

Data Policy at Science Data Bank

JIANG Lulu

Computer Network Information Center, Chinese Academy of Sciences

jianglulu@cnic.cn

Sept 26-27, 2022

A Helmholtz-CAS Workshop organized by the WIKOOP-INFRA Project

Outline

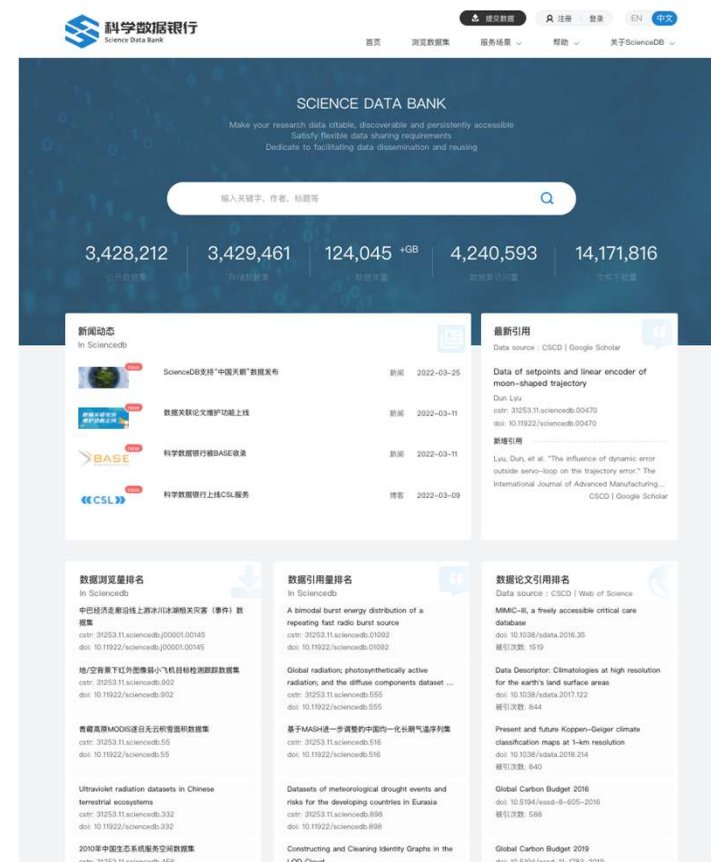
- **Overview on Science Data Bank**
- **Data Policy at Science Data Bank**
- **Facilitate the open data**

1. Overview on Science Data Bank

A non-profit, open and public generalist data repository

Science Data Bank provides data publishing and accessing service to the international science community. It is built and maintained by Computer Network Information Center, Chinese Academy of Sciences.

It ensures the **long-term management, persistent accessible** to the data published on it and aims to foster published scientific data compliance with agreed data standards and conventions, promoting the findability, accessibility, interoperability and reusability of scientific data, and cultivate data sharing culture in China scholarly community.



Service concept of Science Data Bank

Promote FAIR Principles & Endorse TRUST principles

- To make research data sharing easy and low-cost
- To support peer-review of associated manuscript
- To ensure data integrity and long-term accessibility
- To ensure data as open as possible, as closed as necessary
- To facilitate data reuse and protect rights & interests
- To enhance data dissemination and findability
- To track data usages and to incentivize data sharing



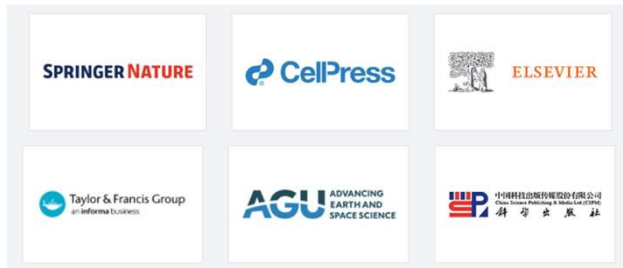
Milestones

Since 2015

ScienceDB was comprehensively upgraded

- Provide solutions for academic journals
- Internationalization and Standardization
 - Promoting FAIR principles

