微信公众号

官网

中国科学院计算机网络信息中心
Computer Network Information Center,CAS

公共事务邮箱
nbsdc@cnic.cn

数据汇交邮箱
datarchive@cnic.cn

联系电话
010-58813784

地址
北京市海淀区东升南路2号院
中国科学院信息化大厦



国家基础学科公共科学数据中心

National Basic Science Data Center

2021 年度报告

目录

中心总体情况

中心概况 01
参建单位 03
组织体系 04
发展历程 05
数据资源 07
年度亮点 09
国际权威数据库 11
科学数据汇交 13
政策和标准体系 15
特色服务与数据软件 17
领域特色工具 19
服务体系 21
合作交流 23

数据资源与服务

科学数据库概况 25
交通运输分中心 27
世界山茶属植物品种数据库 29
机器学习基准图像集 31
环境健康主题数据库 33
中国海岸带植物资源数据库 35
医学物理数据库 37
中国动物主题数据库 39
中国植物主题数据库 41
中国植物DNA条形码数据库 43
化学主题数据库 45
原子分子物理主题数据库 47
光学技术数据库 49
核物理主题数据库 51
宇宙学数值模拟数据库 53
射电天文与深空探测数据库 55
载人航天空间科学与应用数据库 57
材料科学基础数据库 59
干细胞与代谢疾病数据库 61
中国活体人脑成像共享数据库 63
地理空间数据云 65
青海湖区域综合研究专题数据库 67
干旱区药用资源数据库 69
黑龙江省特色基础科技数据库 71

附录

年度大事记 73
典型数据库/集 75
2021年度代表性论文 80
运行保障 81

中心概况

国家基础学科公共科学数据中心（简称“国家基础数据中心”）是中国科学院计算机网络信息中心联合中国科学院、教育部、交通运输部、工信部、国防科工局、国家林草局、黑龙江省等下属研究所共40余家单位共同建设的国家科技资源共享服务平台，旨在汇集管理物理、化学、材料、动物、植物、交通、信息科学等基础学科领域，青海湖、黑龙江、新疆等典型区域长期科研活动积累的科学数据，以及相关领域政府预算资金支持项目汇交的科学数据。经过多年的持续发展，国家基础数据中心已经形成了完善的基础学科数据资源体系，以及分布式科学数据资源统一管理、集成融合、分析挖掘和应用服务的技术体系、标准体系和服务体系，有力支持了数据驱动的基础学科原始创新、多学科交叉融合创新、国民经济发展、人民生命健康和国际交流。2021年入选国家“十三五”科技成就展、中国科学院“十三五”信息化成果展。

目标定位

基础学科数据资源汇聚中心

科学数据自主软件研发基地

交叉学科融合研究分析平台

数据科学理论研究实践基地

参建单位

中国科学院
计算机网络信息中心

北京师范大学

北京大学

中国科学院
新疆生态与
地理研究所

吉林大学

中国原子能
科学研究院

西南大学

中国科学院
生物物理所

黑龙江省科技资源
共享服务中心

北京应用物理
与计算数学研究所

东北农业大学

黑龙省农业科学院
土壤肥料与环境资源
研究所

国家纳米材料中心

中国科学院
长春光学精密机械与
物理研究所

中国科学院
植物研究所

中国科学院
动物研究所

中国科学院
昆明植物研究所

中国科学院
过程工程研究所

中国科学院
紫金山天文台

中国科学院
武汉植物园

中国科学院
心理研究所

中国科学院
上海天文台

中国科学院
广州生物医药与健康
研究院

中国科学院
武汉病毒研究所

中国科学院
空间应用工程与
技术中心

中国科学院
成都生物研究所

中国科学院
上海有机化学研究所

中国科学院
金属研究所

中国科学院
新疆理化技术
研究所

中国科学院
水生生物研究所

中国科学院
长春应用化学研究所

中国科学院
华南植物园

中国科学院
合肥物质科学研究院

中国科学院
文献情报中心

中国科学院
西双版纳植物园

中国科学院
昆明动物研究所

中国科学院
化学研究所

中国科学院
城市环境研究所

青海湖国家级
自然保护区

中国科学院
烟台海岸带研究所

交通运输部
公路科学研究院

交通运输部
天津水运工程科学
研究院

组织体系

科技部

国家科技基础
条件平台中心

中国科学院

中国科学院
计算机网络信息中心

专家委员会

用户委员会

国家基础
数据中心

- 数据资源发展组
- 共享与服务组
- 数据标准组
- 技术与安全组
- 数据汇交组

专家委员会主任



孙九林 院士

专家委员会副主任



陈和生 院士



陈润生 院士



郭华东 院士



于贵瑞 院士



王赤 院士

国家基础学科公共科学数据中心
National Basic Science Data Center

发展历程

基础学科数据是夯实基础科学研究“地基”、构筑国家核心竞争力“大厦”的重要资源。国家基础数据中心是我国科学数据事业的主要建设者，立足17个一级学科数据资源，聚合基础数据、制定通用标准、研发共用技术、创新服务模式，推动基础学科数据开放共享与应用服务，引领大数据驱动的科研模式创新。

