



中国科学院计算机网络信息中心  
Computer Network Information Center  
Chinese Academy of Sciences



Science Data Bank  
科学数据银行

# 科学数据银行 Science Data Bank

中国科学院计算机网络信息中心

报告人: 王鹏尧

2023年11月30日

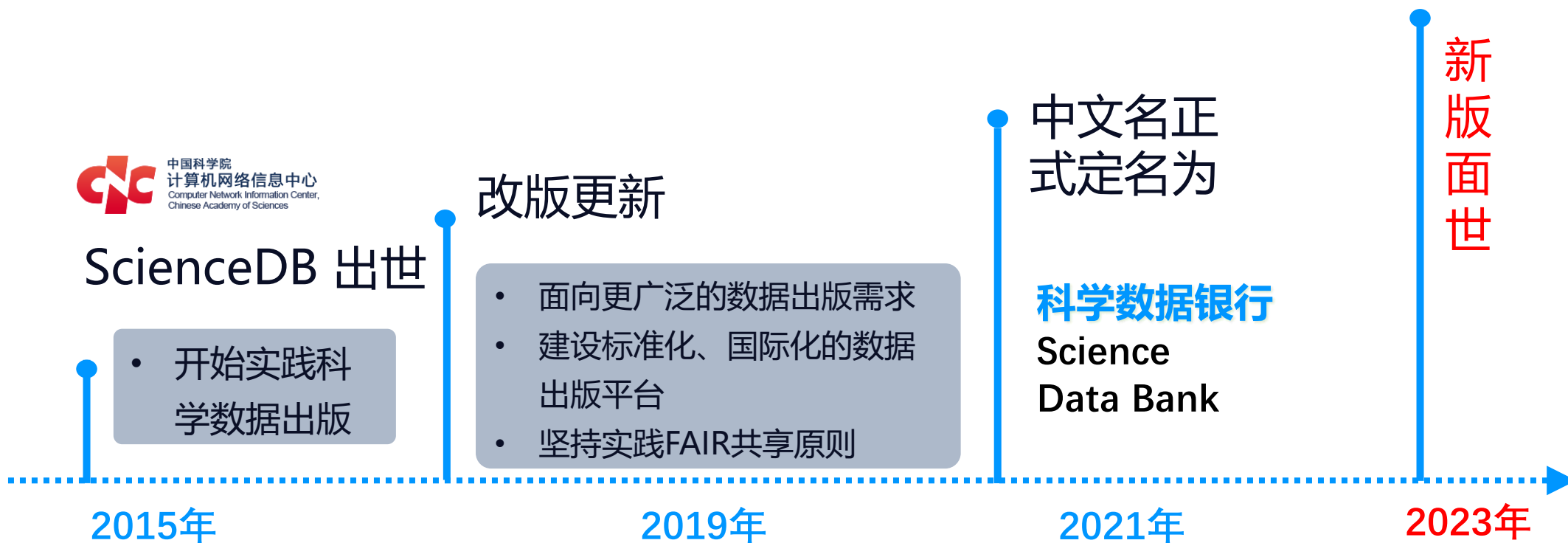
# 报告提纲

---

**一、平台定位与目标**

**二、平台建设成效**

**三、年度功能更新**



**通用型存储库**

不限定学科范围、数据类型

**服务对象**

论文作者、数据作者、期刊编辑部、学术出版商

**业务聚焦**

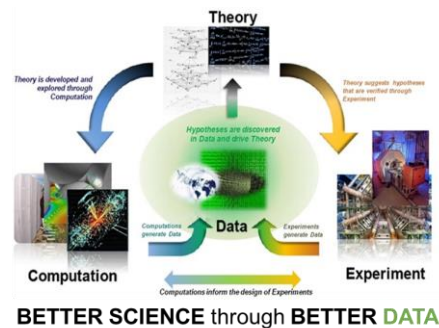
论文关联数据的出版与开放共享

# 一、平台定位与目标

## ➤ 建设初衷—面临的关键问题与挑战

### ① 开放共享难

缺少激励机制，导致数据开放共享意愿低，数据利用率低下，难以充分发挥数据价值



“ The annual costs of replication failures are estimated to be **\$28 billion**

—L. P. Freedman, I. M. Cockburn, T. S. Simcoe

到2020年，欧盟成员国通过开放数据累计节省成本预计达到**17亿欧元**

—Paul Ayris, UCL

### ② 期刊发展困境

**内容层面：**科研诚信问题、稿源问题、期刊论文质量

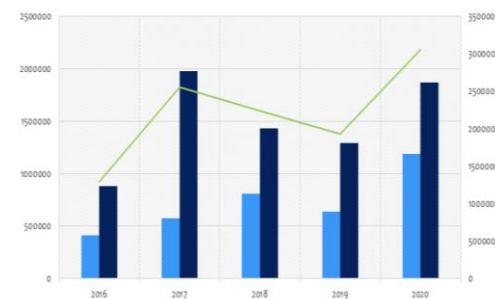
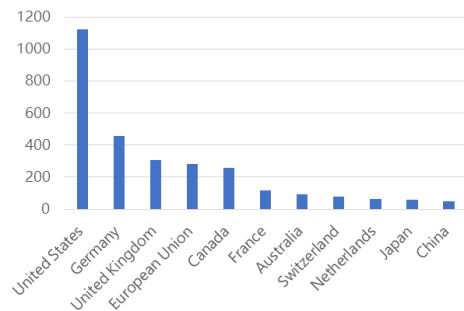
**影响力层面：**论文传播与影响力、期刊如何走上国际等

~ **88.1%** (开放论文关联数据的期刊Q1+Q2区占比)

~ **40%** (开放论文关联数据的论文引用影响力提升)

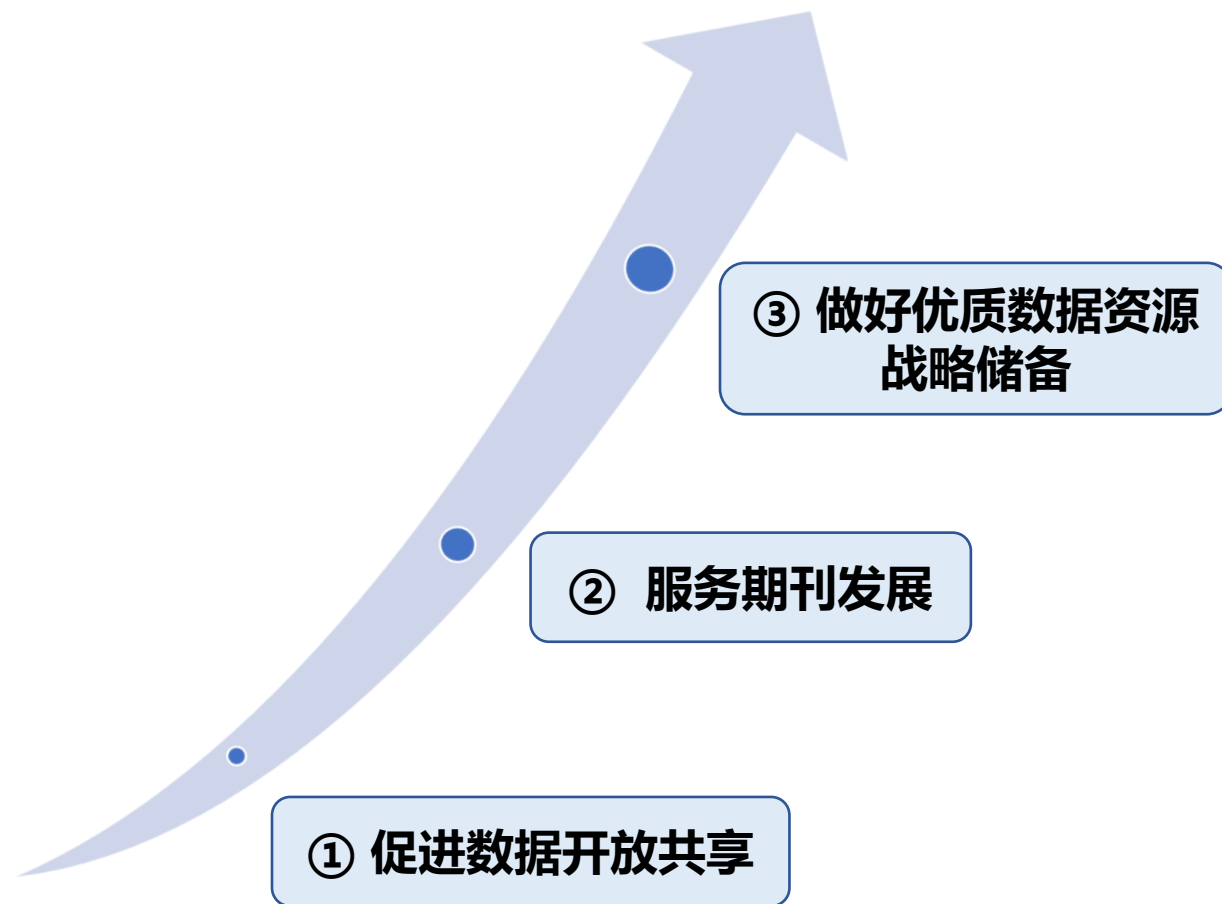
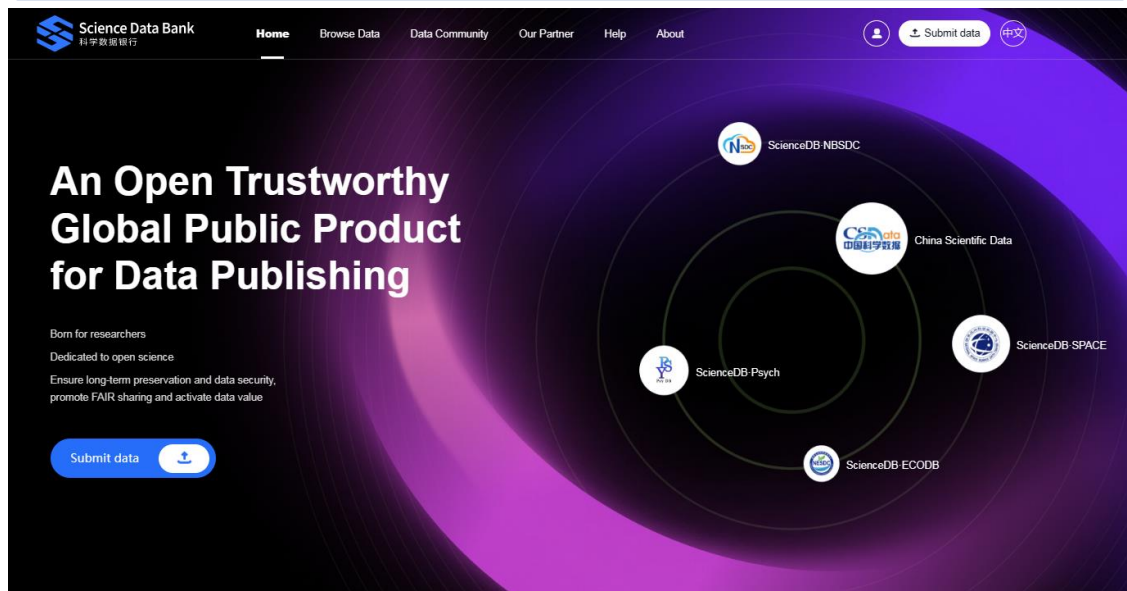
### ③ 优质数据资源战略储备稀缺

