

TRENDS IN OUR HIGH-QUALITY DATA COLLECTIONS

SOFIE LEPLAE | BULK DATA PRODUCTS AND INPADOC DATA OFFICER – PATENT KNOWLEDGE | 4 DECEMBER 2024

OVERVIEW

- Introduction: Data Collections
- Trends:
 1. Increase in data volume and data coverage
 2. Commitment to Data quality
 3. Open Access and Knowledge Sharing

“

BIG DATA IS THE SILENT STORYTELLER,
WHISPERING TRUTHS HIDDEN IN CHAOS
TO THOSE WHO DARE TO LISTEN AND
DECODE ITS PATTERNS.

ChatGPT

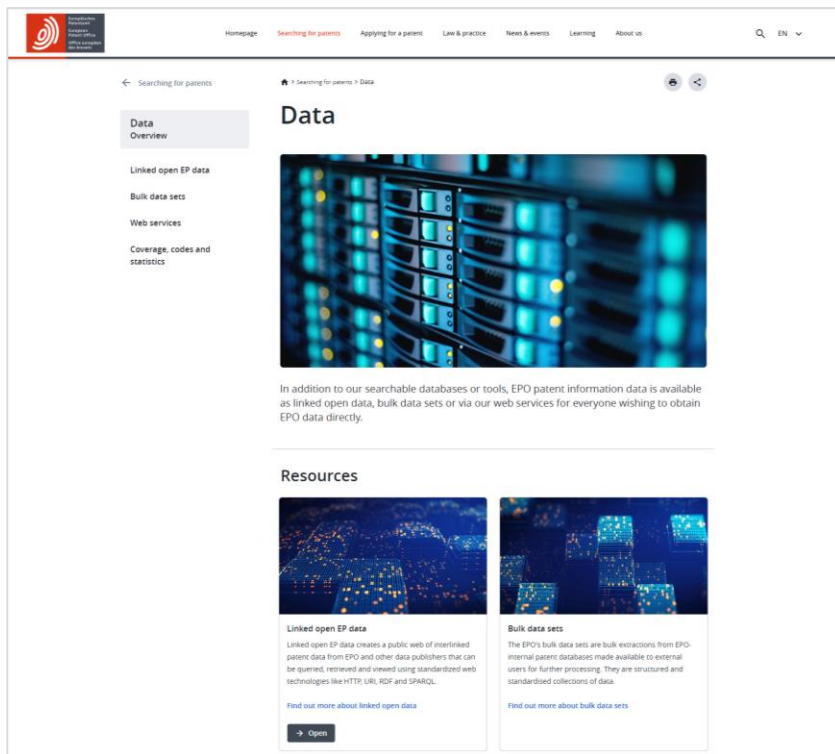
INTRODUCTION

Data Collections



PRIOR ART AND PROCEDURAL DATA COLLECTION

- 156 million patent publications
- Patent documents from 106 countries
- Procedural data: 470 million INPADOC legal events
- Procedural data: 128 million register events
- ...



The screenshot shows the EPO website's 'Data' section. The header includes the EPO logo and navigation links: Homepage, Searching for patents, Applying for a patent, Law & practice, News & events, Learning, and About us. A search bar with 'EN' is also present. The main content area is titled 'Data' and features a large image of server racks. Below the image, a paragraph states: 'In addition to our searchable databases or tools, EPO patent information data is available as linked open data, bulk data sets or via our web services for everyone wishing to obtain EPO data directly.' A 'Resources' section follows, containing two cards. The first card, 'Linked open EP data', describes a public web of interlinked patent data and includes a link to 'Find out more about linked open data' and an 'Open' button. The second card, 'Bulk data sets', describes bulk extractions from EPO internal databases and includes a link to 'Find out more about bulk data sets'.