2021
信息科学分中心启动建设

2021
入选国家“十三五”科技创新成就展
同年，入选中国科学院“十三五”信息化成果展

2020
启动科技项目数据汇交
服务我国基础学科领域国家级项目数据汇交工作，促进科技资源标识和科学数据开放共享

2019
国家基础学科公共科学数据中心
《国家科技资源共享服务平台优化调整名单的通知》（国科发基〔2019〕194号）

2017
国家基础科学数据共享服务平台
《国家科技资源共享服务平台绩效考核与评估结果的通知》（国科发基〔2017〕24号）

2010
科技数据资源整合与共享工程
中国科学院“十二五”信息化专项

2009
基础科学数据共享网 – 理化天文空间生物（一期）
国家科技基础条件平台建设项目

2003
基础科学数据中心
“国家科学数据共享工程”重点项目

2001
科学数据库及其应用系统
中国科学院“十五”信息化建设重大专项

2021
服务800多个国家重点研发计划项目数据汇交

2021
中国人工智能协会共建多语种信息资源
联合人工智能学会多语种智能信息处理专业委员会和《中国科学数据》，推出多语种智能信息处理数据专刊

2021
交通运输分中心成立
交通运输部公路科学研究院、交通运输部天津水运工程科学研究院等单位联合建设

2019
中国科学院基础学科公共科学数据中心
《中国科学院科学数据数据中心体系第一批科学数据中心名单》

2017
GB/T 35294–2017 信息技术 科学数据引用
GB/T 34945–2017 信息技术 数据溯源描述模型
牵头发布两项国家标准

2016
科学大数据工程
中国科学院“十三五”信息化专项

2013
基础科学数据共享网 – 理化天文空间生物（二期）
持续发展资源运行服务

2006
数据应用环境建设与服务
中国科学院“十一五”信息化专项

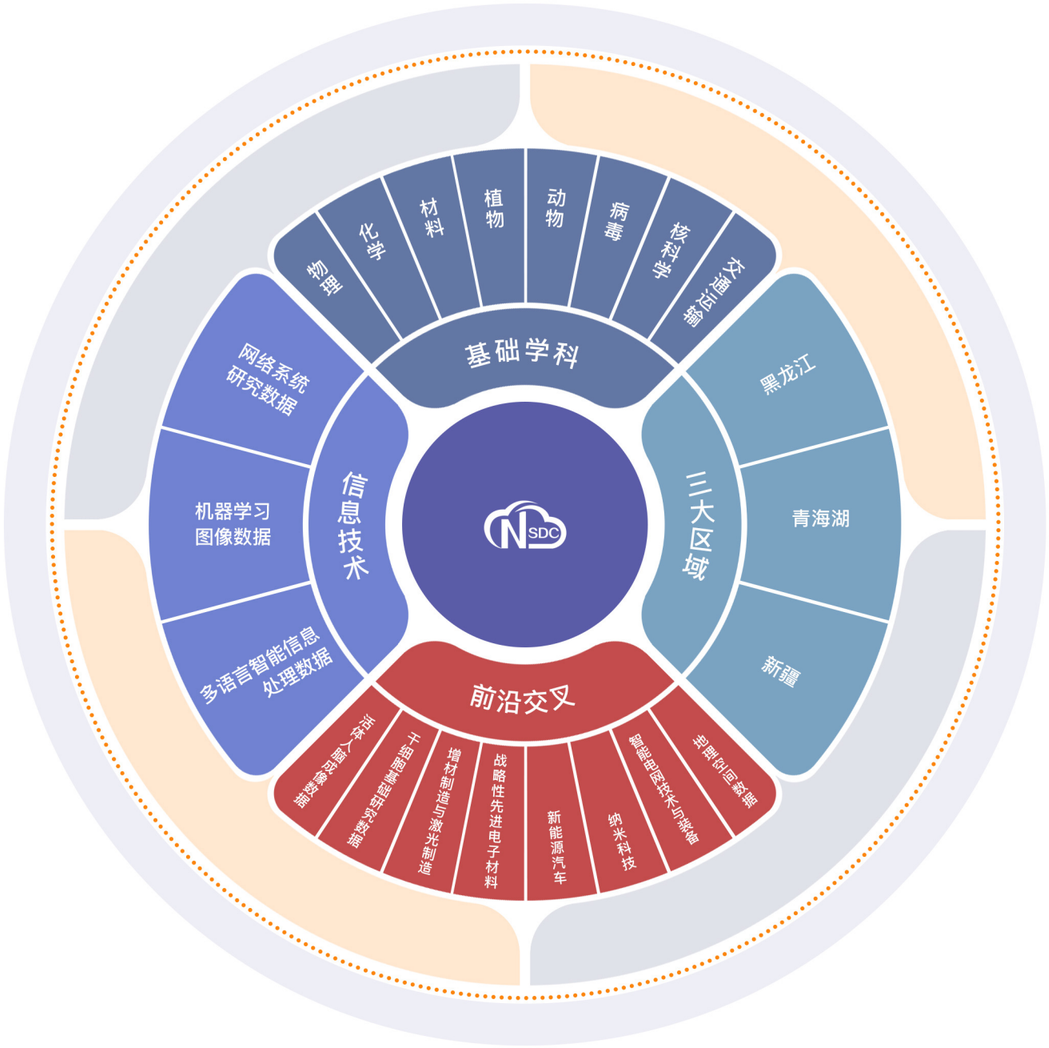
2004
基础科学数据中心及共享服务网
国家科技基础条件平台重点项目

2002
CODATA中国理化数据库
国家科技基础性工作专项

数据资源

国家基础学科公共科学数据中心
National Basic Science Data Center

国家基础数据中心系统整合了物理、化学、材料、动物、植物、交通和信息科学等学科领域，以及青海湖、黑龙江和中亚地区等典型区域长期科研活动中积累的基础科学数据。2021年度，国家基础数据中心继续汇集多个学科数据资源。目前，数据资源覆盖17类一级学科，数据总量2.29PB，包括23个主题库，2373个数据集。自建数据资源数据总量累计达1.13PB ，通过项目数据汇交科学数据量达1.16PB，数据记录10亿余条，并在完善的技术体系、标准体系支撑下构建了基于云门户的“数据服务–信息服务–知识服务”多元化服务体系，有效促进了基础科学数据的资源汇聚、开放共享、多学科交叉融合分析和应用。



17类
一级学科

23个
主题数据库

2373个
数据集

2.29PB
数据总量

10亿条+
数据记录

42个
参建单位

序号	分中心	承建单位	参建单位
01	交通运输分中心	交通运输部公路科学研究院	交通运输部天津水运工程科学研究院

序号	数据库	承建单位	参建单位
01	世界山茶属植物品种数据库	中国科学院昆明植物研究所	
02	机器学习基准图像集	中国科学院合肥物质科学研究院	
03	环境健康主题数据库	中国科学院城市环境研究所	
04	中国海岸带植物资源数据库	中国科学院烟台海岸带研究所	
05	医学物理数据库	吉林大学	
06	中国动物主题数据库	中国科学院动物研究所	昆明动物研究所　上海生命科学研究院 成都生物研究所　武汉水生生物研究所
07	中国植物主题数据库	中国科学院植物研究所	中国科学院文献情报中心
08	植物园主题数据库	中国科学院武汉植物园	
09	中国植物DNA条形码数据库	中国科学院昆明植物研究所	
10	化学主题数据库	中国科学院过程工程研究所	中国科学院上海有机化学研究所 中国科学院长春应用化学研究所
11	原子分子物理主题数据库	北京应用物理与计算数学研究所	
12	光学技术数据库	中国科学院长春光学精密机械与物理研究所	
13	核物理主题数据库	中国原子能科学研究院	
14	宇宙学数值模拟数据库	中国科学院紫金山天文台	
15	射电天文与深空探测数据库	中国科学院上海天文台	
16	载人航天空间科学与应用数据库	中国科学院空间应用工程与技术中心	
17	材料科学基础数据库	中国科学院金属研究所	中国科学院化学研究所 国家纳米科学中心
18	干细胞与代谢疾病数据库	中国科学院广州生物医药与健康研究院	
19	中国活体人脑成像共享数据库	北京师范大学	中国科学院心理研究所　北京大学 中国科学院生物物理所　西南大学
20	地理空间数据云	中国科学院计算机网络信息中心	
21	青海湖区域综合研究专题数据库	中国科学院计算机网络信息中心	中国科学院动物所 中国科学院武汉病毒研究所 中国科学院水生生物研究所 青海湖国家级自然保护区管理局
22	干旱区药用资源数据库	中国科学院新疆理化技术研究所	中国科学院新疆生态与地理研究所
23	黑龙江省特色基础科技数据库	黑龙省科技资源共享服务中心	东北农业大学 黑龙江省科信科技信息中心

年度亮点

科学数据作为国家战略资源，能够有力支撑科学研究、科技创新。国家基础数据中心持续致力基础学科领域科学数据宏观管理与发展，2021年，先后成立交通运输分中心、新增五个优秀科学数据库、启动信息科学分中心建设、入选国家十三五科技创新成就展、中国科学院十三五信息化成果展，联合学术组织、期刊、数据平台构建数据出版社区等。

1. 成立交通运输分中心

2021年10月，国家基础数据中心与交通运输部公路科学研究院在第七届中国科学数据大会上共同签署合作协议，正式建设国家基础学科公共科学数据中心交通运输分中心。



2. 新增五个优秀科学数据库

国家基础数据中心面向全国公开遴选优秀科学数据库。2021年4月–6月，历经课题申请征集、初审、专家现场评审等环节，世界山茶属植物品种数据库、机器学习基准图像集、环境健康主题数据库、中国海岸带植物资源数据库、医学物理数据库5个数据库入选国家基础数据中心。

– 遴选流程 –

2021–04–28	面向全社会公开发布课题指南
2021–06–10	专家组初审
2021–06–18	专家会议评审
2021–06–20	公布入选的5个优秀科学数据库
2021–07–12	召开新建科学数据库工作交流会

3. 入选国家十三五科技创新成就展、中国科学院十三五信息化成果展

2021年10月，国家基础数据中心入选国家“十三五”科技创新成就展。在“基础研究展区”中重点展示了 “数据服务–信息服务–知识服务”三级服务体系、国家重点研发计划项目科学数据汇交进展、以及服务科研创新的典型案例。凸显国家基础数据中心在我国“加强基础前沿研究、提升科技源头创新能力”进程中取得的重大成果。12月，国家基础数据中心入选中国科学院“十三五”信息化成果展。