缺少获国际广泛认可的存储库，导致每年10万+数据集外流（院内1万+）；无法吸引全球数据资源汇聚



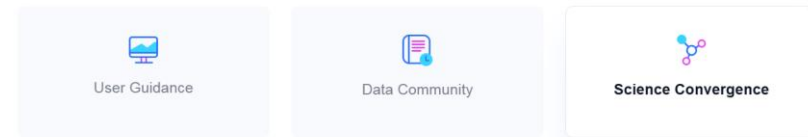
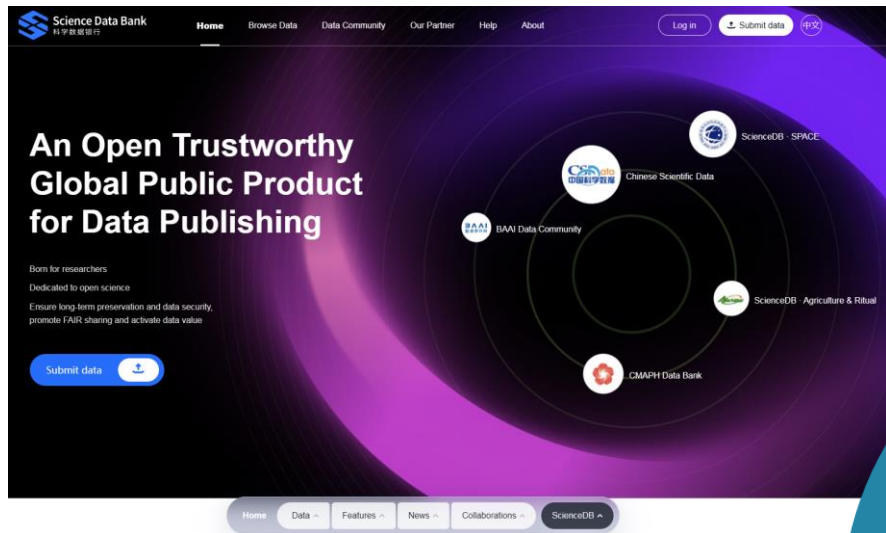
# 一、平台定位与目标

科学数据存储平台  
公益性·通用型·全球化



保护国家科学数据主权 服务期刊高质量发展 推动数据全球开放共享

# 一、平台定位与目标



## Connecting International Science Communities

Promote the discoverability and reusability of data on a global scale.

Users from **90+** countries/regions worldwide published their data on ScienceDB

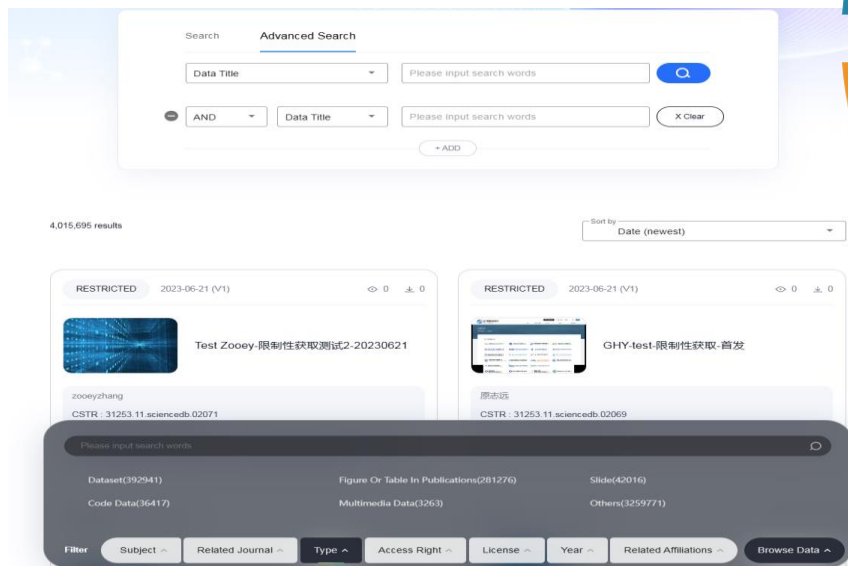
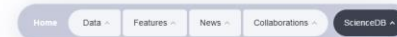
Users cross **190+** countries/regions access and use data from ScienceDB

PAST Chief Scientist, National Astronomical Observatory



"ScienceDB promotes replicable scientific research, enhances data value, and has a positive effect on promoting international collaborative research."