Homepage Searching for patents Applying for a patent Law & practice News & events Learning About us

← Searching for patents Searching for patents > Data

Data

Linked open EP data

Bulk data sets

Web services

Coverage, codes and statistics

In addition to our searchable databases or tools, EPO patent information data is available as linked open data, bulk data sets or via our web services for everyone wishing to obtain EPO data directly.

Resources

Linked open EP data

Linked open EP data creates a public web of interlinked patent data from EPO and other data publishers that can be queried, retrieved and viewed using standardized web technologies like HTTP, URI, RDF and SPARQL.

[Find out more about linked open data](#)

[Open](#)

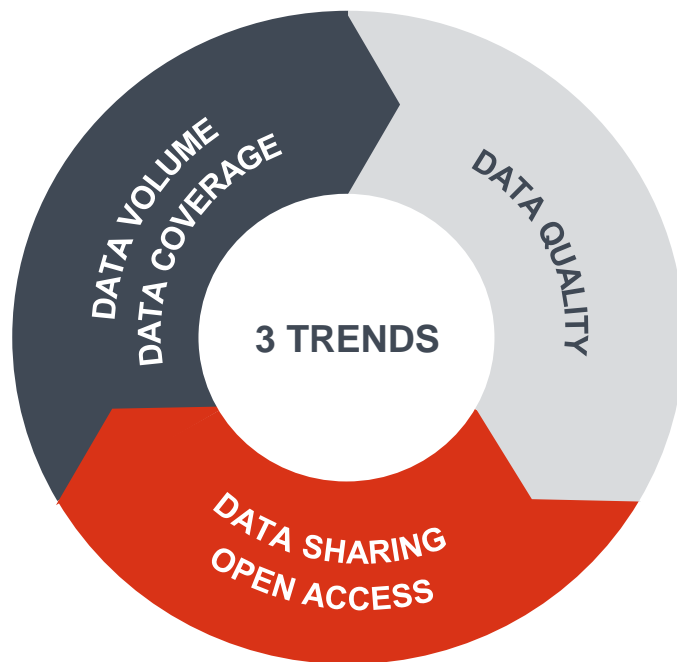
Bulk data sets

The EPO's bulk data sets are bulk extractions from EPO internal patent databases made available to external users for further processing. They are structured and standardized collections of data.

[Find out more about bulk data sets](#)