4. 联合学术组织、期刊、数据平台构建数据出版社区

- 联合中国人工智能学会设立数据论文专刊

为推动人工智能多语种优质数据资源的持续汇聚、传播共享、评测服务、技术交流、国际合作与可持续发展，国家基础数据中心联合人工智能学会多语种智能信息处理专业委员会推出多语种智能信息处理数据专刊，正刊将于2022年出版。

- 联合《中国科学数据》推出前沿学科领域数据论文专刊

联合《中国科学数据》推出《中国活体人脑成像数据集专刊》，致力于打造具有国际影响力的中国活体人脑数据集，提升我国在神经影像领域的国际影响力，正刊将于2022年出版。

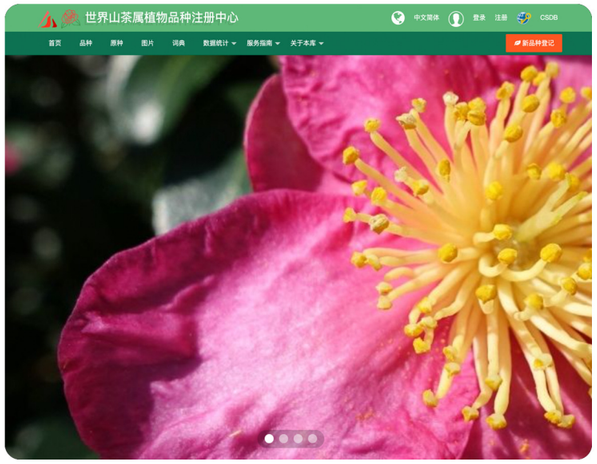
- 联合科学数据银行建设论文关联数据社区







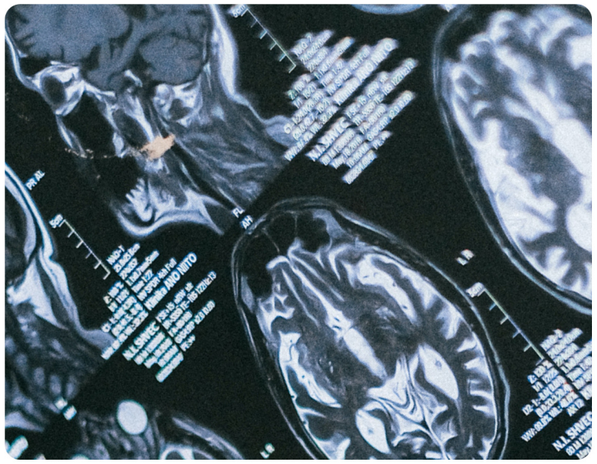
国际权威数据库



世界山茶属植物品种数据库

ICR surpasses by far other camellia databases, not only in terms of quality, but also in terms of the number of il-lustrations and descriptions.

—— Gianmario Motta，President of International Camellia Society



中国活体脑成像共享数据库

BRAIN projects will follow the two no-table big brain data projects (HCP and CoRR) to make tools and data avali-able to the community.

—— Thomas Insel，Founder of Mindstrong Health,Former Director of The Nation al Institute of Mental Health



核物理主题数据库

The new updated CENDL-3.2 is remarkable in those troubled times. I have had no processing errors in the making of such application library forms from CENDL-3.2, unique in my books.

—— Jean-Christophe Sublet，Unit Head(Data Services Unit)，International Atomic Energy Agency

科学数据汇交

国家基础数据中心充分利用依托单位中国科学院计算机网络信息中心基础设施软硬件和数据工作团队的综合优势，积极开展“政府预算资金资助的各级科技计划（专项、基金等）项目所形成的科学数据”的汇交，形成一套面向科研项目全生命周期的完整科学数据汇交管理解决方案。全面支持各级科技计划项目数据汇交，促进科学数据的开放共享与应用服务。



总体成效

主要服务专项信息

项目管理机构	所属专项	项目数(个)
中国21世纪议程管理中心（119个）	公共安全风险防控与应急技术装备	39
	绿色建筑及建筑工业化	21
	水资源高效开发利用	22
	大气污染成因与控制技术研究	14
	深地资源勘查开采	11
	...	
科学技术部高技术研究发展中心（594个）	重点基础材料技术提升与产业化	65
	新能源汽车	47
	增材制造与激光制造	53
	战略性先进电子材料	49
	重大科学仪器设备开发	48
	煤炭清洁高效利用和新型节能技术	42
	纳米科技	42
	云计算和大数据	33
	高性能计算	34
	智能机器人	12
	量子调控与量子信息	21
	现代服务业共性关键技术研发及应用示范	11
	蛋白质机器与生命过程调控	19
	先进轨道交通	12
	可再生能源与氢能技术	7
	综合交通运输与智能交通	7
	大科学装置前沿研究	6
	...	
工业和信息化部产业发展促进中心（82个）	智能电网技术与装备	34
	网络空间安全	15
	...	
中国生物技术发展中心（18个）	合成生物学	5
	数字诊疗装备研发	6
	生物医用材料研发与组织器官修复替代	4
	...	
中国科学技术交流中心（50个）	政府间国际科技创新合作重点专项	30
	战略性国际科技创新合作重点专项	20

截止2021年12月，国家基础数据中心为5家国家重点研发计划项目管理专业机构，累计41个重点专项863个项目提供科学数据汇交支撑服务，目前汇交科学数据总量超过1PB。

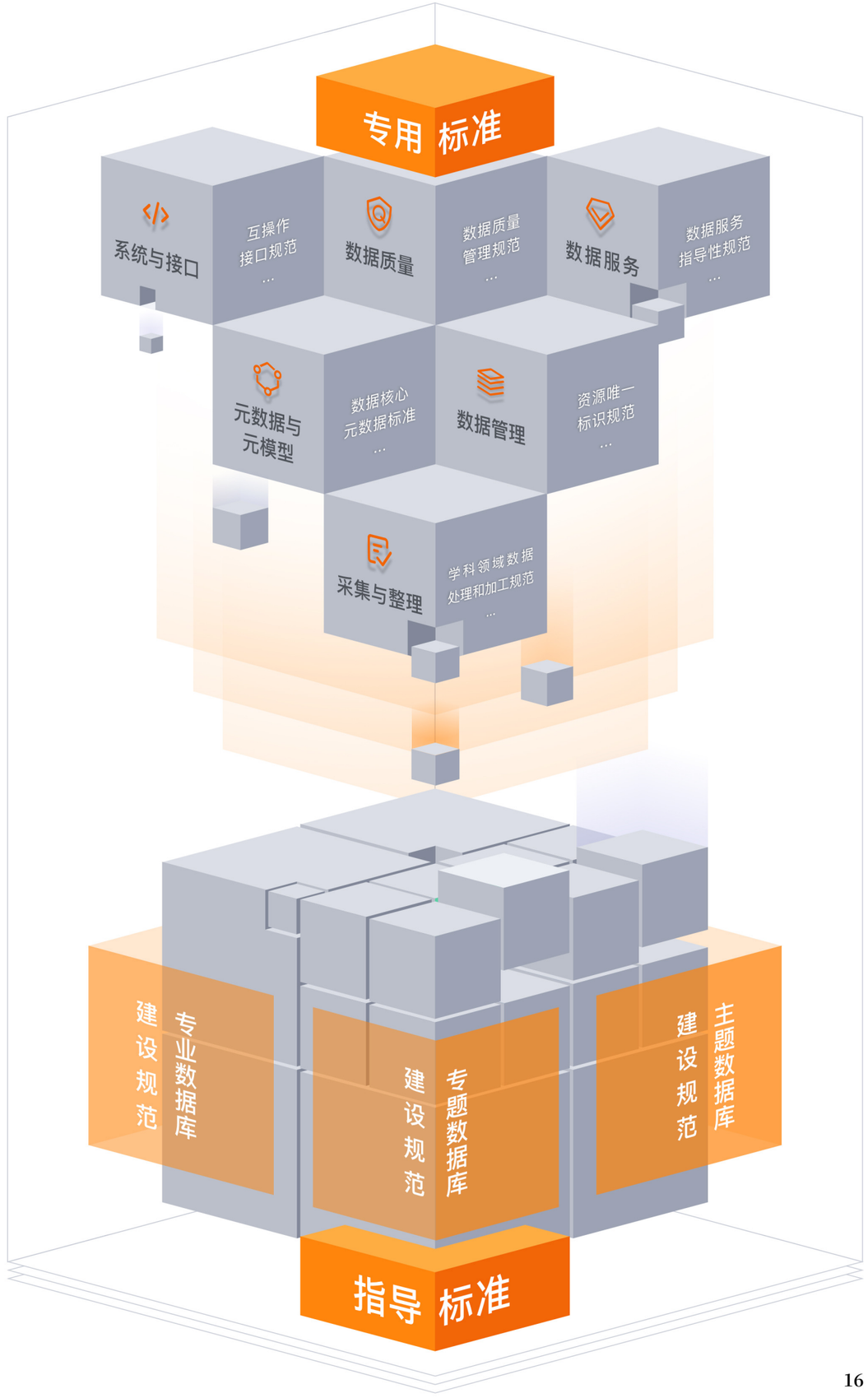
2021年度，国家基础数据中心通过线上/线下的方式组织和参加了涉及“新能源汽车”、“公共安全风险防控与应急技术装备”、“量子调控与量子信息”等多个国家重点研发计划项目重点专项共60余次科技计划项目汇交工作培训会。

