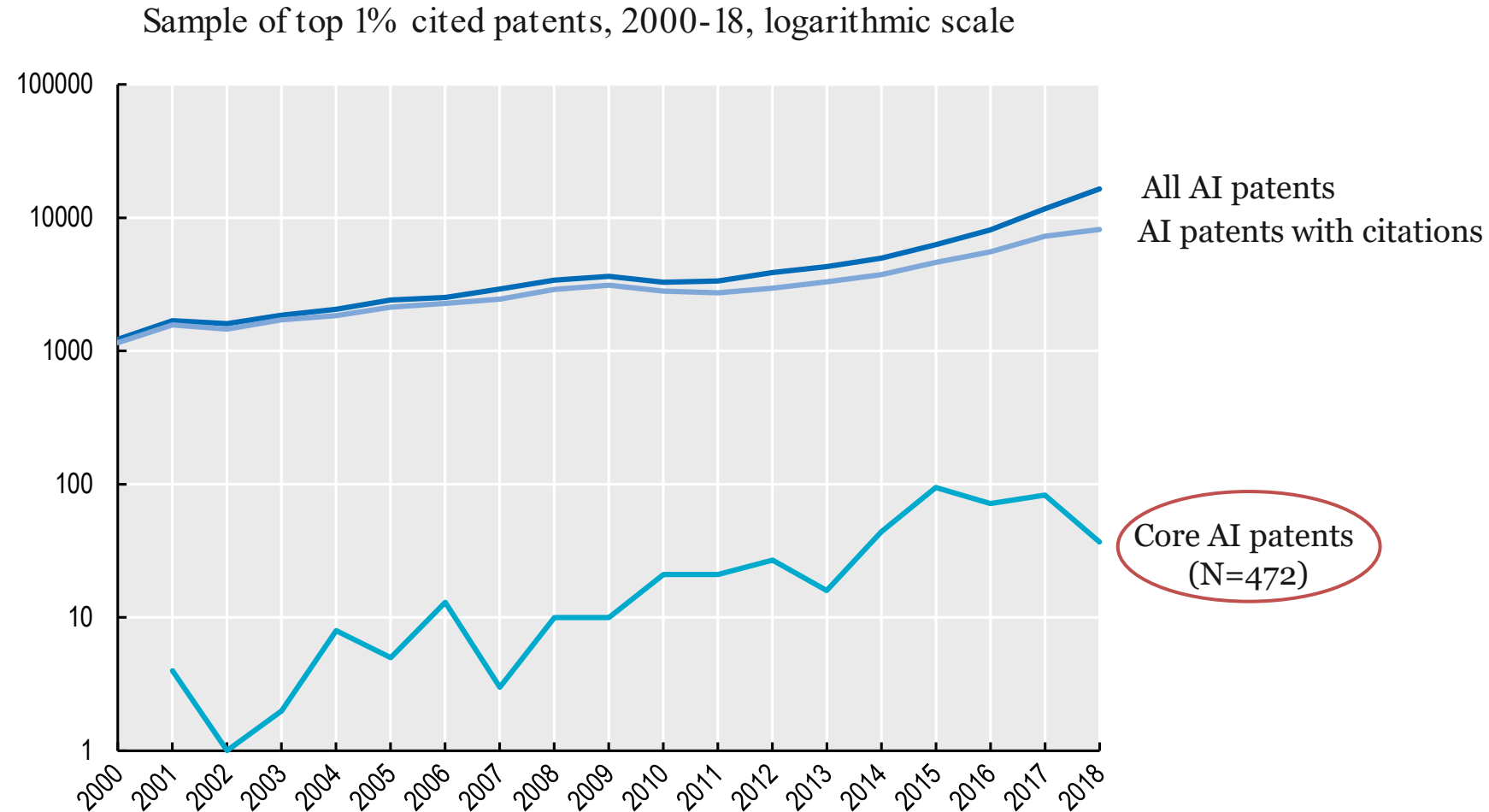
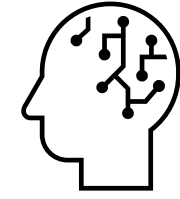
1. Identify AI-related patents based on **IPC/CPC codes closely related to AI and/or AI keyword(s)**
→ Baruffaldi et al. (2020)
2. Use **counts of AI-related forward citations** (normalised by the average number of forward citations received by patents from a reference cohort) and define **top 1% cited AI patents as “core AI”**
3. **Explore AI-related scientific and technological developments** in those “core AI” patents based on their abstracts’ AI keywords



Calvino, F., et al. (2023), "What technologies are at the core of AI? An exploration based on patent data", *OECD Artificial Intelligence Papers, No. 6*, OECD Publishing, Paris, <https://doi.org/10.1787/32406765-en>.