<https://www.epo.org/en/searching-for-patents/data>

TRENDS



1. INCREASE IN DATA VOLUME AND COVERAGE

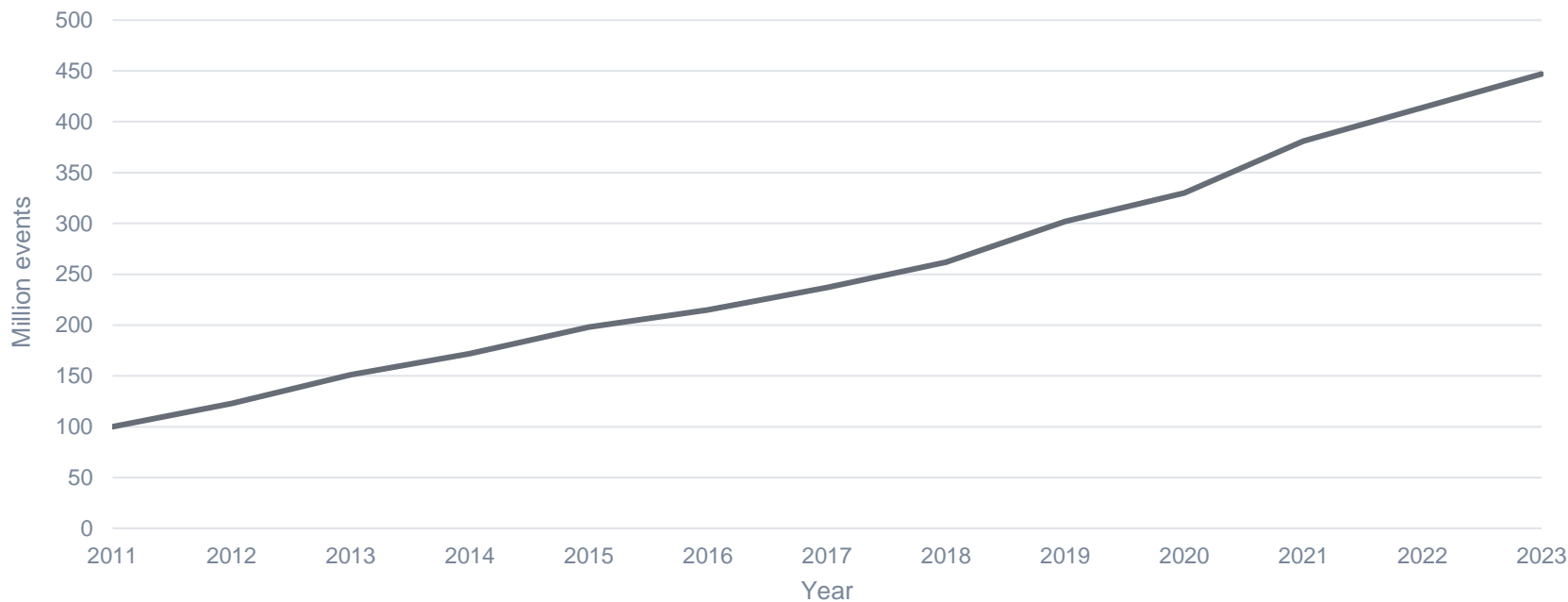


TREND: INCREASE IN DATA VOLUME

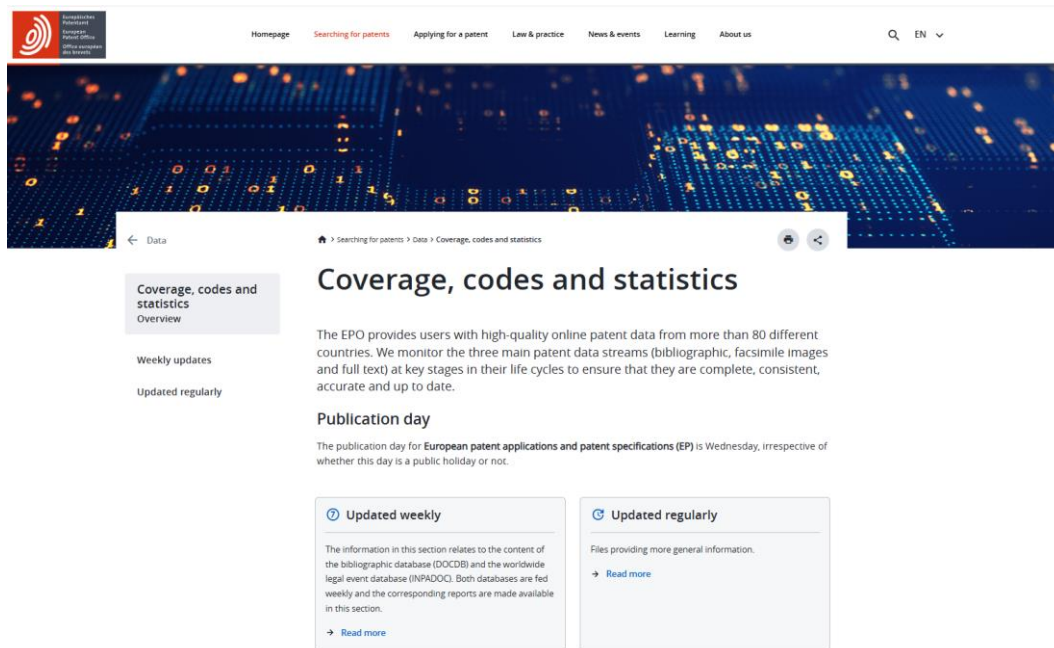
In 2023:

- increase by 7 million patent publications to 153 million
- increase by 4 million patent families
- classification of 880 000 documents into CPC
- increase by 48 million citations to 515 million citations
- increase by 33 million legal events to 447 million

INPADOC LEGAL EVENTS



TREND: INCREASE IN COVERAGE



Homepage Searching for patents Applying for a patent Law & practice News & events Learning About us

← Data

Searching for patents > Data > Coverage, codes and statistics

Coverage, codes and statistics

The EPO provides users with high-quality online patent data from more than 80 different countries. We monitor the three main patent data streams (bibliographic, facsimile images and full text) at key stages in their life cycles to ensure that they are complete, consistent, accurate and up to date.

Publication day

The publication day for **European patent applications** and **patent specifications (EP)** is Wednesday, irrespective of whether this day is a public holiday or not.

Updated weekly

The information in this section relates to the content of the bibliographic database (DOCDB) and the worldwide legal event database (INPADOC). Both databases are fed weekly and the corresponding reports are made available in this section.

→ [Read more](#)

Updated regularly

Files providing more general information.

→ [Read more](#)

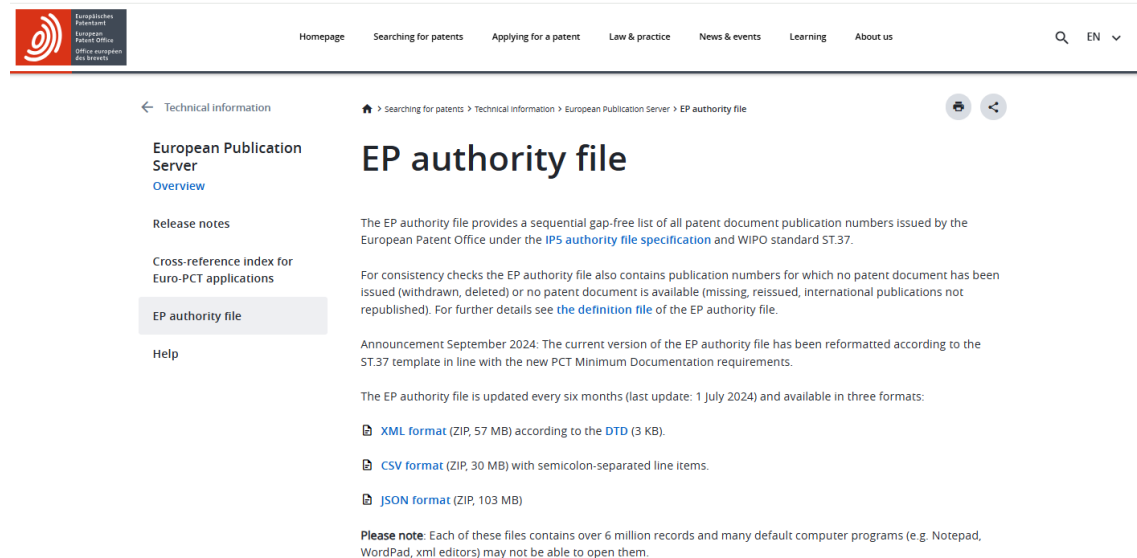
<https://www.epo.org/en/searching-for-patents/data/coverage>