政策和标准体系

为落实《科学数据管理办法》、《国家科技资源共享服务平台管理办法》、《中国科学院科学数据管理与开放共享办法（试行）》等政策文件要求，国家基础数据中心积极推动政策制度和标准体系的研究与制定，内容涉及政策制度、基础标准、数据安全、采集加工、汇交保存、共享利用六方面内容。

在政策制度方面，2021年初步形成国家基础数据中心政策体系。同时结合国家重点研发计划项目科学数据汇交工作实践，形成《国家基础数据中心科技项目数据汇交指南》，制定领域数据管理制度13项。在标准规范研制方面，构建国家基础数据中心标准体系，结合实际业务需要，重点制定了服务平台的3项数据资源元数据标准。年度牵头和参与研制国家标准4项，申请2项。





国家基础学科公共科学数据中心
National Basic Science Data Center

特色服务与数据软件

国家基础数据中心面向科学数据中心体系化、系统化建设管理的需求，研发系列工具，面向不同类型数据用户和应用场景提供分布式数据资源自主管理与服务、数据资源协同管理与外延服务。同时提供特色数据服务，强化数据产品加工和深度分析挖掘、服务国家重大项目和产业升级、充分发挥基础科学数据在我国双创战略实施中的基础支撑作用，提升我国基础科学数据资源的权威汇聚、长期保存、开放共享、分析挖掘、应用服务和驱动创新的整体能力和国际影响力。

专题服务

国家基础数据中心充分发挥多学科的优势，全面支持国家重大发展战略与重点区域发展，特别面向京津冀、长江经济带、粤港澳大湾区、一带一路等重点区域提供多元化的数据服务、专题服务、信息服务、知识服务，为科研机构、高等院校、企事业单位与相关政府部门提供线上、线下相结合的立体服务、跟踪服务和主动服务，取得良好的经济成效。

- 自主关键技术服务
- 一带一路服务
- 重大战略服务
- 大湾区服务
- 交叉学科服务
- 基础地理服务

数据服务

国家基础数据中心通过提供数据集下载服务、数据众包服务、数据集收录和出版服务、以及数据工具和数据标准服务，大力促进数据分析挖掘、交叉融合应用，跨领域数据融合、数据与领域技术融合，促进不同生态系统研究与应用的合作和资源数据整合，实现跨学科、跨领域的合作研究，提升科学数据对科研创新活动的支持能力。

- 数据交叉融合应用
- 数据资源服务
- 数据库(集)收录
- 用户反馈
- 数据工具
- 数据标准



科学数据协同管理平台

http://ds.nbsdc.cn

为科研项目团队打造的一款数据协作空间支撑工具，提供安全、便捷的仓储式科学数据协同管理服务。

- 项目内部共享与协作
  - 便捷邀请成员赋予权限
  - 范围可控、多人协作
- 多种数据文件接入方式
  - 本地文件拖拽上传
  - 专属FTP上传数据
  - 添加网络位置将数据空间挂载至个人系统或计算环境



数据存储库平台

http://idb.nbsdc.cn

提供基础学科领域科学数据全生命周期数据管理服务，支持科技计划项目汇交数据和主题数据库等各类资源集中存储、在线审核与快速发布，提升数据可发现性、可追溯性和可访问性。

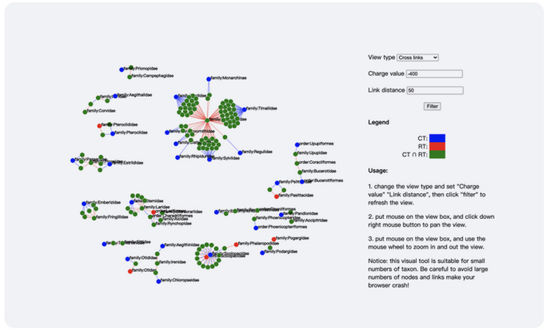
- 科学数据全生命周期管理
- 中心化统一管理与发布
- 第三方软件便捷集成
- 永久唯一标识符分配

领域特色工具

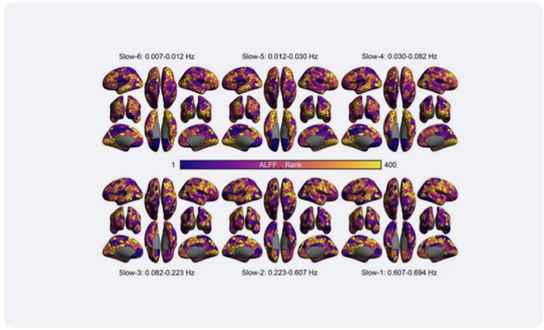
国家基础数据中心加强推动各学科领域数据加工整理、分析挖掘等方面的软件工具与算法研制，部分学科领域根据其学科特点开发系列特色工具，在主题数据库的管理、服务和应用中取得良好成效。

生物分类树工具

依托中国动物主题数据库开发的生物分类树比较分析工具平台，支持分类学家和分类学数据集成者建立、管理、分析比较、共享各种树状的生物分类系统数据。



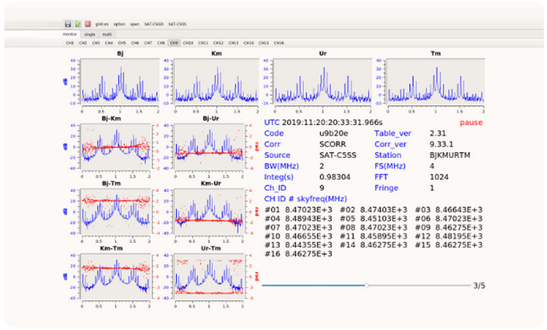
DREAM 人脑功能振荡信号多频段分解软件



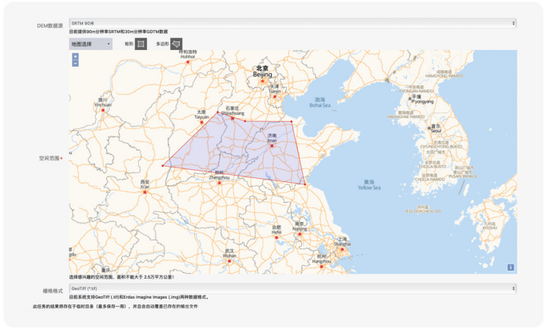
DREAM能够提高不同时空尺度上的神经振荡测量精度，增强可信性和有效性，从而揭示更多人脑功能。其基于人脑神经振荡的“自然对数线性率”等研发，经过数据处理后得到慢波多频段神经振荡幅度的空间次序，揭示了人脑功能个体差异随频段的梯度变化规律。