Acceleration of (core) AI patent activity

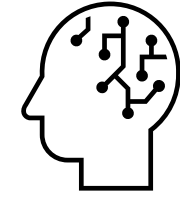


Note: Data refer to patent applications filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies. The sample of top 1% cited patents relies on normalised counts of forward citations.

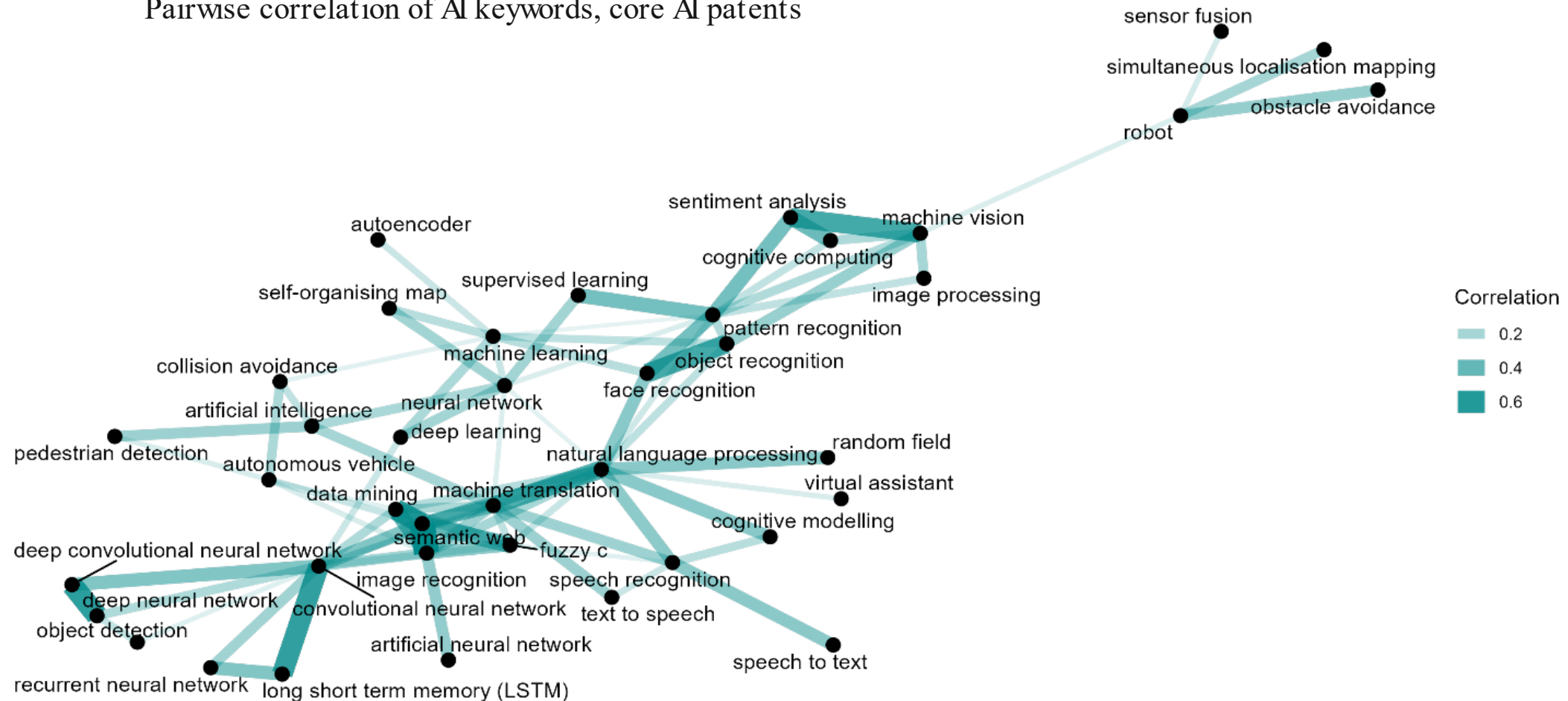
Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, June 2023.



Explore technologies at the core of AI



Pairwise correlation of AI keywords, core AI patents



Note: The figure shows the frequency with which AI keywords occur together rather than separately in an AI core patent. The sample refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies.

Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, June 2023.

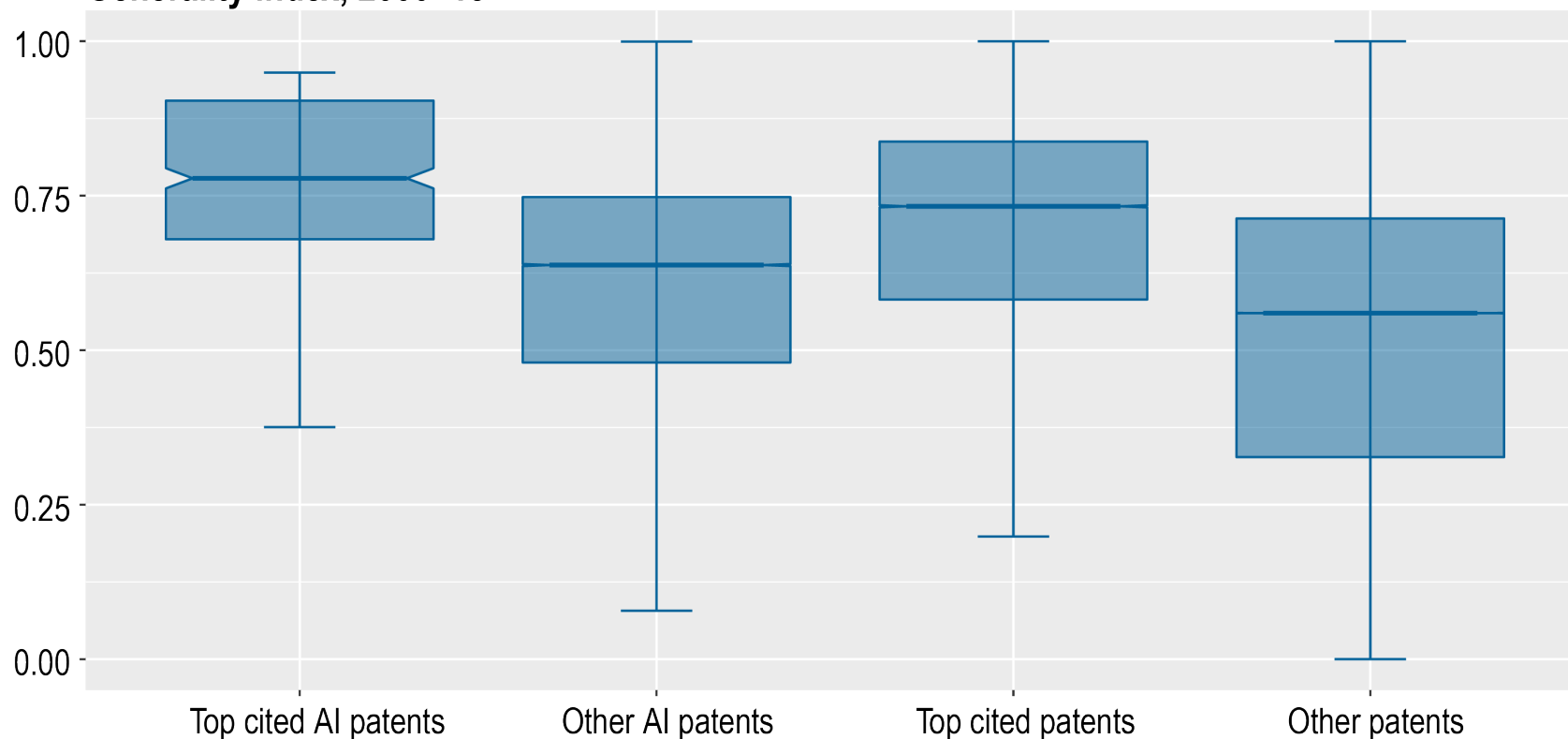


Core AI patents

Patent quality indicators



Generality index, 2000–18



- ...Supporting (core) AI's **GPT nature**.
- While overall patents became less general over time, core AI patents **became more general**.
- Core AI patents are also **broader in scope** and rely on a broader **range of technology fields**
- Although they used to be technologically more radical.

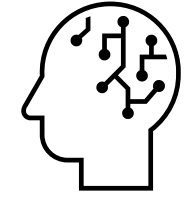
Note: The AI patent sample (up to 472 patents) refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications, excluding patent continuations) that refer to AI technologies. The generality index is constructed based on Squicciarini, Dernis and Criscuolo (2013).

Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, June 2023.



Text mining on patent documents

Explore topics at the core of AI



AI topics of top 1% cited patents, 2000-18

2000-09

Natural language processing
Robotics General AI
Autonomous driving
Networks (deep learning) Algorithms
Computer/image vision Speech
Feature engineering
Recognition/detection

2010-15

Recognition/detection
Natural language processing
Computer/image vision
Chatbot
Algorithms Networks (deep learning)
Autonomous driving
Feature engineering
General AI Robotics
Speech

2016-18

Recognition/detection
General AI
Computer/image vision
Speech Chatbot
Algorithms Robotics
Autonomous driving
Feature engineering
Natural language processing
Networks (deep learning)

- **General AI, robotics, autonomous driving** and **Speech** related technologies are **consistently** at the core of AI development...
- But more recently **deep learning** and **autonomous driving** gain significant **importance**.
- Technologies related to **speech, robotics** and **chatbot** appear **less prevalent**.

Note: The size of the words reflects the frequency with which each topic appears in the sample, which refers to the top 1% cited patent applications (normalised counts of forward citations) filed at the USPTO (patents issued and pre-grant publications) that refer to AI technologies. An AI core patent can have more multiple AI keywords and hence multiple topics. The analysis uses on a sub-sample of core AI patents with AI topics being imputed based on patents with similar IPC characteristics. Core AI patents without any topics (even after the imputation exercise) are dropped.

Source: OECD, STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>, June 2023.

