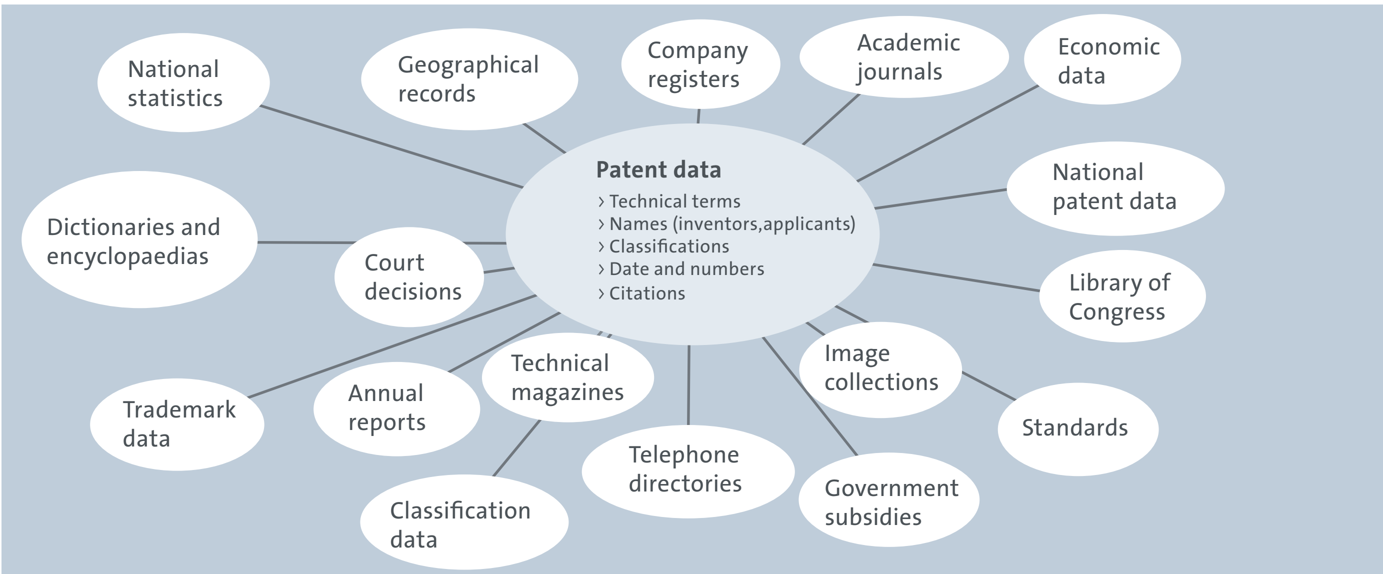
Linked open data from the European Patent Office

Linked open data from the European Patent Office (EPO) offers you new ways of combining patent data and non-patent data in your work.

Linked data is a standard way to represent data on a wide range of topics. Linked data sets have a simple and formalised structure, are machine-readable and are freely available on the internet.

Publishing linked data makes it easier for developers to connect information from different sources, resulting in new and innovative applications. The EPO publishes linked open data on European patents and applications; we call this "linked open EP data".

Patent information and other domains are interlinked



With linked open EP data you can

- retrieve up-to-date information on European patents and, for example, incorporate it into web pages
- connect to data from other patent data publishers
- download the data to create your own database
- easily combine the data with business or other non-patent information
- freely reuse the data on the basis of an open data licence.

What you get

The data set covers the most relevant data elements for patents, including:

- names of inventors and applicants
- dates and numbers
- titles in English
- technical classifications (IPC and CPC)
- citations

It also includes references to the full-text publications stored on the European Publication Server in HTML, PDF and XML. The data is updated weekly.

How it works

Linked open EP data can be queried, retrieved and viewed using standardised web technologies like HTTP, URI and SPARQL. SPARQL is the standardised query language for RDF, in the same way that SQL is the standardised query language for relational databases.

Linked open EP data uses uniform resource identifiers (URI) to identify patent applications, patent publications or other types of information. For occasional use, there is a simple data browser, an API, and a query interface.

EXAMPLE

Here is a SPARQL query which retrieves all applicants and their country as published in publication EP1430076 A2:

```
prefix patent: <http://data.epo.org/linked-data/def/patent/>
prefix vcard: <http://www.w3.org/2006/vcard/ns#>

SELECT ?ApplicantName ?ApplicantCountry  {
     <http://data.epo.org/linked-data/data/publication/EP/1430076/A2/->
          patent:applicantVC ?applicant .
     ?applicant vcard:fn ?ApplicantName;
          vcard:hasAddress/patent:countryCode ?ApplicantCountry .
} LIMIT 100
```

The result is a list of three companies and organisations

3 results in 0 min 0.241 s

Show 10 **entries**

ApplicantName	ApplicantCountry
"I.D.M. IMMUNO-DESIGNED MOLECULES"	"FR"
"INSTITUT GUSTAVE ROUSSY"	"FR"
"INSTITUT NATIONAL DE LA SANTE ET DE LARECHERCHE MEDICALE (INSERM)"	"FR"

Showing 1 to 3 of 3 entries

The example was based solely on data from the EPO's linked open EP data product.
It is also possible to combine external sources, for example:

- technical terms in the patent abstract can be linked to their definition in DBpedia (the linked data version of Wikipedia).
- countries or regions in the applicant or inventor information can be linked to the corresponding geographical entities of the GeoNames database, allowing retrieval of the geographical entities in several languages, the x/y coordinates of the polygon of its borders, etc.

GETTING STARTED

Linked open EP data is free,
so why not take a look and see
what you can do with it?

epo.org/linked-data

Published by European Patent Office, Vienna © EPO, 04-2019