QuickView “VLBI快视”软件

软件能够在VLBI相关处理时，进行参数图形化显示，为数据的实时显示和分析提供了灵活的技术支撑，在探月工程嫦娥五号和首次火星探测天问一号VLBI测轨过程中发挥了重要作用。



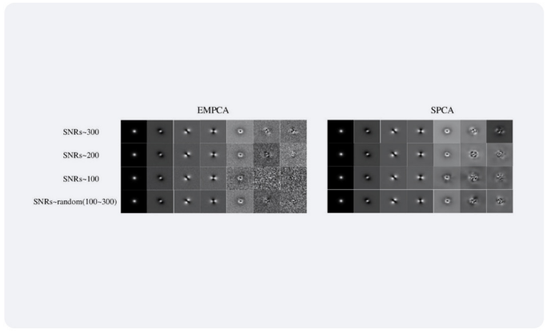
DEM 在线切割模型



地理空间数据云研发并在线集成“DEM在线切割模型”，为用户提供基于网络的简单、快速、交互式的在线获取、分析处理和可视化的先进工具，支持用户在线提交模型条件，实时运行并直接下载处理后数据结果。

图像处理光滑点扩散函数技术

技术结合了Moffat基函数和期望最大化主成分分析法的优点，从而实现了既光滑、正交性好，而又不过拟合的理想效果。该技术已在中国空间站工程巡天望远镜的科学数据处理管线中使用。



植物DNA条形码管理系统



植物DNA条形码管理系统通过本地部署对项目信息、样本信息、序列信息、测试图谱、文件、图片等多类型数据实现统一、标准化管理，确保数据库安全、稳定运行与服务。

服务体系

总体成效

国家基础数据中心门户“国家基础学科数据云”（www.nbsdc.cn）于2021年7月正式改版上线，基于G-FAIR原则，构建安全可管理的一体化数据基础设施，提供包括数据采集、数据发现、数据出版、技术服务、数据分析挖掘、图谱关联等数据全生命周期管理服务。截止2021年底，平台发布科学数据集2000余个，覆盖17类一级学科，具有典型学科多样性特点。同时平台实现对论文、专著、报告、软著、软件工具、标准、项目、机构等多种类别数据资源的在线管理和关联发布，逐步推动建设开放科学生态，有力支持了基础领域多学科数据的汇聚、融合和共享应用。

G-FAIR

可管理 Governable	数据汇交 数据安全 分级分类管理
可发现 Findable	数据标识 数据检索 关联推荐
可访问 Accessible	数据下载 数据服务 访问控制
可操作 Interoperable	元数据服务 数据接口
可重用 Reusable	数据出版 数据引用 版权声明

一体化基础设施

2021年度，国家基础数据中心依托中国科技云建成数据专有环境



高可靠、高稳定的基础设施运行环境	7×24小时实时监控 进口大容量UPS系统 国际最高标准机房 专业化的科学数据存储基础设施
高性能、高可靠、弹性扩展云存储环境	在线盘阵、近线盘阵和近线磁带库三级存储框架 服务器处理节点达到128个 正常运行率达99.9%
高速稳定的网络设施	基于IPv4、IPv6协议数据存储服务 中国科技网 网络总带宽达到10GB/S

合作交流

联合组织数据竞赛

2021年4月作为协办单位，支持北京大学组织“第三届全国高校数据驱动创新研究大赛”，有效推动了基础学科领域数据资源在教育领域的开放共享与融合创新。



加强宣传推广

以国家基础数据中心门户和官方公众号为核心构建宣传矩阵，促进前沿学术交流，开展科普教育，推动科学数据开放共享。

- 年度推送科学数据支撑成果、动态300篇次+
- 追踪国际前沿，推动数据伦理、数据互操作等理论研究
- 10+媒体联动宣传
  服务国内科研用户60万人+



中国科技资源共享网
官网、“锐共享”公众号

中国科学院官网

中国科学院
“科苑网信”公众号

中国科学院
“科苑数据”公众号

中国科学院计算机网络信息中心
官网、公众号

《中国科学数据》
官网、公众号

国家基础数据中心
官网、公众号

*各渠道新闻发布情况（以“第七届（2021）中国科学数据大会在呼和浩特市成功召开”新闻为例）

科学数据库概况

在2002年科技部科技基础性工作专项“CODATA中国理化数据库”和2003年国家科技基础条件重点项目“基础科学数据中心”的基础上，经过近20年的持续建设和发展，国家基础数据中心聚焦物理、化学、材料、动物、植物、信息科学、交通运输、重点区域、前沿学科等学科方向，以我国科学家在科学研究活动中获取的观测数据、探测数据、试验数据、实验数据、调查数据、考察数据、统计数据、模拟数据等为核心，建设自主数据库，目前共建有23个主题数据库，数据总量1.13PB，在支持科研创新、多学科交叉融合、国家战略、产业经济发展、全民健康等方面发挥了积极作用。