2. COMMITMENT TO DATA QUALITY

Authority File – API led data exchange
– Business Rule Engine



AUTHORITY FILE



The screenshot shows the EPO website's 'EP authority file' page. The top navigation bar includes links for Homepage, Searching for patents, Applying for a patent, Law & practice, News & events, Learning, and About us. The left sidebar contains links for Technical information, European Publication Server (with an Overview link), Release notes, Cross-reference index for Euro-PCT applications, EP authority file (highlighted), and Help. The main content area features the title 'EP authority file' and a description: 'The EP authority file provides a sequential gap-free list of all patent document publication numbers issued by the European Patent Office under the [IPS authority file specification](#) and WIPO standard ST.37.' It also mentions that the file contains publication numbers for which no patent document has been issued (withdrawn, deleted) or no patent document is available (missing, reissued, international publications not republished). An announcement from September 2024 states that the current version of the EP authority file has been reformatted according to the ST.37 template. The file is updated every six months (last update: 1 July 2024) and is available in three formats: XML format (ZIP, 57 MB) according to the DTD (3 KB), CSV format (ZIP, 30 MB) with semicolon-separated line items, and JSON format (ZIP, 103 MB). A 'Please note' section at the bottom states that each of these files contains over 6 million records and many default computer programs (e.g. Notepad, WordPad, xml editors) may not be able to open them.

<https://www.epo.org/en/searching-for-patents/technical/publication-server/authority-file>

API LED DATA EXCHANGE

<https://www.epo.org/en/about-us/office/strategic-plan-2028>



Strategic Plan 2028
EUROPEAN PATENT OFFICE

STRENGTHEN EUROPEAN CO-OPERATION

Ever since the European Patent Organisation came into being, European co-operation has consistently delivered a more user-friendly patent system for users. Indeed, the preamble of the EPC highlighted

Furthermore, the catalogue of co-operation activities will contribute to new BCAs tailored to the specific needs of member states and their stakeholders.

■ A cloud-based Search system (based on ANSERA) that supports search and examination activities at NPOs and is in use by over 1 200 examiners in 28 countries;

■ Data exchange modules to enable the transmission of prior-art data from member states to the EPO and the provision of patent data, enabling EPO examiners to carry out national searches with different modules in use in eight countries already; and

also to define new initiatives and opportunities for cooperation in the interest of member states and users. It will also include pursuing technical and procedural co-operation, encompassing stronger IT co-operation, greater convergence of practice and search co-operation.

countries already, with a further six implementing actively and another four in the preparation phase; several additional countries have signalled their intention to proceed in the near future. So far, participating countries have received over 9 000 filings of different types;

with NPOs the possibility of creating a digital toolkit that covers all the functionalities of an integrated front and back office. This will enable all users in a particular country to use the same tools, whether they file nationally with the EPO or WIPO or a combination of these, and to offer a workflow system that manages the data from filing to

BUSINESS RULE ENGINE

<https://publication-bdds.apps.epo.org/raw-data/products/public/product/31>

A	B	C	D	E	
Publication Number Formats					
COUNTRY / TERRITORY	KC	NUMBER			
CC Description	DOCD	until 31-12-1999	as from 01-01-2000	Type of document	
B					
AL Albania	B	n(nnnnn)	n(nnnnn)	"Granted patent"	
	S1	n(nnnnn)	n(nnnnn)	"Registered design"	
	Y		n(nnnnn)	"Granted utility model"	
AM Armenia	A1	n(nnnnn)	n(nnnnn)	"Patent of invention"	
	A1		n(nnnnn)	"Description of invention protected by patent"	
	A1		n(nnnnn)	"Basic patent of invention"	
	A2	n(nnnnn)	n(nnnnn)	"Provisional patent of invention"	
	A2		n(nnnnn)	"Description of invention protected by provisional patent"	
	A2		n(nnnnn)	"Provisional patent of invention"	
	A		n(nnnnn)	"Patent of invention"	
	B1	n(nnnnn)	n(nnnnn)	"Patent granted on the basis of an official document issued by the former USSR or former Soviet Republic"	
	B1		n(nnnnn)	"Description of invention protected by a patent"	
	B2	n(nnnnn)	n(nnnnn)	"Patent of invention protected initially by a RA temporary patent"	
	B2		n(nnnnn)	"Description of invention protected by patent"	
	B2		n(nnnnn)	"Basic patent of invention initially protected by provisional patent"	
	B		n(nnnnn)	"Conclusion on patentability"	
	S	n(nnnnn)	n(nnnnn)	"Industrial design"	
	U	n(nnnnn)	n(nnnnn)	"Utility model"	
AP African Regional Industrial Property Organization	A	n(nnnnn)	n(nnnn)	"Patent"	
	A0	yyynnnn	yyynnnn	"Application filed, as announced in the Gazette published by this office"	