— PHD. LI



# 报告提纲

---

一、平台定位与目标

二、平台建设成效

三、年度功能更新

## 二、平台建设成效

### 获国际主要出版商推荐

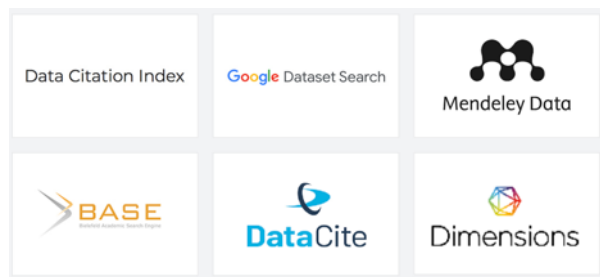
服务全学科领域，为全球**万余种**期刊的论文管理数据提供服务支持服务



Spring Nature 推荐的中国首家全学科通用型存储库

### 获国际学术平台收录

平台发布数据融入**全球**最主要的开放数据资源**传播生态链**



国际学术平台基本全覆盖

### 服务重要科研成果开放共享

服务FAST、联合国可持续发展计划(SDGs)、世界第三极、中国人脑毕生发展研究等重大科研数据成果发布

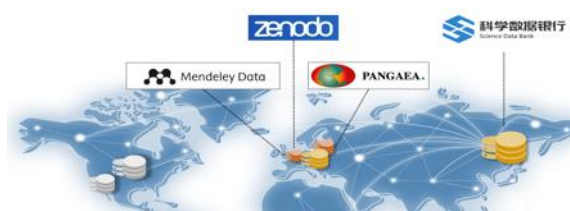


### 服务全球数据共享与开放数据获取

数据贡献者来自全球**105**个国家

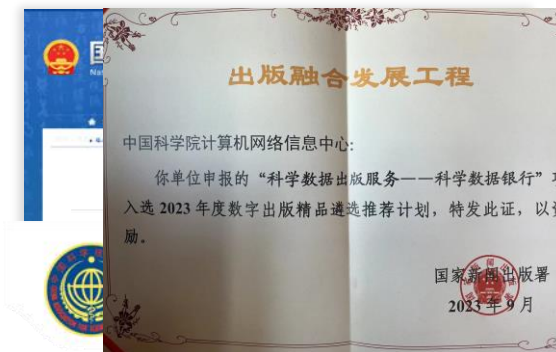
公开数据集超700万，平台访问8亿次，数据下载2300万次

### 推动构建全球开放数据资源镜像网络



国际TOP3存储库镜像合作国内独家

### 成为期刊主管部门推荐数据平台



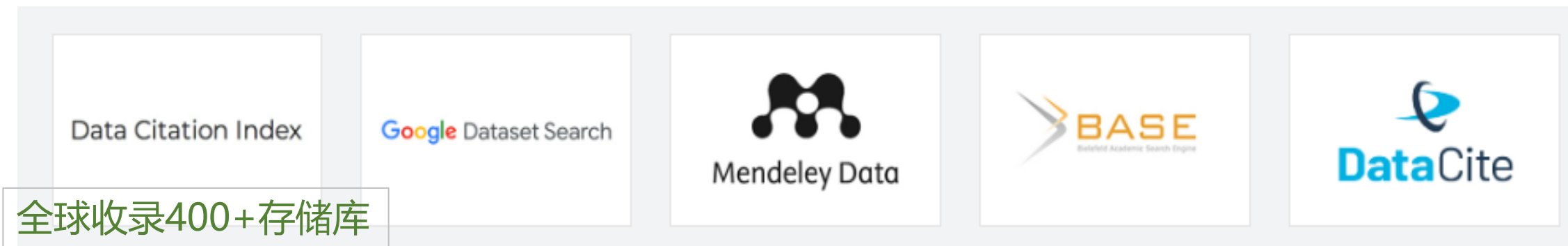
# ① 建设获得国际认可的科学数据出版全球公共产品



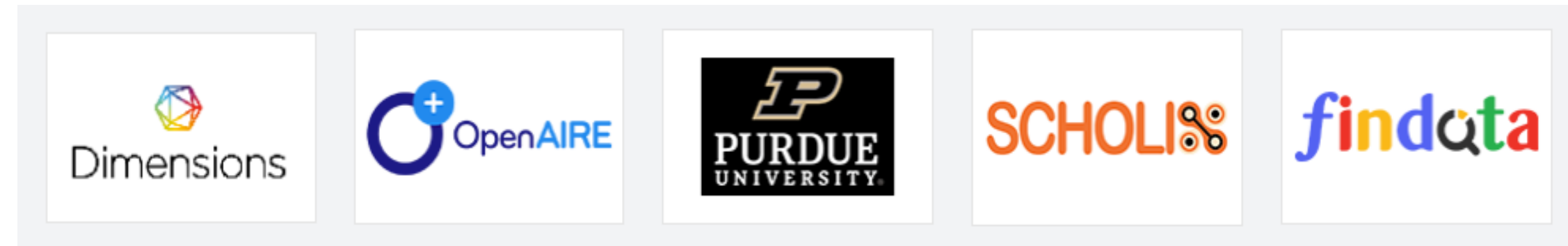
获国际主流出版商推荐 成为全球万余种期刊推荐平台

# ① 建设获得国际认可的科学数据出版全球公共产品

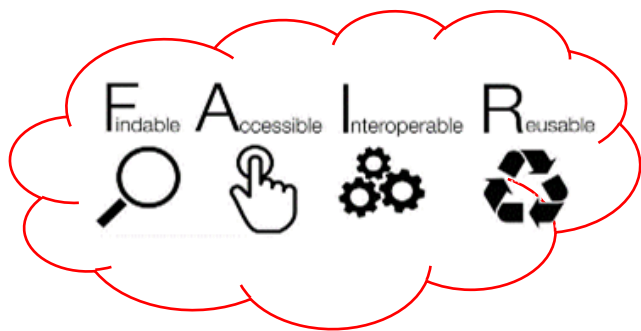
全球收录400+存储库



This row contains five logos: Data Citation Index, Google Dataset Search, Mendeley Data, BASE (Biology Abstract Search Engine), and DataCite.



This row contains five logos: Dimensions, OpenAIRE, Purdue University, SCHOLIX, and findqta.



获国际学术平台收录 促进开放数据的全球化传播

# ① 建设获得国际认可的科学数据出版全球公共产品

发表在ScienceDB平台上的数据可以快速进入国际传播渠道  
让  
我国期刊及论文走上国际及带来引流  
成为可能

**A bimodal burst energy distribution of a repeating fast radio burst source**

CSTR 31253.11sciencedb.01092  
DOI 10.11022sciencedb.01092

Di Li, Pei Wang, Weiwei Zhu, Bing Zhang, Xinwin Zhang, Ruan Dan, Yongkun Zhang, Yi Feng, Ningyu Tang, S. Chatterjee, J. M. Cordes, M. Cruces, Shi Dai, V. Gajjar, G. Hobbs, C. Jin, M. Kramer, D. R. Lorimer, Chenchen Miao, Chenhui Niu, Jiarui Niu, Zhichen Pan, Lei Qian, L. Spiller, D. Werthimer, Guoqiang Zhang, Fayin Wang, Xiaoyao Xie, Youling Yue, Lei Zhang, Qijun Zhi, Yan Zhu

This dataset has been curated by **NAOC - The Chinese National Astronomical Science Center (CRAFTS) Data Community**

This dataset has been included in **NSFC**

**Associated publication**

**A bimodal burst energy distribution of a repeating fast radio burst source**  
Li D., Wang P., Zhu W.W. et al. A bimodal burst energy distribution of a repeating fast radio burst source. *Nature* 598, 267–271 (2021).  
<https://doi.org/10.1038/s41586-021-03878-5>  
doi:10.1038/s41586-021-03878-5

**Energy levels and magnetic dipole transition parameters for the Nitrogen isoelectronic sequence**

ScienceDB, July 2022  
<https://doi.org/10.57760/sciencedb.j00113.00022>

Posted by Mu-Hong Hu, Nan Wang, Pin-Jun Ouyang, Feng, Xin-Jie, Yang, Wu, Chen-Sheng

**Associated publication**

**Energy levels and magnetic dipole transition parameters for the nitrogen isoelectronic sequence**  
Mu-hong Hu, Nan Wang, Pin-Jun Ouyang, Xin-Jie Feng, Yang Yang, Chen-sheng Wu  
2022, *Chinese Physics B* - Article  
[Add to Library](#)

**1. Magnetic properties of oxides and silicon single crystals**

Author: Zhong-Xue Huang (黄志学) [claim]; Rui Wang (王瑞) [claim]; Xin Yang (杨鑫) [claim]; Hao-Feng Chen (陈浩锋) [claim]; Li-Xin Cao (曹立新) [claim]

Description: The data related to the article "Magnetic properties of oxides and silicon single crystals" published in the journal Chinese Physics B. The magnetic properties of single crystals Si, SrTiO<sub>3</sub>, LaAlO<sub>3</sub>, MgO, and (La,Sr)(Al,Ta)O<sub>3</sub> were investigated...

Publisher, Year: Science Data Bank, 2022-06-20  
Document Type: dataset; [Dataset]  
Content Provider: 科学数据银行 (ScienceDB)  
Science Data Bank (ScienceDB)

## ② 服务我国期刊开展论文关联数据出版工作

### ✓ 中国科学院

- 国内率先出台政策，推动科技期刊论文关联数据汇交
- 鼓励期刊建立数据政策，纳入创办期刊的发展规划合理性考核项

### ✓ 国家新闻出版署

- 《2021/2022年度期刊核验表》明确列出：学术期刊“是否向作者推荐...进行数据汇交和共享”

### ✓ 中国科协

- 《2021年世界一流科技期刊建设工作要点》，试运行“期刊论文关联数据汇交和应用服务平台”，推动建立论文发表相关数据存缴、发布、引用、共享机制
- 2022年度中国科技期刊卓越行动计划选育高水平办刊人才子项目，强调“加强和规范论文关联数据汇交管理与开放共享”

### ✓ 中国科协办公厅、中国科学院办公厅

- 2022年11月 《关于组织开展期刊论文关联数据汇交工作的通知》

相关工作获得院、中宣、科协等大力支持

### 中国科学院厅局文件

传播字〔2021〕2号

#### 中国科学院科学传播局关于进一步加强和规范论文关联数据管理的通知

院属各单位、院机关各部门：

为贯彻落实《国务院办公厅关于印发科学数据管理办法的通知》（国办发〔2018〕17号）、《中共中央办公厅 国务院办公厅印发〈关于进一步弘扬科学家精神加强作风和学风建设的意见〉》（中办发〔2019〕35号）关于科学数据管理的相关要求，根据《中国科学院关于印发〈中国科学院科学数据管理与开放共享办法（试行）〉的通知》（科发办字〔2019〕11号）和《中国科学院办公厅关于印发〈中国科学院科学数据工作要点〉的通知》（科办〔2020〕