交通运输分中心

世界山茶属植物品种数据库

机器学习基准图像集

环境健康主题数据库

中国海岸带植物资源数据库

医学物理数据库

中国动物主题数据库

中国植物主题数据库

中国植物DNA条形码数据库

化学主题数据库

原子分子物理主题数据库

光学技术数据库

核物理主题数据库

宇宙学数值模拟数据库

射电天文与深空探测数据库

载人航天空间科学与应用数据库

材料科学基础数据库

干细胞与代谢疾病数据库

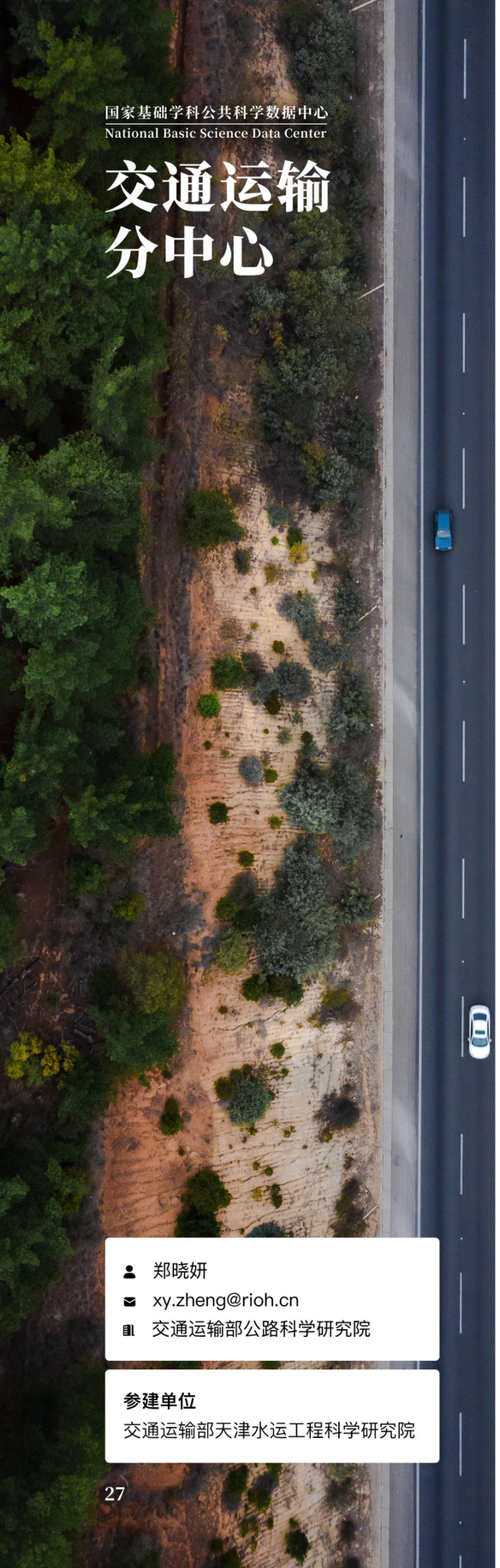
中国活体人脑成像共享数据库

地理空间数据云

青海湖区域综合研究专题数据库

干旱区药用资源数据库

黑龙江省特色基础科技数据库



国家基础学科公共科学数据中心
National Basic Science Data Center

交通运输分中心

郑晓妍
xy.zheng@rioh.cn
交通运输部公路科学研究院

参建单位
交通运输部天津水运工程科学研究院

交通运输分中心依托交通运输部公路科学研究院、交通运输部天津水运工程科学研究院等单位联合建设，主要建设公路科学数据中心和水运工程科学研究与应用数据库。

分中心汇聚交通行业的海量数据，覆盖了足尺环道、道路桥隧、路网监测、智能交通、客运货运等各个领域，形成了沿海主要港口设施维护管理数据、近岸水环境数据、环保监测数据、地下水监测数据等特色数据集，建设了一带一路、京津冀一体化等八大专题资源库。其中公路科学研究院建设的足尺环道是目前除美国以外唯一具备实车加速加载能力的足尺路面试验研究设施，基于该设施获取了5000万次实车加载作用下沥青路面全寿命、全时空、全环境下服役性能观测数据，数据具有全面性和权威性特点。相关数据在道路服役性能研究、出行服务与疫情防控、货运监测与冷链运输、车路协同与风险防控等实际应用中发挥了重要作用，支撑服务了多项国家重点研发计划项目，为水运领域重大工程建设、国家重大国家战略实施提供了有效数据基础。

3项
专著

1项
译著

50篇+
支撑论文

30项+
软件著作权

20项+
专利

40项+
支撑国家重点研发计划项目

典型案例

支撑我国长寿命道路服役性能研究验证大幅度提高我国路面长期性能

足尺环道服役性能观测数据支持发展沥青路面全寿命、全时空、全环境试验与观测技术，突破了长期以来依靠构建数学模型和智能算法才能验证道路服役性能的瓶颈，相关成果支撑了“足尺环道高通量响应数据挖掘与长寿命道路服役性能验证技术”（No.2017YFC0840201）等2个国家重点研发计划的实施工作，推动建立了道路全寿命周期多元服役信息平台，对于促进我国长寿命道路基础设施建养技术体系的形成发挥了重要作用。



支撑高寒高海拔西藏公路路面结构设计与养护维修助力脱贫攻坚

高寒高海拔地区气象恶劣，严重考验交通基础设施的服役性能、安全性能和耐久性能。公路院高原公路科技攻关团队对西藏1.2万公里在役公路的交通状况、路面结构、路面材料、施工方式与病害类别进行数据采集，建立国内首个高寒高海拔西藏公路路面病害基因库，并搭建了青藏高原混凝土气象观测站，持续监测气象环境和土壤参数。相关数据为藏区交通运输体系安全畅通和“四好农村路”建设提供了坚实保障，助力西部地区脱贫攻坚。



“水运工程科学研究与应用数据库”助力悬浮隧道基础理论体系研究

悬浮隧道是未来实现深海峡湾跨越的重大交通运输工程。科研团队基于水运工程科学研究与应用数据库，利用数学模型、物理模型、理论分析、设计评估等多种手段，开展水动力与结构、锚固系统、连接结构、工程材料、施工工法及装备等系列专题研究，实现了无人值守、高精度、智能化现场监测等，将局部冲刷坑的测量精度由传统的“分米级”提升至“厘米级”，极大提高了现有测量精度，推动水动力学、流固耦合、结构岩土、工程材料发展和研究，助力交通强国建设。



“水运工程科学研究与应用数据库”助力长江电子航道图升级完善

长江电子航道图作为船舶通航和航运管理的重要基础手段，在长江航运发展中具有重要地位。水运工程科学研究与应用数据库为“电子航道图APP”提供了气象数据、公众出行数据、船舶数据、航道沿线地物、沿途货源、多式联运等海量实时数据和历史数据，升级完善“电子航道图APP”，向航运企业、航行船舶、社会大众、港航管理单位等提供全线电子航道图、航道维护尺度、水位、航标等在线式不间断航道信息服务，提升长江港航参与者的满意度，有效推进交通强国和国家综合立体交通网建设，提供了长江电子航道图的不间断航道信息服务。



国家基础学科公共科学数据中心
National Basic Science Data Center

世界山茶属植物品种数据库

世界山茶属植物品种数据库涵盖了全球所有国家山茶属中的茶花、茶叶、茶油品种，发表和使用的名称超过5万条，图片数目超过5.7万张。数据库以中科院昆明植物所成员为核心，建立了全世界主要国家和地区协同的数据更新工作机制。在科学数据共享机制的促进下，数据持续不定期更新。注册用户可以上传照片、建立在线云相册，工作团队进行鉴定和认证工作，保证数据库的持续运维以及数据信息的权威性和完整性、使用与分享的便利性。经过多年努力，世界山茶属植物品种数据库已成为世界范围内信息全面，领域内最权威的山茶属植物品种数据库之一，为提升世界茶花、茶油、茶叶产业的健康发展提供积极支持。

25,918个
收录品种

50,850个
收录品种名称

199个
茶花原种

57,345张
茶花影像资料

王仲朗
wang@mail.kib.ac.cn
中国科学院昆明植物研究所


Camellia japonica 'El Beso'


Camellia japonica 'Tumida' Yves Chapel 2021


Camellia japonica 'Maria Otero del Rio'


Camellia japonica 'Tumida' Verschaffelt 1848

服务成效

标准	1项	世界山茶属植物品种库标准规范
论文	1篇	Wang Y, Zhuang H, Shen Y, Wang Y, Wang Z (2021) The Dataset of Camellia Cultivars Names in the World. Biodiversity Data Journal 9: e61646. https://doi.org/10.3897/BDJ.9.e61646
专著	1项	中文版《国际山茶会志》
软著	2项	世界山茶属植物品种注册平台 2019SR0845134 生物多样性电子图书馆系统 2020SR0719405
软件	1项	Camellia Dictionary

典型案例

广东从化阿婆六茶花谷引用世界山茶属植物品种数据库的二维码作为山茶花铭牌

世界山茶属植物品种数据库为国际杰出茶花园“广东阿婆六茶花谷”的国内外600多个品种、近4万株茶花介绍提供数据支撑，茶园通过采用“世界山茶属植物品种数据库”二维码作为铭牌，公众可在每个品种的详页上浏览和下载二维码图片。2020年4月，广东阿婆六茶花谷被国际山茶协会网络理事会认证为“国际杰出茶花园”，“国际杰出茶花园”是世界茶花业界的至高荣誉。



全面规范化处理茶花命名混乱、异物同名问题

数据库为山茶属植物资源的挖掘利用提供便利。日本是茶花生产利用大国，在束花茶花方面拥有很多古老品种，但是日文茶花品种中存在大量的不同品种但使用相同名称的问题。例如‘Asahi-zuru’等日本原产品种，现有日文书籍中品种的异物同名多达2200余个，世界山茶属植物品种数据库根据《国际栽培植物命名法规》做了全面且规范化的处理，使这个困扰茶花界百年的难题得以圆满解决，为世界茶花、茶油、茶叶产业的健康发展提供积极支持。



