



CnOpenData

User Manual

Version 2021

CnOpenData Our Team



Hello! We are a group of IT guys, "tech geeks" is our label, but we love this kind of joyful working mode of writing code and capturing data.

By chance, we have established contact with some university professors to provide them with data collection and collation services. Over the past few years, we have established stable and joyful relationships with a number of university professors, which surprisingly sparked us with a certain influence in a small field, and more and more scholars started to contact us for data services. As our clients increased, we had the idea of establishing a data platform.

The Internet generates massive amounts of data, records a wealth of information, and presents it in an open format. However, due to **the huge amount of data, the extreme difficulty of data collection, and the time-consuming cleaning and integration**, scholars face many obstacles in using data, which is exactly what we IT guys are best at. Of course, we are backed by a strong academic team, who were once our clients and are now our advisors!

Simplicity, clarity and usefulness are the goals we pursue.

Data is scattered gold, we are willing to be the miner, and look forward to cooperating with you like the gold miner!

CnOpenData Team
2019-5-1



Patent Data

1. Patent Innovation Data

- Provides data on patent applications and patent grants from 1985 to 2020, as well as data on patent status information.
- Patent Data are divided into published inventions, granted inventions, utility models and designs.
- Transaction data for each patent is recorded in detail.
- **Fields:** application publication date, PTC publication and application data, applicant, application date, etc.

2. Chinese Patent Citation Data (Chinese Edition, World Edition)

- Provides detailed citation information in China and abroad for all Chinese patents from 1985 to 2019 (the World edition has now been updated to 2020, the Chinese edition will be upgraded soon). The World Edition contains information on Chinese patents cited by Chinese applications and applications from other countries around the world, details of citations to the patent, details of other patents cited by the patent and the legal status of each patent in each country to date.
- Includes Basic Information Table, Patent Citation Table, Patent Cited Table, Patent Status Table.
- **Fields:** patent ID, IPC classification number, the priority date of patents, status type, etc.

3. Green Patent Innovation Data

- Produced in separate tables according to seven categories: waste management, nuclear power, transportation, energy conservation, agriculture and forestry, alternative energy production, administrative supervision and design, each table includes three categories of published invention, granted invention and utility model.
- Includes Published Invention Table, Granted Invention Table, Utility Model Table.
- **Fields:** application number, PTC application and publication data, national priority, patent grant number, etc.

4. Chinese Patent Text Data

- A large sample size of over 24 million text-based data, covering the period 1985-2020, consisting of patent description and patent claim information.
- **Fields:** patent ID, description.

5. Patent Citation and Patent Cited Data for Chinese Industrial Companies, Tax Survey Companies

- The data on Chinese industrial enterprises/tax survey enterprises are carefully matched with Chinese patent innovation data by reference to scientific data matching and processing methods. The duration of data for Chinese industrial enterprises is up to 2013, the data for patents is up to 2018, and the data for tax survey enterprises covers the period 2008-2015.
- The four categories of Invention Publication / Invention Grantation / Utility Model / Design are provided in detail using the following tables: Basic Information Table, Patent Citation Table, Patent Cited Table, Patent Status Table.
- **Fields:** company name, year, date of publication of application, patent agent, publication number of the cited patent application, application number, etc.

6. Global Patent Data

- Information on patent applications, patent grants and patent citations and patents cited in all countries worldwide.

7. China Business Registration Information

- Highly well-researched and extensive information on all registered enterprises in China, covering five areas: basic information, key concerns, intellectual property rights, corporate development and business status, with over 40 tables in total.
- **Fields:** holding company, change record, annual report, website filing, patent information, equity freeze, financing information, administrative license, etc.

8. PPP Project Data

- Data on government-enterprise urban infrastructure projects, covering basic project information, the preparation phase, procurement phase and implementation phase.
- Some of the tables at each stage are the contents of files, with the file name as its main display.

9. Name Data of Listed Companies and their Subsidiaries

- The names of listed companies and their subsidiaries and affiliated companies for each year since 1990.
- **Fields:** stock code, fiscal year, name of an associated company, the relationship of an associated company to a listed company, etc.