我院率先出台政策加强论文关联数据管理工作

中国科协办公厅 中国科学院办公厅  
关于组织开展期刊论文关联数据汇交工作的通知

科协改革进行时 2022-11-23 19:45 发表于北京

点击 科协改革进行时 关注我们

中国科协所属全国学会、协会、研究会，中国科学院院属各单位、机关各部门：

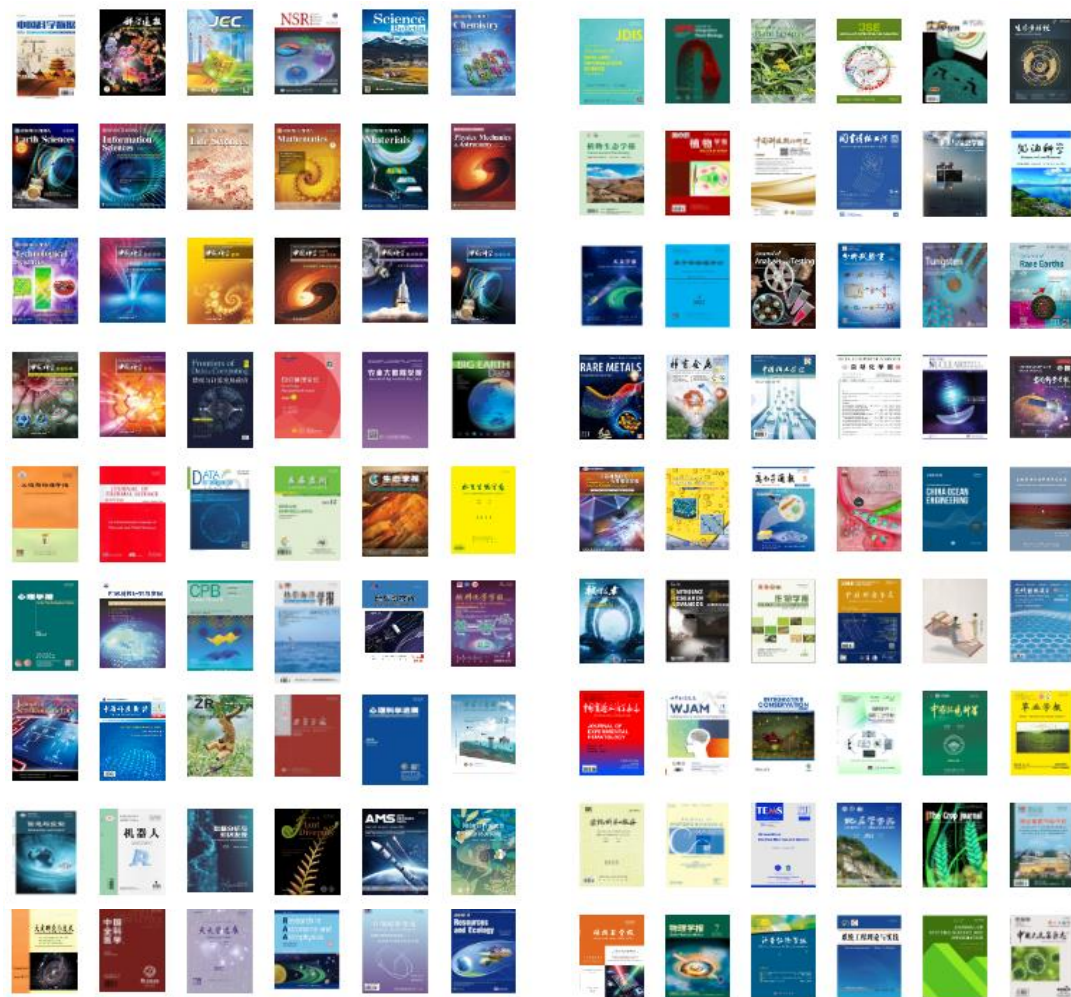
为落实《中华人民共和国数据安全法》和《科学数据管理办法》，提升我国科学数据管理水平和利用能力，保障国家科学数据安全，推动科学数据开放共享，更好发挥科学数据作为国家科技创新和经济社会发展重要基础性战略资源的支撑服务作用，中国科协、中国科学院决定联合开展论文关联数据汇交工作，现将相关事宜通知如下。

2022年11月  
中国科协办公厅、院办公厅  
联合发文

2021年度期刊核验表		国家新闻出版署制表 第4页	
出版能力	平均期印数（万册）	平均期发行量（万册）	年度出版期数（期）
年度总学术期刊出版有关情况（非学术期刊无须填写）			
期刊数量	年收文章（篇）	批发文章（篇）	开放获取情况 <input type="checkbox"/> 否 <input type="checkbox"/> 部分论文 <input type="checkbox"/> 全部论文
	是否将论文数据加工整理并保存		<input type="checkbox"/> 是 <input type="checkbox"/> 否
	是否向作者推荐使用我国自主研发的数据存储平台（如“科学数据银行（ScienceDB）”）进行数据汇交和共享		<input type="checkbox"/> 是 <input type="checkbox"/> 否
数字出版平台建设或合作情况	与国际出版商合作情况		

连续2年进入国家新闻出版署期刊  
年检核验项

# ② 服务我国期刊开展论文关联数据出版工作



请按照下方页面操作指引，完成论文关联数据出版的第一步！

一、论文关联数据出版的意义

二、论文关联数据出版的范围

三、前期数据准备

四、前期数据审核

六、数据审核

九、帮助和支持

预览

11 导出word

导出WORD文件 下一步

The Significance of Paper Related Data Sharing

"Paper Related Data" refers to the data generated through basic research, applied research, experimental testing, etc. to support the publication of academic papers, as well as data obtained through observation, inspection, investigation, inspection and testing, etc. and used to form papers' claims and the raw data or its derived data support the research conclusions of the paper. The sharing of paper related data is an important evidence for the research results and conclusions of the paper, improving the verifiability and transparency of the research, the sharing of data improves the efficiency and advancement of scientific papers, the reuse of data saves manpower and material resources, allowing scientists form an innovative research, paper related data as a new type of academic achievement will promote innovative cooperation and talent training among disciplines, and promote open scientific practice.

二、Types and Requirements of the Paper Related Data

A	B	C
79	Light: Advanced Manufacturing	<a href="https://www.light-am.com/news/contact_us">https://www.light-am.com/news/contact_us</a>
80	图书情报工作	<a href="https://www.lis.ac.cn/CN/column/column38">https://www.lis.ac.cn/CN/column/column38</a>
81	生态学报	<a href="https://www.ecologica.cn/stxb/news/view/2">https://www.ecologica.cn/stxb/news/view/2</a>
82	能源化学 (英文) Journal of Energy Chemistry	<a href="https://www.jenergychem.com/EN/news/ne">https://www.jenergychem.com/EN/news/ne</a>
83	分析测试技术与仪器	<a href="https://www.fxcsjyyq.net/fxcsjyyq/news/g">https://www.fxcsjyyq.net/fxcsjyyq/news/g</a>
84	声学学报	<a href="https://jac.ac.cn/news/tongzhigonggao/0c3">https://jac.ac.cn/news/tongzhigonggao/0c3</a>
85	应用声学	<a href="http://yxsc.cnjournals.cn/ch/reader/view_ne">http://yxsc.cnjournals.cn/ch/reader/view_ne</a>
86	Journal of Systems Science and Complexity (JSSC)	<a href="https://sysmath.cjoe.ac.cn/jssc/EN/column/">https://sysmath.cjoe.ac.cn/jssc/EN/column/</a>
87	催化学报	<a href="https://www.cjatal.com/CN/news/news729">https://www.cjatal.com/CN/news/news729</a>
88	分子催化	<a href="https://www.jmchina.org/ch/reader/view_r">https://www.jmchina.org/ch/reader/view_r</a>
89	Machine Intelligence Research	<a href="https://www.mi-research.net/news/Guide%">https://www.mi-research.net/news/Guide%</a>
90	Zoological Research (ZR) 动物学研究	<a href="https://www.zoores.ac.cn/news/guideline.ht">https://www.zoores.ac.cn/news/guideline.ht</a>
91	新型炭材料 (中英文) New Carbon Materials	<a href="http://xxtl.sxicc.ac.cn/news/qikandongtal/G">http://xxtl.sxicc.ac.cn/news/qikandongtal/G</a>
92	工业仪表与自动化装置	<a href="http://yb-zdh.shaanqu-group.com/docs/ne">http://yb-zdh.shaanqu-group.com/docs/ne</a>
93	自动化学报	<a href="http://www.aas.net.cn/news/tqzx.htm">http://www.aas.net.cn/news/tqzx.htm</a>
94	植物科学学报	<a href="https://www.whzwxj.cn/news/tongzhigong">https://www.whzwxj.cn/news/tongzhigong</a>
95	IEEE/CAA Journal of Automatica Sinica	<a href="https://www.ieee-jas.net/news/Information/">https://www.ieee-jas.net/news/Information/</a>
96	力学学报	<a href="https://lxbx.cstam.org.cn/news/kexueshuju">https://lxbx.cstam.org.cn/news/kexueshuju</a>
97	模式识别与人工智能	<a href="http://mssb.jugk.net/">http://mssb.jugk.net/</a>
98	地质科学	<a href="http://www.dzqx.org/dzdxs-upload/news/geology/20234283258994.docx">http://www.dzqx.org/dzdxs-upload/news/geology/20234283258994.docx</a>
99	系统科学与复杂性(英文版)	<a href="https://sysmath.cjoe.ac.cn/jweb_xtkoysx/CN/news/news110.shtml">https://sysmath.cjoe.ac.cn/jweb_xtkoysx/CN/news/news110.shtml</a>
100	系统科学与数学	<a href="https://sysmath.cjoe.ac.cn/jweb_xtkoysx/CN/news/news110.shtml">https://sysmath.cjoe.ac.cn/jweb_xtkoysx/CN/news/news110.shtml</a>
101	色谱	<a href="https://www.chrom-china.com/CN/column/column21.shtml">https://www.chrom-china.com/CN/column/column21.shtml</a>
102	邮电设计技术	<a href="http://ydsjts.paperopen.com/#/NewsDetails?Ite">http://ydsjts.paperopen.com/#/NewsDetails?Ite</a>
103	中国医学影像技术	<a href="http://www.cjilt.com/cjilt/ch/first_menu.aspx?pa">http://www.cjilt.com/cjilt/ch/first_menu.aspx?pa</a>
104	当代经济	<a href="http://www.djzz.com/news/132.html">http://www.djzz.com/news/132.html</a>