Recommended by more and more journals



2015

ScienceDB was launched

Initiate data publishing in China with *China Scientific Data* (a data journal)



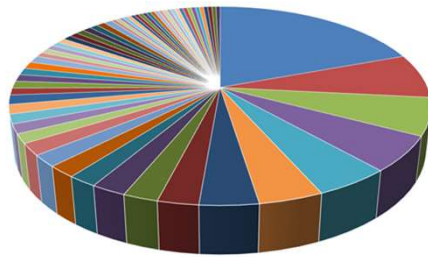
2019

2021

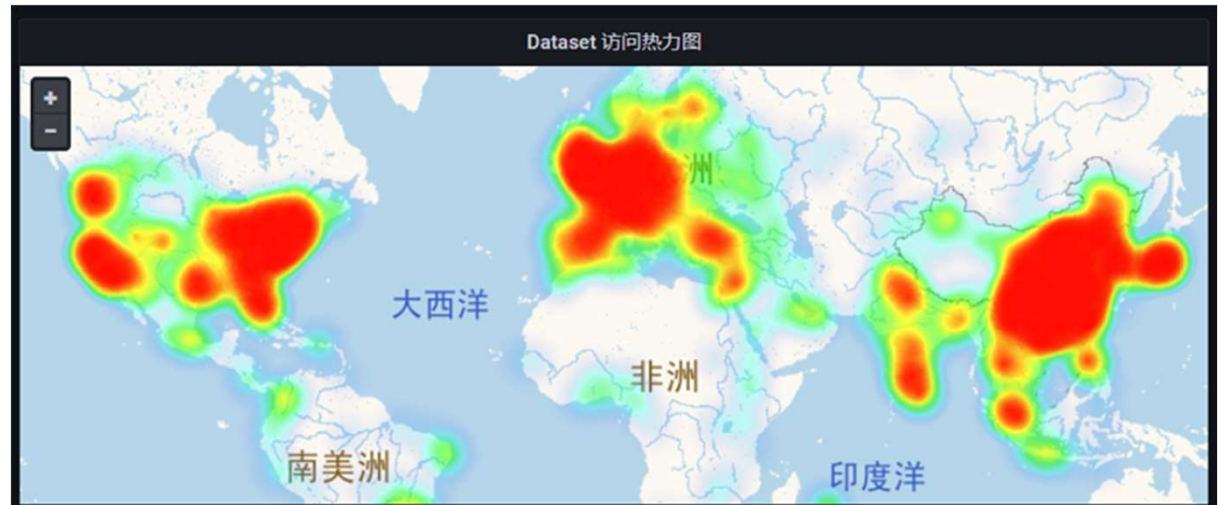
ScienceDB is a generalist Data Repository

- Multi-disciplinary
- Open service designated to global researchers / data depositors
- Accept any research data type
- Support any file size

Till now, ScienceDB has provided data access service for researchers cross over **170 countries**, whilst people from more than **60 countries** have shared their data by ScienceDB.



- | | | | | |
|-------------------------|------------------|-------------|----------------|-------------|
| • United States | • United Kingdom | • Brazil | • Netherlands | • Germany |
| • France | • Iran | • Australia | • Finland | • Turkey |
| • Czechia | • Greece | • India | • South Africa | • Italy |
| • Japan | • Kazakhstan | • Spain | • Taiwan | • Austria |
| • Canada | • Denmark | • Ethiopia | • Indonesia | • Vietnam |
| • Algeria | • Argentina | • Belgium | • Chile | • Kenya |
| • Mexico | • New Zealand | • Portugal | • Saudi Arabia | • Thailand |
| • the republic of korea | • Ukraine | • Angola | • Azerbaijan | • Belarus |
| • Bulgaria | • Cambodia | • Cameroon | • Colombia | • Croatia |
| • Egypt | • Ireland | • Israel | • Malta | • Mauritius |
| • Morocco | • Norway | • Peru | • Philippines | • Poland |
| • Republic of the Congo | • Romania | • Singapore | • South Korea | • Sri Lanka |
| • Sweden | • european union | • Benin | • Malaysia | |