3. OPEN ACCESS AND KNOWLEDGE SHARING

Maximise impact






OPEN ACCESS: FREE PRODUCTS

Home > Basket

Basket

Continue shopping
Check out

3 items | 0,00 €

Item	Amount	Qty	Delivery	Total
 14.6 Decisions of the EPO boards of appeal 2024 CLD Edition: 2024 Type: Frontfile Period: Annual (1 January - 31 December)	0,00 €	1	Standard	0,00 € ✕
 EP full-text data 2024 ABDWL Edition: 2024 Type: Frontfile Period: Annual (1 January - 31 December)	0,00 €	1	Standard	0,00 € ✕
 14.1 EBD - EP bibliographic data 2024 PRDWL Edition: 2024 Type: Frontfile Period: Annual (1 January - 31 December)	0,00 €	1	Standard	0,00 € ✕

Coupon code

Apply

Subtotal: 0,00 €
Total amount 0,00 €

Continue shopping
Check out

<https://shop.epo.org/en/login>

KNOWLEDGE SHARING: ONLINE SEMINARS

[INPADOC worldwide legal event data for advanced users \(October 2024\)](#)

[INPADOC worldwide legal event data for beginners \(September 2024\)](#)

[INPADOC worldwide legal event data for advanced users \(April 2024\)](#)

[INPADOC Legal event data and the unitary effect \(March 2024\)](#)

[EPO Patent information tools](#)

KNOWLEDGE SHARING: PATENT KNOWLEDGE NEWS

[Searching and retrieving Unitary Patent information](#)

(Focus on INPADOC – 29.01.2024)

[Patent protection: need more time?](#) (Supplementary protection certificates (SPCs) and how to find them in Espacenet 12.02.2024)

[Protection beyond 20 years: data on SPCs and other term extensions for pharmaceutical patents](#) (Part I: Basic concepts at a glance - 22.04.2024)

[Searching and understanding data on Supplementary Protection Certificates \(SPCs\)](#) (Part II: SPC bibliographic data in the EPO's databases - 21.05.2024)

[Searching and understanding data on supplementary protection certificates \(SPCs\)](#) (Part III: SPCs in legal event data - 01.07.2024)

[Case number for UPC opt-outs in INPADOC](#) (21.10.2024)

[Patent Knowledge News](#)

CONTACT US FOR MORE INFORMATION

1 Enquiry

Contact us

This contact form has no legal validity in proceedings under the EPC or PCT. [Read more](#)

Topic •

Searching for patents / patent knowledge / data products

Subject •

Access, Subscriptions and Prices

Enquiry number ?

Your reference

Your enquiry •

Select file ?

Your preferred language for reply •

☐ English ☐ French ☐ German

2 Subject •

Access, Subscriptions and Prices

Please select

Espacenet

Patent Information Centres (PATLIB network)

Patent information / patent knowledge in general

Patent Knowledge News

Patent Translate

Data & services

Access, Subscriptions and Prices

Asian / worldwide patent information

Bulk data sets (DOCDB / INPADOC / full-text data, etc.)

EP Bulletin search

EP full-text search

European Patent Bulletin

European Patent Register / Federated Register / Register Alert

European Publication Server

Global Dossier

Global Patent Index

Linked open EP data

Open Patent Services (OPS) / web services / API

Patent analytics / patent insight reports

Please contact us via our [contact form](#).

You may select the **subject** that best fits your inquiry

You can **upload** supporting information.