构建数据社区**291**个（覆盖400余本期刊），服务**100**余个期刊建立数据政策

# ② 服务我国期刊开展论文关联数据出版工作

中国科学数据

中国河北晚侏罗世蝶螈亚目（两栖纲，有尾目）基于属种干沟青龙螈 (*Qinglongtriton ganguensis*) 骨骼高精度 $\mu$ CT三维数据集

Dataset of 3D high-resolution  $\mu$ CT scan of fossil specimens of *Qinglongtriton ganguensis*, a basal salamandroid (Amphibia, Urodela) from the T1 linear traverse of Mahai, Shrivine, China

在生物演化的长河中，99.9%以上的物种已灭绝

摘要：古生物学研究进入数字化时代，与传统的化石机械清理和薄片技术相比，X射线断层扫描技术（high-resolution X-ray tomography或 $\mu$ CT）以其高精度、高效率和无损检测等优势迅速获得了古生物研究人员的青睐。干沟青龙螈 (*Qinglongtriton ganguensis*) 是发现于中国河北晚侏罗世地层山地的蝶螈亚目的基干属种，距今约1亿六千万年。本文提供了该属种正型标本 (PKUP V0226) 和特选白化标本 (PKUP V0224、PKUP V0224) 的 $\mu$ CT数据集。正型标本上并未发现用于3D打印的 $\mu$ CT切片及5种标本中的重要数据集和切片文件。这些高精度 $\mu$ CT技术在古生物学研究中的广泛应用，该技术的应用发现了青龙螈多个独特生物学特征，包括独特的骨化轴突等，为研究其生物学演化提供了重要依据。本数据集为目前 $\mu$ CT技术在古生物学研究中的应用提供了参考，同时其数据集、视图和可用于3D打印的虚拟模型也为公众了解和研究精美化石提供了便利渠道。

关键词：河北；晚侏罗世；蝶螈亚目；干沟青龙螈； $\mu$ CT数据；骨骼化石



丰富出版内涵 提升科研诚信

Data in Brief

Authors are invited to submit a Data in Brief article alongside their research article. The Data in Brief article is an attractive alternative to supplementary material. Data in Brief articles ensure that your data and the metadata to understand it, which is normally buried in supplementary material, is actively reviewed, curated, formatted, indexed, given a DOI and publicly available to all upon publication. If your research article is accepted, your Data in Brief article will also be published in the new, Open Access journal, Data in Brief. Please use the Data in Brief template available through the below link: [Data in Brief Template](#)

Include interactive data visualizations in your publication and let your readers interact and engage more closely with your research. Follow the instructions [here](#) to find out about available data visualization options and how to include them with your article.

CPB Chinese Physics B

Announcement

Chinese Physics B begins to accept and publish "Data Papers" in 2021. The Data Paper is dedicated to publishing original and important resources (datasets) that are of high value and interest to a broad readership in physical sciences. Manuscripts reporting datasets or a major informational database should contain new insights derived from an analysis of the data sets. Data Articles can be organized like Research Articles and have no critical limits on the length and figure/table numbers. To assure the reuse of the dataset, more detailed description is required in the Method section, including research methods, the obtained data source, and data usage notes (such as the URL link, software used, and date file types). The data will be deposited in an accredited data repository.

11 March 2021 View: 303

拓展数据论文类型 丰富稿源

# ③ 积极做好优质数据资源战略储备

机构名称	初始化需归档数据量(条) =C+D+E	提交率=(归档中+已归档)/初始化待归档数据量
中国科学院福建物质结构研究所	2669	100.00%
中国科学院昆明植物研究所	708	100.00%
植物研究所	625	100.00%
中国科学院上海药物研究所	549	100.00%
中国科学院微生物研究所	328	100.00%
中国科学院高能物理研究所	321	100.00%
中国科学院西双版纳热带植物园	303	100.00%
中国科学院青岛生物能源与过程研究所	205	100.00%
中科院生物物理研究所	177	100.00%
中国科学院南京地质古生物研究所	175	100.00%
中国科学院上海高等研究院	161	100.00%
中国科学院广州生物医药与健康研究院	152	100.00%
中国科学院上海光学精密机械研究所	149	100.00%
中科院古脊椎所	145	100.00%
中国科学院青藏高原高原研究所	115	100.00%
中国科学院心理研究所	92	100.00%
中国科学院上海硅酸盐研究所	91	100.00%
中国科学院过程工程研究所	91	100.00%
中国科学院合肥物质科学研究院	89	100.00%
中国科学院苏州生物医学工程技术研究所	88	100.00%
沈阳应用生态研究所	86	100.00%
中国科学院自动化研究所	84	100.00%
中国科学院近代物理研究所	81	100.00%
中国科学院广州地球化学研究所	74	100.00%
中国科学院地球化学研究所	64	100.00%
中国科学院云南天文台	61	100.00%
中国科学院西安光学精密机械研究所	59	100.00%
中国科学院南京地理与湖泊研究所	59	100.00%
西北高原生物研究所	55	100.00%
中国科学院上海天文台	52	100.00%
中国科学院数学与系统科学研究院	41	100.00%
中国科学院山西煤炭化学研究所	41	100.00%
中国科学院青海盐湖研究所	39	100.00%
中国科学院长春光学精密机械与物理研究所	36	100.00%
中国科学院计算技术研究所	33	100.00%
上海技术物理研究所	29	100.00%
中国科学院半导体研究所	27	100.00%
中国科学院计算机网络信息中心	27	100.00%
中国科学院新疆天文台	27	100.00%
中国科学院工程热物理研究所	23	100.00%
中国科学院空天信息创新研究院	23	100.00%
中国科学院广州能源研究所	11	100.00%
中国科学院沈阳自动化研究所	8	100.00%
中国科学院理论物理所	6	100.00%
声学研究所	5	100.00%
中国科学院地质与地球物理研究所	5	100.00%
国家授时中心	5	100.00%
中国科学院电工研究所	4	100.00%
中国科学院东北地理与农业生态研究所	4	100.00%
中国科学院软件研究所	4	100.00%
中国科学院武汉岩土力学研究所	3	100.00%
中国科学院文献情报中心	2	100.00%
中国科学院分子细胞科学卓越创新中心	1	100.00%
中国科学院分子植物科学卓越创新中心	1	100.00%
中国科学院南京土壤研究所	1	100.00%

University Of Chinese Academy Of Sciences (4,629)	Fujian Institute Of Research On The Structure Of Matter (1,527)	Hefei Institutes Of Physical Science, Chinese Academy Of Science, CAS (1,220)
Shanghai Institute Of Organic Chemistry (934)	University Of Science And Technology Of China (641)	Institute Of Chemistry (545)
Northeast Institute Of Geography And Agroecology (508)	Institute Of Zoology, CAS (441)	Shanghai Institute Of Optics And Fine Mechanics (433)
Kunming Institute Of Botany (391)	Institute Of Botany, The Chinese Academy Of Sciences, CAS (374)	Institute Of Plasma Physics (322)
Institute Of Genetics And Developmental Biology (286)	Changchun Institute Of Applied Chemistry (279)	Chinese Academy Of Sciences, CAS (252)
Xishuangbanna Tropical Botanical Garden, CAS (251)	Institute Of Hydrobiology (231)	Technical Institute Of Physics And Chemistry (213)
Shanghai Institute Of Materia Medica (194)	Kunming Institute Of Zoology (187)	National Space Science Center (182)
Institute Of Oceanology, CAS (181)	Xinjiang Technical Institute Of Physics & Chemistry (167)	Dalian Institute Of Chemical Physics, CAS (165)

2023 数据标题 请输入检索词 检索 导出全部

通知所选数据作者 通知所有数据作者 检索到 11062 条数据

数据标题	数据DOI	本机构作者	论文标题	操作
<input type="checkbox"/> Experimental demonstration of optical trapping and manipulation with multifunctional metasurface	10.6084/m9.figshare.c.5738714	Li Xingyi Zhou Yuan Ge Suyang 查看更多	Experimental demonstration of optical trapping and manipulation with multifunctional metasurface	发送邮件 查看邮件记录 预览 获取归档地址
<input type="checkbox"/> CCDC 2168963: Experimental Crystal Structure Determination	10.5517/ccdc.csd.cc2bqx11	Zheng Chang Tang Yu Yu Biao	Tr[(N-carbazoly)phosphine Gold(I) Complexes: Structural and Catalytic Activity Studies	发送邮件 查看邮件记录 预览 获取归档地址
<input type="checkbox"/> Additional file 4 of m6A RNA methylation impairs gene expression variability and reproductive thermotolerance in Arabidopsis	10.6084/m9.figshare.21622458.v1	Wang Ling Zhuang Haiyan Fan Wenwen 查看提交	m6A RNA methylation impairs gene expression variability and reproductive thermotolerance in Arabidopsis	发送邮件 查看邮件记录 预览 获取归档地址

