

Methodological notes for patent counts

International
Energy Agency

iea

The PASTAT dataset presents patent statistics and indicators that are suitable for tracking innovation in clean energy technologies. The patent statistics presented here are constructed using data extracted from the OECD STI Micro-data Lab: Intellectual Property Database, <http://oe.cd/ipstats>. Only published applications for "patents of invention" are considered (*i.e.* excluding utility models, petty patents, etc.).

The relevant patent documents are identified based on the basis of a classification scheme (the [Y02 scheme](#)) developed by the European Patent Office for technical attributes of technologies that are related to climate change mitigation (CCMTs). The Y02 subclasses include areas related to specific clean energy technologies, namely Y02B (CCMTs related to buildings), Y02C (greenhouse gases- capture and storage/sequestration or disposal), Y02E (greenhouse gases - emissions reduction technologies related to energy generation, transmission or distribution), Y02T (CCMTs related to transport), etc. In addition, the ancillary Y04S scheme identifies CCMTs related to smart grids.

The categories were defined with the help of experts in the field, both from within the EPO and with the help of external partners, including the IEA. The inventories of the Y tags are updated by running search algorithms designed by expert examiners and relying on other CPC symbols, IPC symbols and keywords. Relevant documents are then 'tagged' in the EPO's [World Patent Statistical Database](#) making it easier to retrieve patent document that cover these technologies.

For the purposes of this database the IEA has mapped broad clean energy technology areas from the scheme (See table below with relevant Y02 classes).

Table 1. Mapping of technology fields with relevant Y02 classes

Technology Fields	Classes
Solar	Y02E 10/40-60
Wind	Y02E 10/70
Hydro	Y02E 10/20
Other Renewables	Y02E 10/10 & Y02E10/30
Bioenergy	Y02E 50/10-30



Technology Fields	Classes
Nuclear	Y02E 30
Grid	Y02E 40 & Y02E 60/60 & Y04S
Storage (not e-mobility)	Y02E 60/10-16 & Y02E 70/30
Hydrogen and Fuel Cells	Y02E 60/30-50 & Y02T 90/40
Carbon Capture and storage	Y02C
E-mobility	Y02T 10/60-72 & Y02T 90
Other Road Transport	Y02T 10/10-40 & Y02T 10/60 & Y02T 10/80-92
Air, Rail & Marine	Y02T 30-70
Building Energy Efficiency	Y02B except Y02B 10
Industry Energy Efficiency	Y02P except Y02P 60
Renewable Energy Integration in Buildings	Y02B 10
Agriculture Energy Efficiency	Y02P 60
Other Energy Efficiency	Y02E 20 & Y02D

In addition, to data on energy-related CCMTs, the database includes data on fossil fuel technologies. This has been based on the development of search strategies that are comparable in nature to the Y02 scheme (See [documentation here](#)). The searches are grouped into three broad areas: upstream technology fields; processing and downstream technology fields; and transmission and distribution technology fields. However, unlike the Y02 scheme, these documents are not tagged directly in PATSTAT, and as such algorithms have been developed independently. The relevant code can be found in a [public repository here](#).

Note that the fossil fuel technology searches are not meant to be exhaustive, focussing on the supply of energy. The large volume of patenting activity for fossil fuel end technologies such as engines would skew the sample heavily towards these activities. (While such improvements to end-use technologies improve the attractiveness of continued fossil fuel use, many could also support non-fossil fuel energy such as bioenergy or hydrogen.) In addition, technologies that are designed to improve devices for the combustion or other energy-related uses of fossil fuels, including reducing their greenhouse gas emissions, are not included



due to significant overlaps with energy efficiency technologies among the end-use technologies identified in the Y02 scheme. Given their different scope the absolute counts related to fossil fuel technologies and energy-related CCMTs should not be compared directly.

In both cases the counts presented are based on fractional counts by country of residence of the inventor(s); e.g. for a patent listing inventors from two different countries, each country will obtain a count of 0.5, to avoid double-counting of inventions. The year of invention is identified by the first filing date worldwide, under the Paris Convention since the priority date is considered to be closest to the actual date of invention. In addition, in this database we only include patents that have been protected in at least two intellectual property offices (commonly referred to as 'Claimed Priorities'). Previous work has indicated that this ensures that only relatively high-value patents are included in the counts (see [Hascic et al. 2015](#) for a discussion).

And finally, note that counts for aggregate technological domains are provided separately to avoid double-counting of inventions. For example, the count of "clean energy technologies" is less or equal to the sum of its sub-components. This is because patents are commonly classified in more than one technology class. Therefore each patented invention is counted only once when aggregating across technological domains. The same holds at the sub-domain level.