国家基础学科公共科学数据中心
National Basic Science Data Center

机器学习基准图像集

机器学习基准图像集面向跨媒体智能关键技术突破，聚焦农业本体图像识别、农业病虫害图像识别、农业知识图谱、农业知识的推理和演化等，由中国科学院合肥物质科学研究院承担建设。其中农业病虫害研究图库（IDADP）是机器学习基准图像集中的核心数据库，田间尺度作物病害图像数据资源超过500GB，作物病害图像涵盖了大田作物和典型设施蔬菜瓜果，为作物病害图像识别研究提供机器学习构建高质量的农业病虫害图像识别模型所需的样本数据，同时为农业领域的病虫害防治研究提供宝贵的基础数据资源，推动农业病虫害图像识别技术及其应用的发展。



袁媛
yuanyuan@iim.ac.cn
中国科学院合肥物质科学研究院

服务成效

1篇
数据论文
《IDADP-葡萄病害识别研究图像数据集》，中国科学数据

1个
专利
一种基于知识图谱的农业信息可视化交互式系统的构建方法. 发明专利, 申请号：202111322203.9

3项
软著
农业知识图谱问答系统（2021SR1433120）
基于深度学习的黄瓜病害识别系统（2021SR1312240）
生成式文本摘要系统（2021SR1433121）

典型案例

为全国农业科教云平台提供农作物病害图像数据和诊断技术

全国农业科教云平台由农业农村部科技教育司牵头开发，服务用户400万，涵盖农业专家、基层农技人员、农民等。机器学习基准图像集为全国农业科教云平台提供黄瓜、西红柿和水稻病害图像数据集、病害图像识别模型及支持多平台的WebAPI，目前手机端的稻麦病害诊断系统已经正式上线服务，为用户提供田间病害图像快速识别方式，提高了农业病虫害防治技术的智能化水平。



为“识农APP”提供病虫害诊断数据支撑

“识农APP” 是深圳市识农智能科技有限公司研发的一款软件产品，将人工智能技术应用于农作物病害、虫害、螨害、缺素症状的诊断以及杂草的识别。用户可通过“识农APP”直接拍摄或上传病害症状照片，2秒内获得对应病虫害的诊断反馈，并推送相关的解决方案。机器学习基准图像集为“识农APP”提供农业病害识别模型训练所需的水稻、葡萄等基础数据样本，简单切实地解决了困扰中国农业作物“看病难”，农技服务“无标准”的痛点。



国家基础学科公共科学数据中心
National Basic Science Data Center

环境健康主题数据库

环境健康主题数据库由中国科学院城市环境研究所承建。面向环境健康需求，汇集基于时空变化的环境污染数据，形成环境健康大数据平台，服务“美丽中国”和“绿色丝路”建设。目前已建有多个子数据库，其中生物体污染数据子库收录我国乃至一带一路沿线国家超过10万条生物体污染物的含量数据，环境声音数据子库收录来自超过2万民众的800多万条记录，抗生素抗性基因数据子库收录超过40万条的环境抗生素抗性基因分布数据。致力于贯彻可持续发展目标理念，为环境保护问题提供数据支撑，为目标区域的污染防治和人群健康提供科学依据。

10万条+
生物体污染含量数据

800万条+
环境声音数据

40万条+
环境抗生素抗性基因分布数据

吴扬雨
yywu@iue.ac.cn
中国科学院城市环境研究所

服务成效

2篇
数据论文

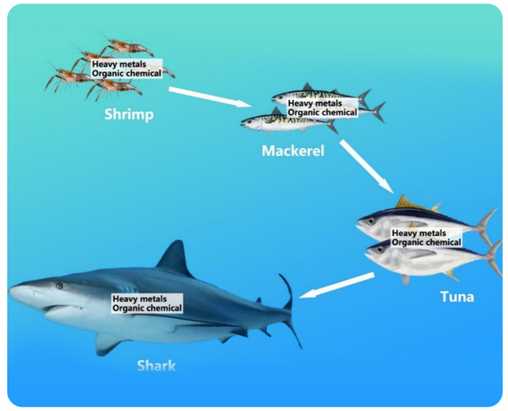
刘潇博，黄海宁，吴扬雨，黄乾生*. 中国沿海水生生物污染数据集及食用水产品的健康危害评估[J]. 生态毒理学报，2021, 16(3): 252-263. DOI: 10.7524/AJE.1673-5897.20200903002

Liu XB, Lin C, Wu YY, Huang HN, Zhu LT, Jiang R, Huang Q*. Dataset-based assessment of heavy metal contamination in freshwater fishes and their health risks. Environ Sci Pollut Res Int. 2022 Feb 28. doi: 10.1007/s11356-022-19427-0. Epub ahead of print. PMID: 35224696.

1项
软著

著作权人：中国科学院城市环境研究所 软件名称：物种—污染大数据平台（简称：DMMS）。登记号：2021SR0179177

典型案例



为中科院A类战略性先导科技专项“美丽中国生态文明建设科技工程”提供数据支持

生物体污染数据子库服务于“美丽中国生态文明建设科技工程”项目“粤港澳大湾区城市群生态建设工程与生态系统智能管理示范”，提供了珠江流域等重点区域的生物体污染物负荷清单，为生态修复和环境污染物治理提供了数据支撑；承建福建省“一带一路”环境健康联合实验室，汇集沿线国家生物体污染物清单，为“绿色丝路”建设提供数据支撑。



服务抗生素抗性基因传播认知和科普宣传

随着抗生素药物的发现及使用，越来越多的耐药菌株由此产生。通过对耐药基因的鉴定挖掘能够一定程度上帮助科研人员揭开耐药机制，为疾病的治疗、药物研发提供参考。抗生素抗性基因数据子库汇集了不同环境下的抗生素抗性基因分布及迁移特征，数据支持发表高引用论文10余篇。2021年，参与世界卫生组织确定的世界提高抗菌药物认识周(World Antimicrobial Awareness Week 2021)，通过数据展示和宣传科普提高全球对抗生素耐药性的认识。

中国海岸带植物资源数据库

中国海岸带植物资源数据库是一个以海岸带蓝碳为导向、主张海岸带植物资源保护与利用一体化、兼顾生态环境保护的公众科普为特色的数据库，由中国科学院烟台海岸带研究所、烟台大学、浙江海洋大学共同建设。按照用户需求分为高等植物、大型藻类和微型藻类三个子数据库，按照资源类型组织成四大模块，并链接至实体的种质库和标本库，以及非实体的基因数据库和专家知识库。相关数据填补了我国海岸带植物数据库方面的空白，在利用海岸带植物对滨海盐碱地土壤修复研究方面支持了解析菌根化田菁耐盐机制，对海岸带蓝碳研究和生态环保科普产生积极作用。

王寅初
ycwang@yic.ac.cn
中国科学院烟台海岸带研究所

服务成效

1篇 论文 《海洋生物产品服务平台的构建与应用》

1000人次+ 技能培训 累计开展海岸带生物资源利用技能培训1000余人次

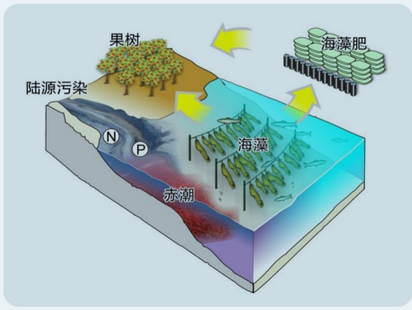
1项 软著 海洋生物产品服务平台 登记号：2019SR0516221

1020条+ 植物资源信息 采集海岸带植物资源信息1020条+

典型案例

为长岛和黄河口区域的海藻资源调查与海藻场产业化提供数据支撑

中国海岸带植物资源数据库长期为长岛和黄河口区域海藻保护利用一体化研究提供数据支撑。自2009年起对黄渤海海藻资源的变迁进行了连续的调查和分析，形成了中国海岸带藻类数据，通过聚焦天然海藻资源的保护与利用，开展长岛和黄河口区域的海藻场养护与多重生态系统构建工作，形成了4900亩示范区。同时在海藻植物资源产业化方面，开展了基于发酵工艺的新型海藻肥研发与应用工作，于2015年实现产业化，基于海藻发酵技术形成的海藻肥生产线已实现年产能30万吨。