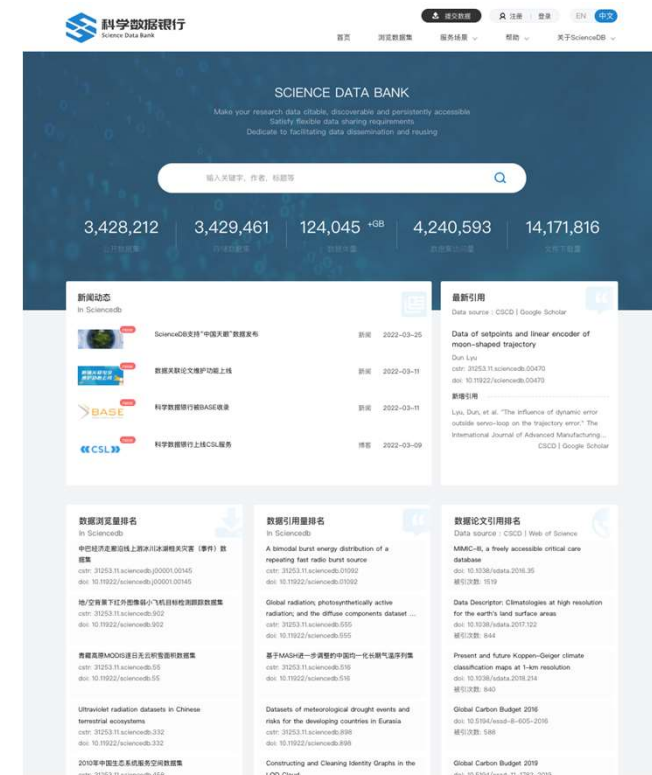
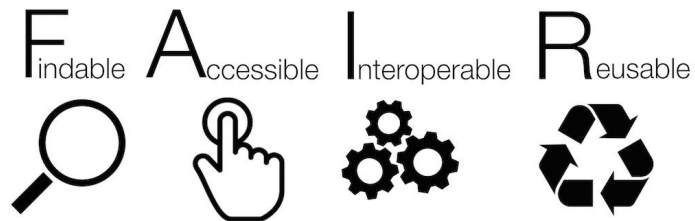


Outline

- Overview on Science Data Bank
- **Data Policy at Science Data Bank**
- Facilitate the open data in China

2. Data Policy at Science Data Bank

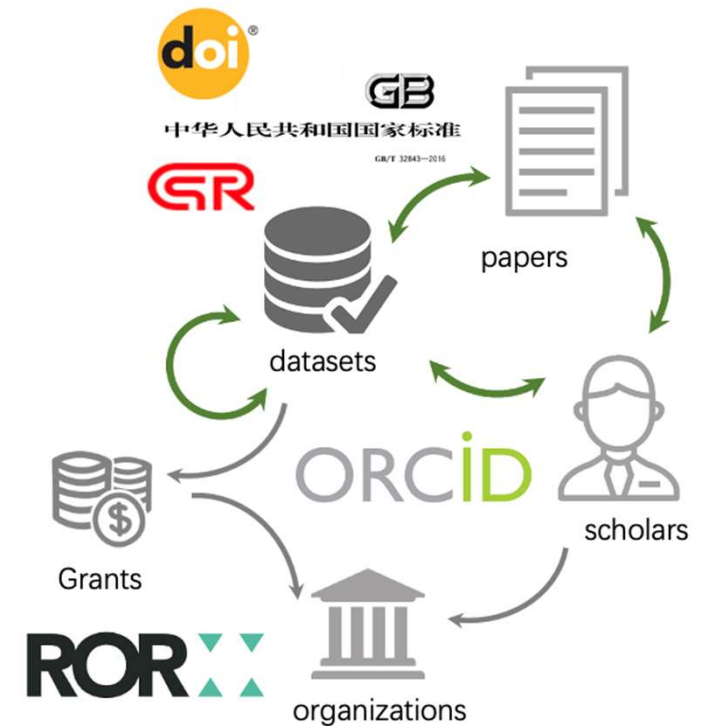
- Data Ingestion Phase
- Data Review and Curation Phase
- Data Published Phase