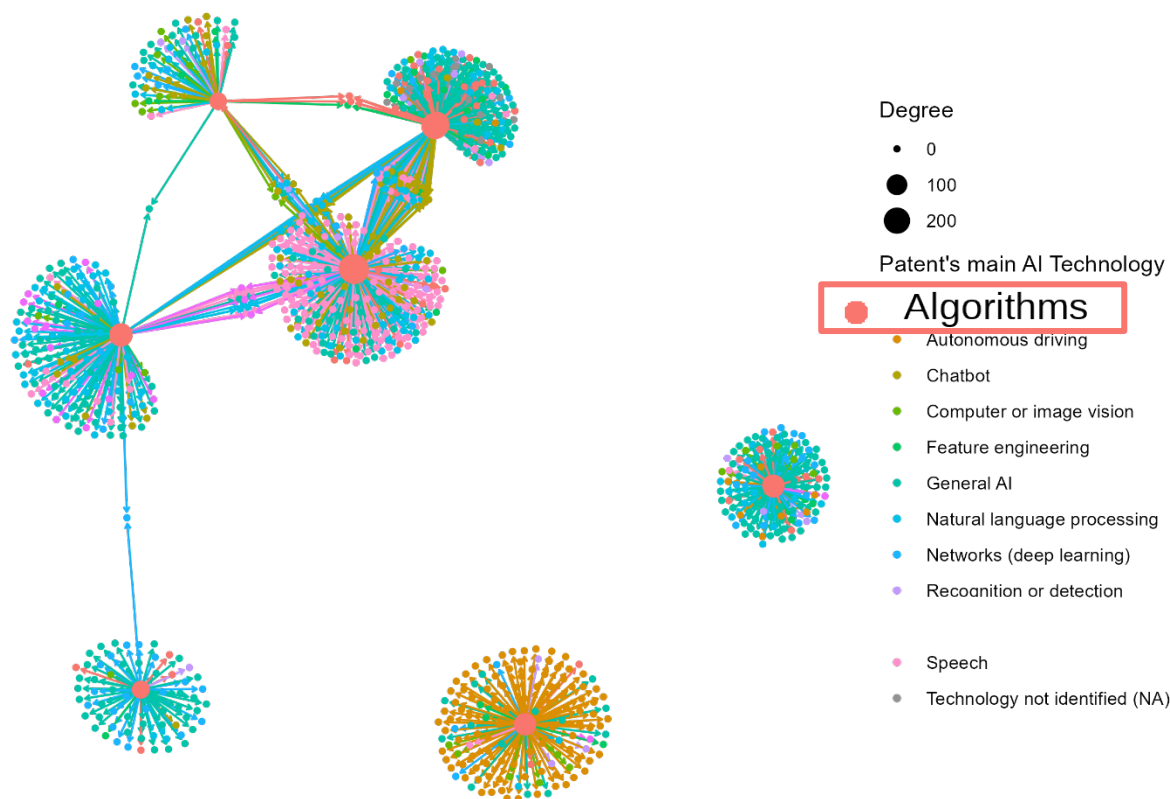


Citation linkages – Core vs. non-core AI patents

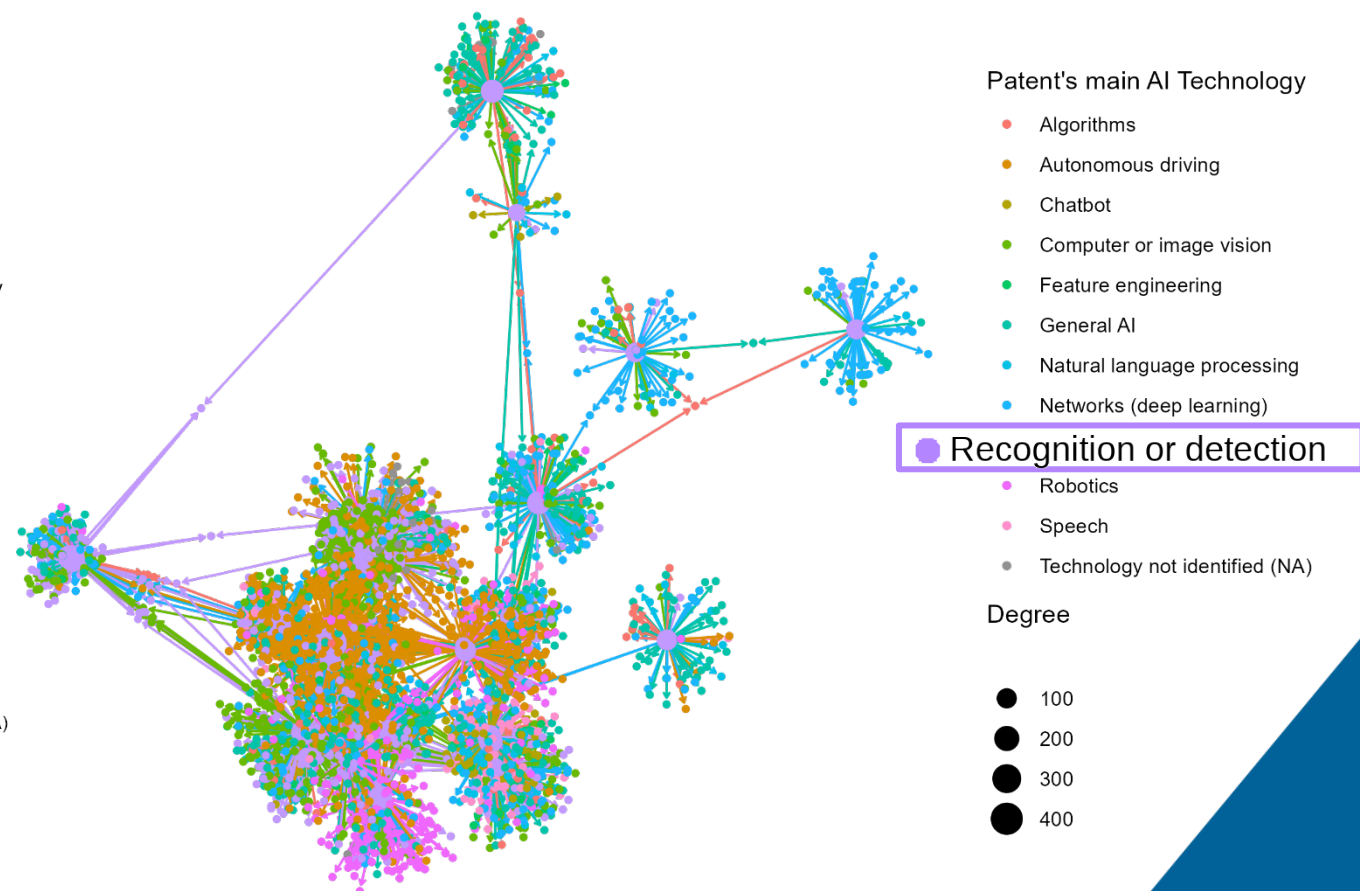
Technologies tend to develop from one another



Algorithms
core AI patents with AI citations, 2000-18



Recognition/ detection
core AI patents with AI citations, 2000-18





THANK YOU



Contact : helene.dernis@oecd.org

STI Micro-data Lab : sti.microdatalab@oecd.org

More information on OECD work on IP statistics : <http://oe.cd/ipstats>