Listed Company Data

10. Stock Exchange Comment Letter - open

- Covering comment letters from the SSE and comment letters and their responses from the SZSE, including all data before 2020.01.01, providing a data basis for case studies and empirical studies on a large sample.
- **Fields:** company code, company name abbreviation, date of letter, type of comment letter, title, document name, etc.

11. (Green) Patent and their Citation and Cited Data of Listed Companies

- Provides data on (green) patent citations and cited details for listed companies for the year to the end of 2019.
- A table of basic information, patent citation and cited data, and status tables under each of the four sections: Published Inventions / Granted Inventions / Utility Models / Designs (none for Green Patents).
- **Fields:** patent ID, national priority, brief description, patent image, inventor, date of publication of transaction data, etc.

12. Annual Reports/Prospectuses/Social Responsibility Reports of Listed Companies in China

- PDF documents of more than 3,000 listed companies' annual reports/prospectuses/social responsibility reports from 1998 to 2019 (continuously updating).

13. Annual Reports/Prospectuses/Social Responsibility Reports of Listed Companies in US



14. Community and Second-hand Houses Transaction Data in China

- Information on the distribution of community and the second-hand houses transaction data in 28 provinces, autonomous regions, municipalities and 86 prefecture-level cities in China.
- Three sub-tables with information on communities, data on second-hand houses sold and second-hand houses for sale.
- **Fields:** province, city, developer, building type, house type, building structure, age of the house, etc.

15. Land Transaction Data

- Full range of data on leases, mortgages, transfers.
- Contains Basic Information Sheet on land supply results, an Information Sheet on Payment Agreement by Instalments, Land Rental Sheet, Land Mortgage Sheet, Land Transfer Sheet.
- **Fields:** administrative district, e-supervision number, payment period number, lease term, name of the land mortgagee, original and current land use rights holder, etc.

16. Data on Judicial Documents

- Providing a full range of case files from 1996 to 2016, including a full range of text files.
- **Fields:** full text, case type, case number, cause of action, court of appeal, legal basis, etc.

17. Government Procurement Data

- We have launched Government Procurement Data, which covers multi-dimensional information and provides a quality data resource for relevant research.
- **Fields:** contract amount, geographical area, supplier, purchaser, field of industry, etc.

18. Big Data on Population Migration

- It contains three sub-modules: information on the origin of in-migration by region, the destination of out-migration by region and the intensity of out-migration within each city, covering the general trend and detailed data on in-migration and out-migration of cities and provinces, providing quality data samples for relevant studies.
- **Fields:** inbound/outbound city, migration size index, inbound/outbound population share, intra-city travel intensity, etc.

19. National Standard Data

- This database covers all recommended and mandatory standards in force in China, those are about to be implemented and those have been abolished.
- Contains Mandatory Standards Data Sheets and Recommended Standards Datasheets.
- **Fields:** standard number, standard type, standard status, implementation and the competent unit, similar standards, etc.

20. National Fund Project Data

- Contains all NSF and SSF projects in 2019 and earlier, including detailed project leader information, status, results, and more, in addition to information about the projects that were established.
- Social science data includes all data up to 15th July 2019, natural science data includes all data for the period Jan.1987 to Jan. 2020.
- **Fields:** project leader, date of completion, award, completion status, 6/4/2-digit application code, corresponding discipline and academic department, etc.

21. Medical Big Data

22.China Administrative Divisions Data - open

23.Minimum Wage Data by Region

Sub-regional Data

24.COVID-19 Protection Patent Information Data - open

25.COVID-19 Prevention Standard Information Data - open

26.COVID-19 Companies Donation Data - open

27.The Number of Confirmed Cases of COVID-19 in Communities in China

- The database has a total of 11,868 rows of data and can be used to explore the trends of the pandemic in both time and space.
- **Fields:** address details, latitude and longitude, and the province, city and district to which they belong to, as well as the number of confirmed cases and the time of update.