逐步督促科研人员完成数据归档工作

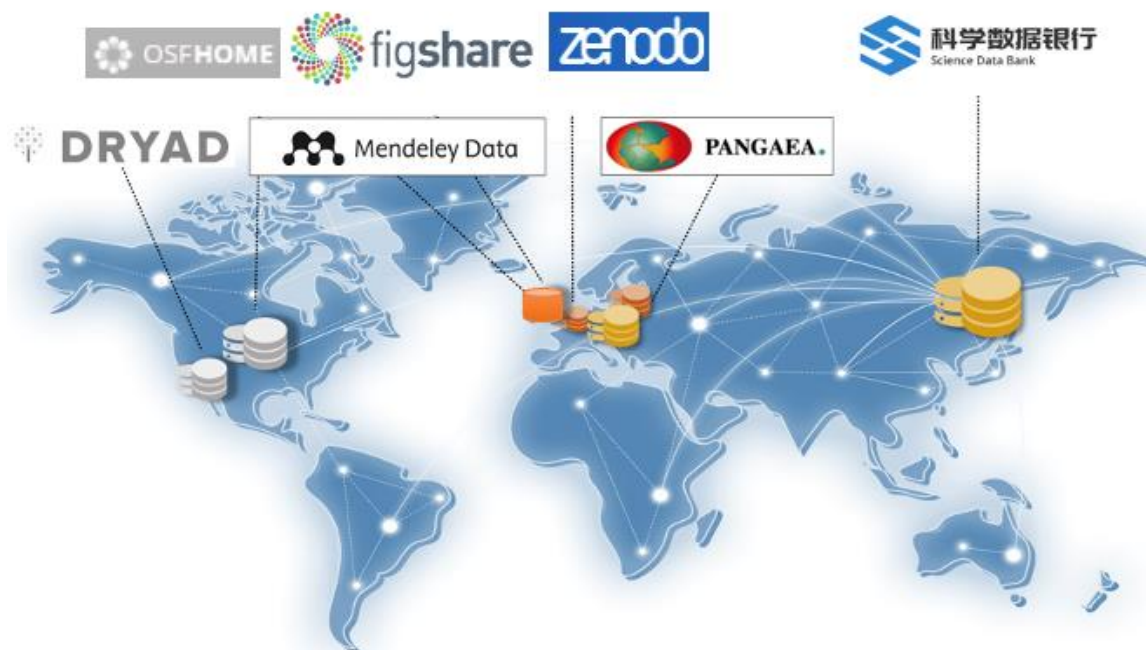
- 10月31日前，完成「待归档数据列表」内数据归档（清零）  
通知数据作者，按《科学数据银行用户使用手册（数据银行注册用户）》步骤提交数据  
通过后台发送邮件通知  
与出全部数据清单，通过邮件或企业微信等方式下发给科研人员，科研人员自行认领、提交数据
- 逐步完成「待归档数据列表」外的历史发布数据归档  
告知科研人员，主动将之前发布在海外平台的数据，归档到ScienceDB  
告知科研人员按照期刊征稿文件要求，在发论文时，选择将数据首发/备份发布在ScienceDB上。  
目前ScienceDB已获得多个国际数据库索引，被列入国际知名期刊推荐的数据发布存储库列表，科研人员有期刊需求的前提下，优先选择ScienceDB存储数据，首发在ScienceDB上的数据，优先再次归档在ScienceDB上，前减除了后续维护人做数据的工作量
- 要求本单位科研人员，产出的论文关联数据 首发或备份在ScienceDB

**保障2022年度数据归档  
正在开展2023年度数据归档工作**

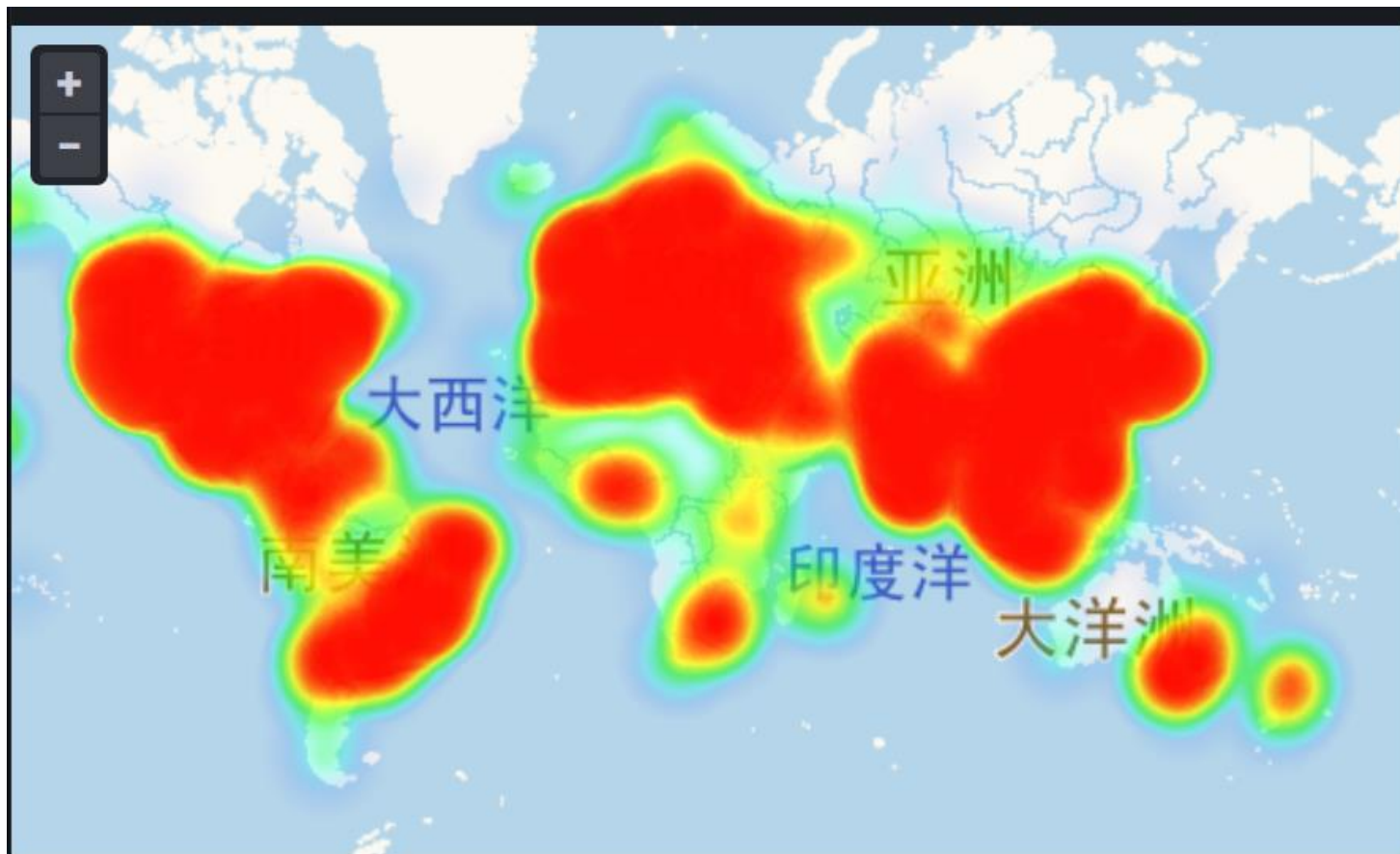
### ③ 积极做好优质数据资源战略储备



2023.1.10 与Figshare线上签约



- 实现与**国际3大主流**存储库**Figshare、Zenodo、Mendeley Data**以及**Pangaea**等国际品牌合作
- 实现外流数据的合法合规回流
- 为国内科研人员提供数据访问便捷服务