2. Data Policy at Science Data Bank

• Data Ingestion Phase

- Only the signed-in accounts have the permission to upload the data
- Data depositor shall have the corresponding intellectual property rights
- Ownership of the data uploaded on ScienceDB belongs to depositors or the public (if CC0 license chose)
- Depositors shall agree with T&C of the site ahead when create a new submission
- All the data are allocated a globally unique and persistent identifier when create a submission——both DOI number and CSTR number
- Embargo period is allowed to set on files



2. Data Policy at Science Data Bank

• Data Review and Curation Phase

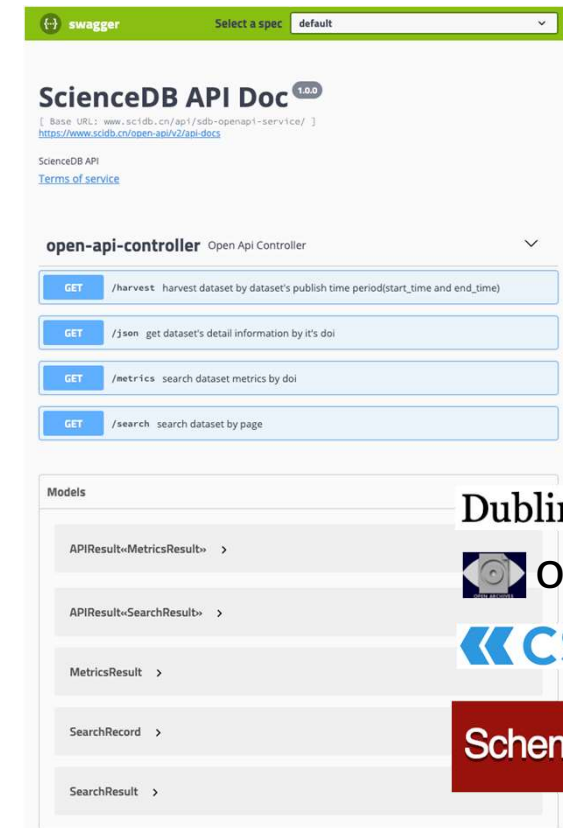
- All the data requires be reviewed by ScienceDB curator with a transparent criteria (see "4. Data review criteria" at https://www.scidb.cn/en/data_policy)
- By difference status of data in the workflow, relevant files are separated stored in distinct physical environment.
- Equipped with digital digest, like MD5, system is able to ensure data integrity during transmission cross distinct physical devices.



2. Data Policy at Science Data Bank

- Data Published Phase

- All the metadata at ScienceDB are under CC0 license
- All the data published at ScienceDB have a clear license
- All the files published online are followed by a MD5 code to ensure integrity
- Integrate the CSL standards to facilitate formal citation of data
- Tracking data usage and citation metrics automatically



Outline

- Overview on Science Data Bank
- Data Policy at Science Data Bank
- Facilitate the open data in China

3. Facilitate the open data in China

- Promote Chinese academic journals to join into open data
 - Provide solutions for academic journals called Data Community
 - Provide [Customized Service for Journal Data Policy](#)
- Promote more researchers in China focusing on open data and best practice through various training courses

At ScienceDB, compared with 2021 data there have been **21 times** as many depositors from China during 2022 YTD, whilst the growth rate of scholars from CAS is about **2,910% YOY**.

THANKS!

JIANG Lulu

jianglulu@cnic.cn

Computer Network Information Center, Chinese Academy of Science

Sept 26-27, 2022

A Helmholtz-CAS Workshop organized by the WIKOOP-INFRA Project