COVID-19 Pandemic Data

28.Bank Branches in China

- Contains information on the distribution of all bank branches in China, using the status of financial licenses held by bank branches as a clue, divided into financial license holding table/institution out of control table/institution withdrawal table.
- **Fields:** branch of the institution,name of the institution,location of the institution,issuing authority ,date of issuance,status of holding the certificate.

Bank Data

29.Data on Penalties Imposed on Bank Branches in China for Non-Compliance

- In addition to covering the entire administrative penalty decisions published on the official website of the CBRC from 2014-2020, the latitude and longitude of the penalized banks have been integrated to facilitate location and spatial measurement analysis.
- **Fields:** penalized entity,legal person,cause,basis,penalty decision,name of the authority that imposed the penalty,time of the penalty,latitude and longitude of the penalized bank.



CnOpenData Catalogue

30. Local Government Message Board Text Data

- A comprehensive crawl of local government message board Q&A text content, spanning nearly 10 years (1st Jan. 2011 -31st Dec. 2020), and will continue to be updated in the future.
- **Fields:** question and reply content,number of likes,section,user nickname, etc.

Text
Data

31.CCTV News Feeds Text Data

- Text-based unstructured data, with concise and rich fields, for the period from 15th June 2016 to 31st Dec. 2019, and continuously updating.
- **Fields:** date of publication,ID column,news title and content.

32.Rumour Data-open

33.Poem of Tang & Song Dynasty - open

- Contains six tables, including a poetry table and author information table, with up to 250,000 lines.
- **Fields:** Tang poetry, Song lyrics, Song poetry author/name/text, etc.

34.Air Quality Site Monitoring Data - open

- Daily data, location information, monitoring data for air quality of each city (since 13th May 2014).
- **Fields:** monitoring point code,name,city,latitude,longitude,data collection date,time,city,air quality index.

35.Historical Weather Data in China - open

- Historical weather data in all cities (including counties) in China since 2011, with indicators including temperature, weather, wind direction wind, air quality (since 2016).
- **Fields:** province,district,date,minimum and maximum temperature,wind direction,air quality index,air quality registration.

36.Typhoon Data

- Contains basic typhoon information, typhoon track information and typhoon radius table.
- **Fields:** typhoon number,Chinese and English names,date of occurrence,typhoon intensity,latitude and longitude,wind custom,direction of movement,path number,wind speed level, etc.

Meteorological
Data

37.Airport Information Data - open

- World airport data, Chinese airport data (by 15th Jan. 2020).

38.Railway Information Data

- Contains information on all train times in 12306 (year 2007, 2011, 2013 and the following year), a list of all HSR lines and a timetable for the opening of each HSR station.

Traffic
Data

39. Chinese Film Box Office

- Contains basic film information sheet and box office details sheet (by Entertainment Data 15th Oct. 2019).
- **Fields:** title, year of release, genre, Cumulative box office, director and leading actor, single week and cumulative box office, etc.

40. Film Rating Data

- Comprehensive coverage of all film-related information, including a large number of concise short pieces as well as quality full-length film reviews.
- **Fields:** film title, douban rating, rating ID, rating, review content, actor name, date and place of birth, etc.

41. International Football Match Results Data - open

- Data on the results of football matches from 1872 to 2017.
- **Fields:** date of the match, name of the home and visiting team, name of the match, full team score (excluding penalties).

42. Data on Defaulters List

43. GUBA Comment Data

44. Chinese Provincial Party Media Newspaper Data

Data is constantly being updated

**CnOpenData Will be with You
All the Way Through This Ascetic Academic Journey.**



The Peak of the Academia Lies at the Apex of Data

Our Services

Online-Published Data

Non-Public Customer Service Data

Some of the non-public data has been used in several academic papers published in top journals.

Personalized Data Customization Service

It's our strength, all you need to do is to illustrate your requirements and then we will response to you as soon as possible.

The larger the data and the more difficult the crawl, the more confident we are!



 **CnOpenData**
Official Accounts



 **CnOpenData-ybl**
Sales Manager-Yang



 **CnOpenData-wsw**
Sales Manager-Wang



 **CnOpenData-dxy**
Business Service-Ding

Official link: www.cnopendata.com

Email: wangshiwei@cnopendata.com