作者国家数量

105

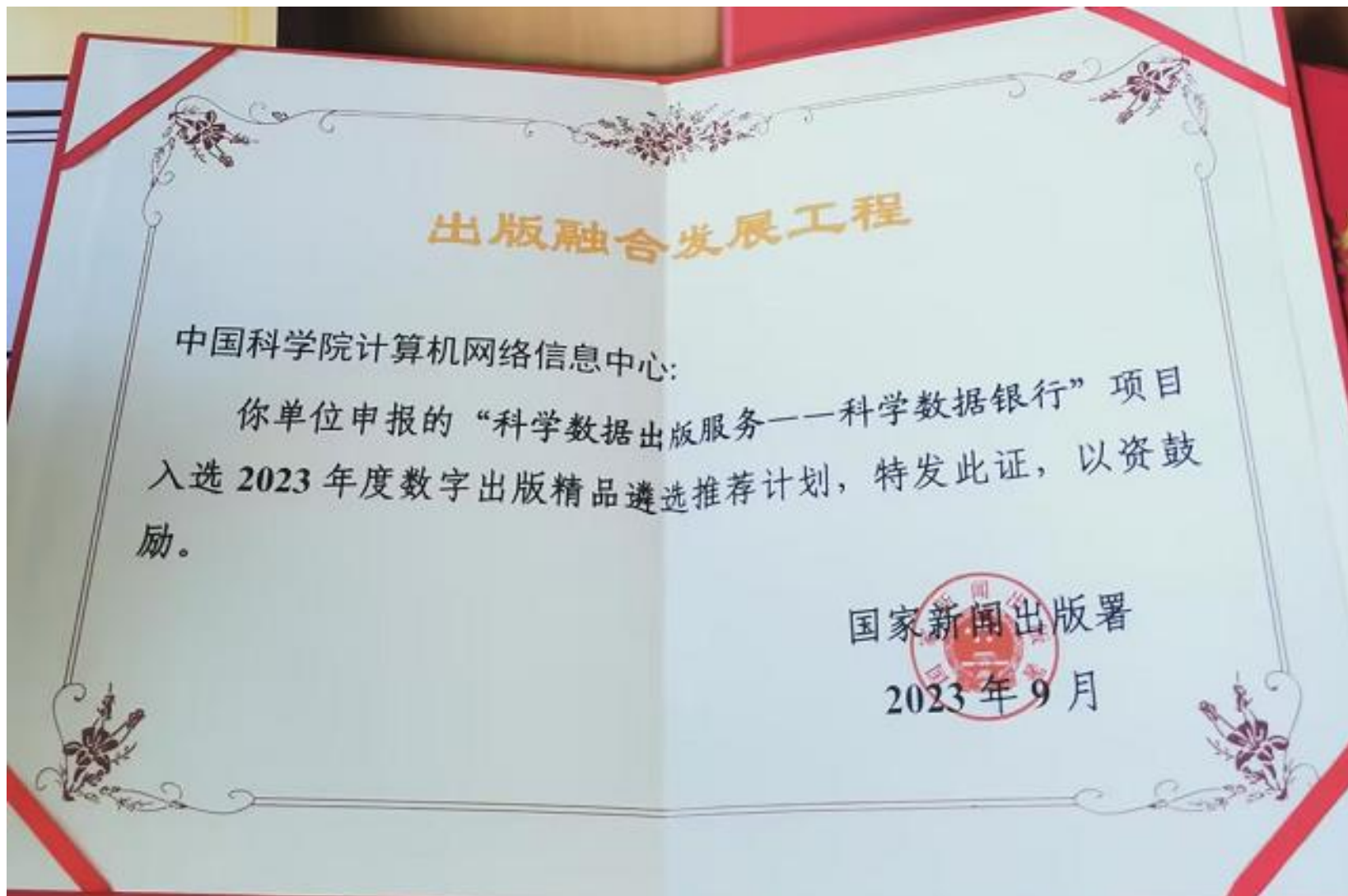
服务期刊/机构

5873

■ 相比2022年

➤ 数据作者国家数量增长40%+

➤ 访问用户的国家数量增长30%+



入选国家新闻出版署 出版融合发展工程  
2023年度数字出版精品

# 报告提纲

---

一、平台定位与目标

二、平台建设成效

三、年度功能更新

# 三、年度功能更新

## ➤ 新版本

突出社区功能，优化社区模版和检索功能

## ➤ 生态化

同论文采编平台，机构知识库，数据中心进行一系列集成

## ➤ 智能探索

探索AI技术在数据出版平台应用

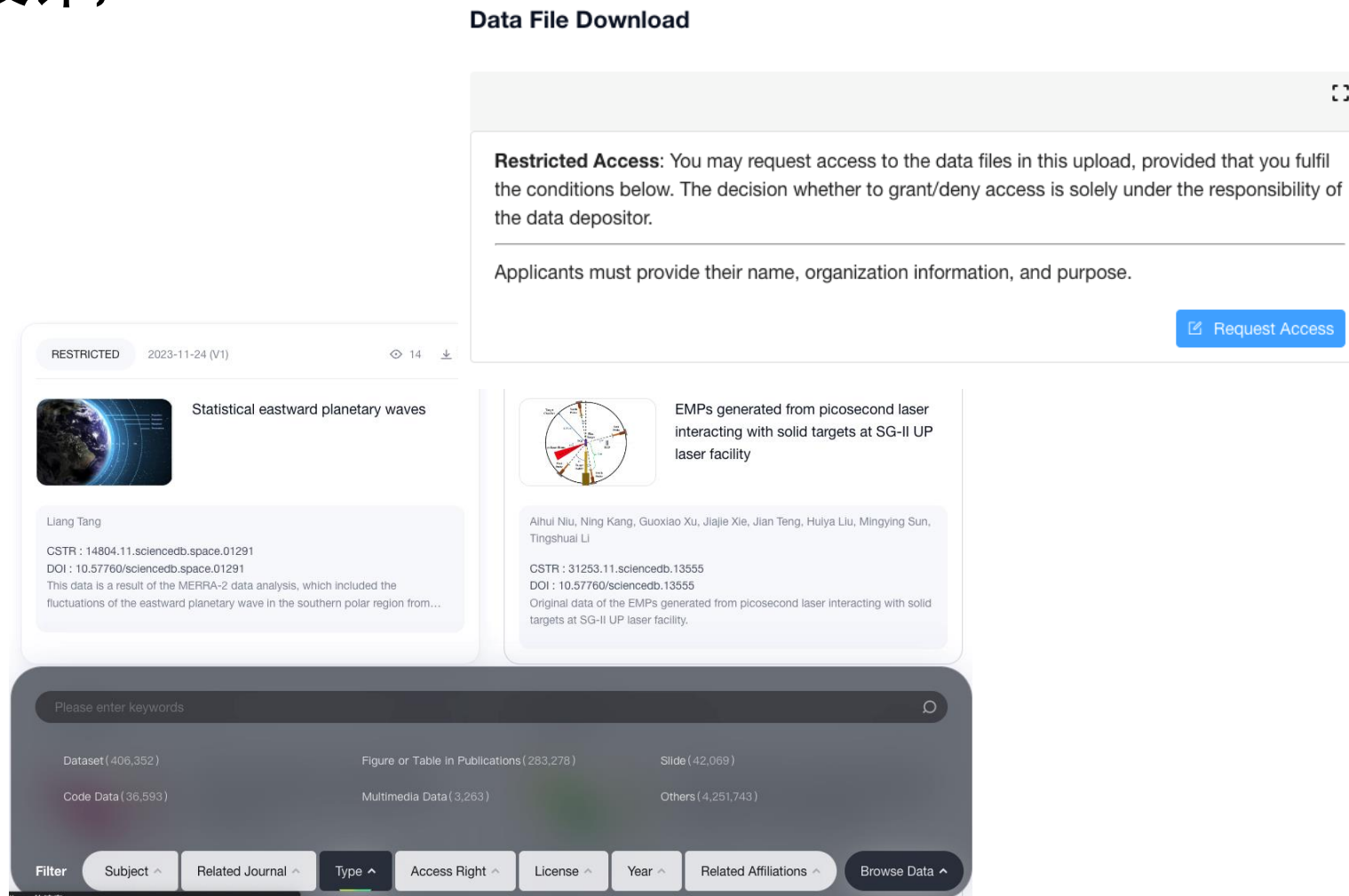
## ➤ 架构深耕

深耕优化升级文件存储，版本控制等技术方向

# ① 新版本

保留原版本所有功能，优化UI设计，  
新增多种功能服务

1. 定制化检索
2. 限制性获取模式
3. 灵动岛过滤
4. 期刊数据政策定制工具
5. 科学数据链认证
6. 文件粒度共享机制
7. FTP访问申请和计量统计
8. 帮助中心
9. 反馈助手
10. 数据集打包
11. 社区列表
12. 机构数据管理
13. Github第三方导入



## ② 生态化

### ■ 标识体系

CSTR, DOI, ORCID, ROR

### ■ 机构库

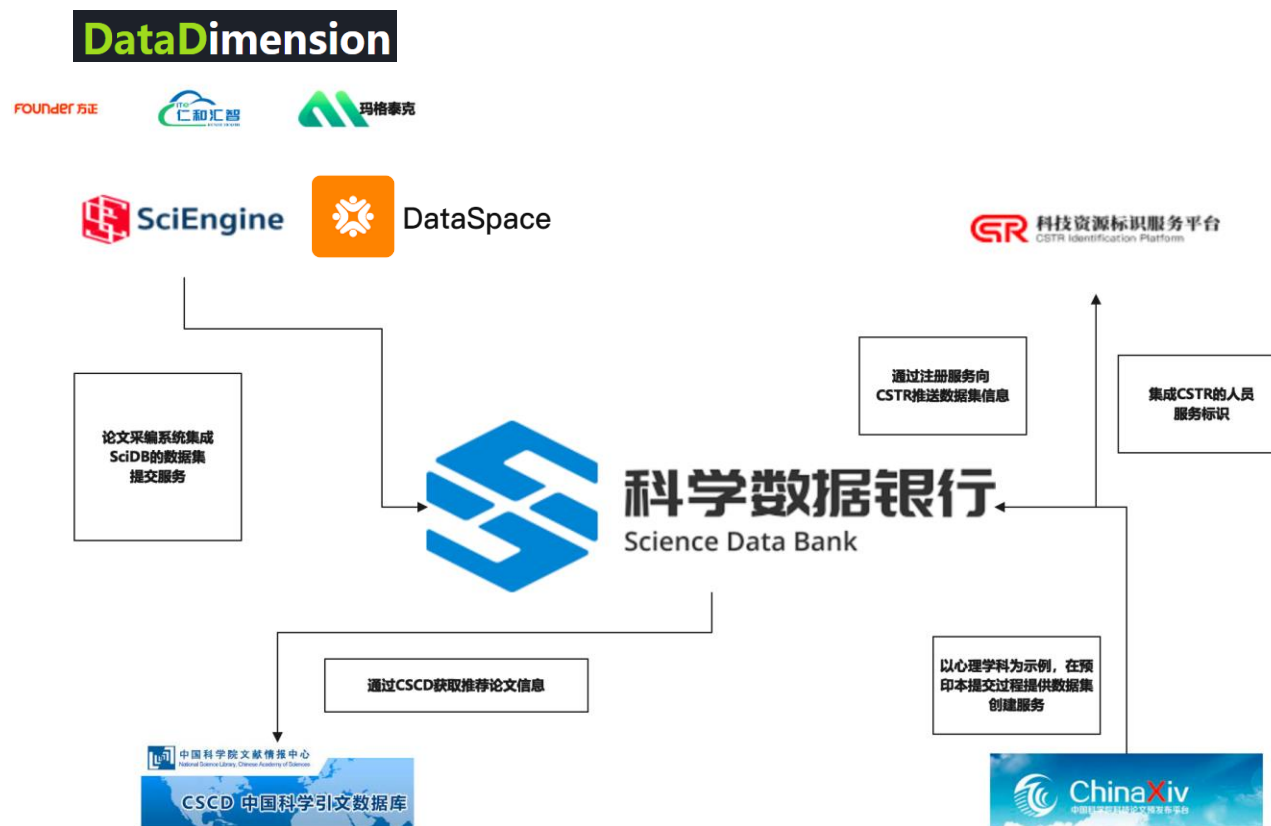
Dataspace, DataDimension

### ■ 论文采编平台

SciEngine, 方正鸿云, 仁和, ChinaXiv

### ■ 学术平台

CSCD, PubScholar



### ③ 智能探索



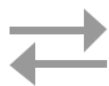
GB学科分类  
辅助判别

数据集阅读  
推荐

数据集集关  
键词生成

数据集封面  
图辅助生成

Scopus数据  
清洗



### Azure 的大语言模型服务



Embedding

GTP-3.5-turbo

DALL.E

## ④ 架构深耕

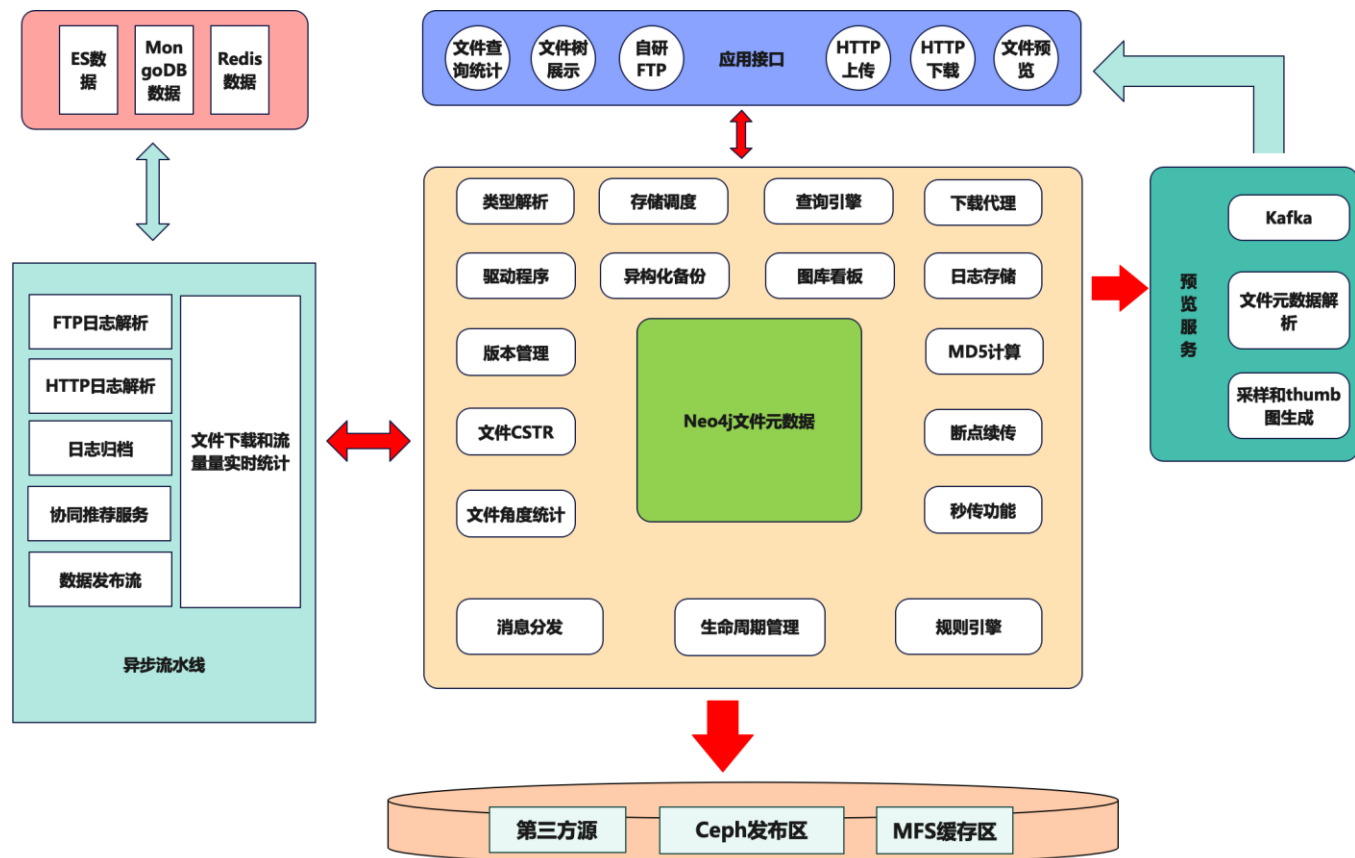
问题场景: 数据集实体文件版本管理 大体量的数据等突出点

原理: 简化一下

底层技术: Golang/图数据库

成果效应:

1. 较之前方案节约存储空间30%
2. 实现数据集版本控制和文件粒度版本比较
3. 为整个ScienceDB未来扩展提供基础支撑, 可扩展多种异构的存储源
4. 专利: 一种基于图数据库的数据集文件版本控制方法和系统一作



## ④ 架构深耕

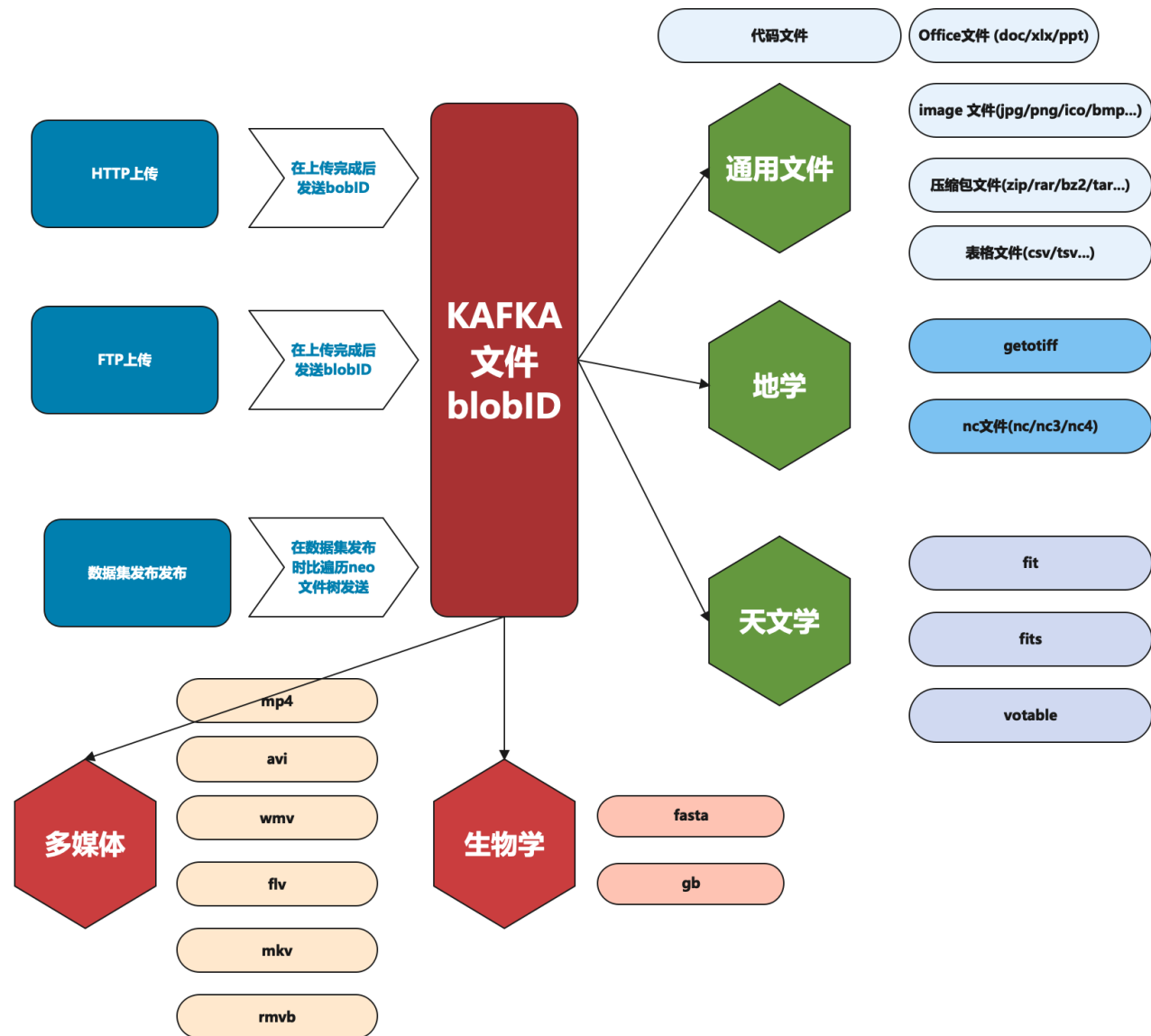
问题场景: 开源的预览组件支持文件有限, 性能低, 用户获取数据效率低

原理: 简化

底层技术: Python/Kafka

成果效应:

1. 70%的文件实现在线预览
2. 实现可持续扩展的科学数据文件预览服务
3. 专利提交中





中国科学院计算机网络信息中心  
Computer Network Information Center,  
Chinese Academy of Sciences

**感谢用户长期以来的信任和支持!**