为黄河三角洲生态修复提供适宜植物资源的数据支持

位于黄河的入海口山东省东营市的黄河三角洲湿地是世界暖温带唯一一块保存最完整、最典型、最年轻的湿地生态系统。受历年自然和人为因素影响，黄河三角洲湿地急需保护与修复。依托中国海岸带植物资源数据库以及黄河三角洲滨海湿地生物多样性信息系统，科研人员对黄河三角洲以及近岸植物600余条资源记录进行性状信息比较，选择鳗草在东营黄河三角洲周边近岸修复并构建了60亩以上的海草场，推动海岸带生态系统重建。山东电视台新闻联播、中央电视台“远方的家”栏目系列节目《行走海岸线》分别进行了报道。



医学物理数据库

医学物理数据库是我国首个医学物理交叉学科数据库，由吉林大学承担建设。该数据库覆盖核医学、放疗、医学辐射防护等方向所需的从原子核到器官尺度的微观物理、宏观剂量、生物效应及各类医学技术的数据，数据库主要包括五方面数据——重离子肿瘤治疗数据：各类重离子微观反应截面、宏观剂量和生物效应数据；X射线放疗数据：高能光子的微观反应截面及宏观剂量数据；同位素药物数据：^{32}P等治疗用和^{24}Na等诊断用的共50余种放射性药物的衰变数据和内照射剂量数据；医疗辐射防护数据：医疗进程中与术后辐射防护、放射性药物和仪器废料处理等公共卫生相关数据；医学影像等其他技术方向所需的相关数据。相关数据为我国的医学物理产业发展提供基础研究和服务支撑。

服务成效

- 我国首个服务于医学物理这一新兴交叉学科的基础数据库
- 已全面推广服务东北各三甲医院及六家高校的医疗和基础研究工作
- 文章：放射性同位素药物理特性数据集、MV法模拟质子治疗中靶区定位、MC法模拟X射线透射金属微孔板等

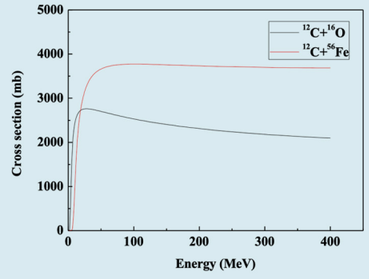
- 重离子肿瘤治疗数据
- X射线放疗数据
- 同位素药物数据
- 医疗辐射防护数据
- 医学影像数据

杨东
dyang@jlu.edu.cn
吉林大学

典型案例

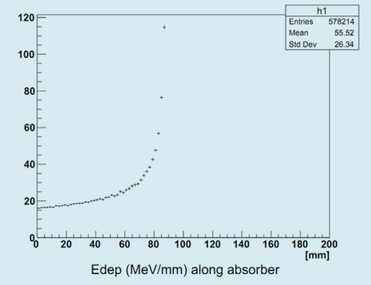
为安徽大学重离子肿瘤治疗微观数据的AI计算研究提供数据支撑

重离子与人体及机房仪器中的元素核反应的微观截面数据是重离子肿瘤治疗的重要基础数据。安徽大学牛中明教授2021年至2022年初利用AI技术对重离子肿瘤治疗中的微观截面数据问题进行了深入的研究，利用医学物理数据库重离子肿瘤治疗数据集中全套微观截面数据作为学习样本，有效解决该研究中学习样本的数据可靠性和准确性，助力推动重离子肿瘤治疗理论研究。



为长春工业大学重离子放疗的Monte Carlo模拟研究提供数据支撑

重离子放疗中的剂量分布、辐射防护等问题的研究，长期受困于研究成本、物理规则本身以及测量手段等诸多限制因素。吉林工业大学程道文副教授2021年6月至2021年12月针对重离子放疗剂量分布和辐射防护问题进行Monte Carlo模拟研究，对微观截面库数据准确性有高标准要求。医学物理数据库的重离子肿瘤治疗数据集中全套微观截面数据全面支撑了此研究，从而提升了模拟结果准确性。





国家基础学科公共科学数据中心
National Basic Science Data Center

中国动物主题数据库

林聪田
linct@ioz.ac.cn
中国科学院动物研究所

参建单位
昆明动物研究所
成都生物研究所
上海生命科学研究院
武汉水生生物研究所

中国动物主题数据库以动物学科领域数据为主体，依据国内外相关动物学研究成果进行建设，是目前国内动物学领域覆盖全面、内容科学的权威数据库。该数据库涵盖中国鸟类、兽类、两栖爬行动物、鱼类、昆虫及其它无脊椎动物。2021年度新增动物声音谱图数据库及濒危动物遗传资源数据库。截至2021年底，总计在线子库36个，服务数据量超过1.6TB，总记录数超过910万条，下载量达2TB左右，累计支撑超过20项科学研究项目及公众科普活动，对动物学研究、生物保护决策及科学普及具有重要意义。

36个
在线子库

1.6TB+
服务数据量

910万条+
总记录数

20项+
科研项目及科普活动

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	中国动物志数据库	7.84GB	0.5GB
2	地方动物志数据库	1GB	0.1GB
3	鸟类声音语谱图数据库	2GB	2GB
4	濒危动物遗传资源数据库	3GB	3GB

服务成效

支撑3篇论文

Xue, Y., Lin, C., Wang, Y. et al. Ecological niche complexity of invasive and native cryptic species of the Bemisia tabaci species complex in China. J Pest Sci (2021). https://doi.org/10.1007/s10340-021-01450-8（发表于昆虫学国际知名期刊Journalof Pest Science）

薛延韬; 王江宁; 林聪田; 韩艳; 丁晓庆; 纪力强. 基于动物志的鸟类形态特征数据集. 中国科学数据(中英文网络版), 2021, 6(03): 196-203

Wang, Zhaojun; Wang, Jiangning; Lin, Congtian; Han, Yan; Wang, Zhaosheng; Ji, Liqiang. Identifying Habitat Elements from Bird Images Using Deep Convolutional Neural Networks. Animals, 2021, 11(5): 1263.

支撑1个标准

中国动物分类代码 第1部分：脊椎动物

支撑2个项目

物种多样性信息平台
XDA19050200
中国科学院A类先导专项课题

濒危动物遗传资源保护及利用技术
2016YFC0503205
国家重点研发计划重点专项课题

典型案例

为《中国生物物种名录》2021版提供主要数据支撑

中国生物物种名录2021版是由中国科学院A类先导专项支持，主要收录中国已经发表物种的清单，涵盖动物、植物及微生物，包含物种56000个，种下单元4081个，是目前国内整合动物门类最多的名录。多年来，中国动物主题数据库持续为中国生物物种名录提供稳定、科学的数据源，为其定制多项服务功能，构建数据采集到名录发布工作流，为名录可持续发展奠定了坚实的基础。



支持中国科学院成都文献情报中心生物多样性调查

中国科学院成都文献情报中心学科咨询与服务部开展生物多样性调查，亟需西北部分地区的动物物种的统计情况。中国动物主题数据库按照用户需求提取西北部分动物名录及分布数据，通过比较、分析与合并等一系列工作，收集产出了多个目标县域的鸟类、兽类、两栖爬行类、昆虫等类群名录，数据包括纲、目、科、种及县级分布数据等，保证了该单位生物多样性项目的顺利实施。



国家基础学科公共科学数据中心
National Basic Science Data Center

中国植物主题数据库

中国植物主题数据库是以Species 2000中国节点和中国植物志名录为基础建设而成的综合型数据库。该数据库系统整合了植物彩色照片、植物志文献记录、化石植物名录、标本以及药用植物数据，建成27个数据子库及7个特色软件工具，通过免费共享的方式为公众、科研人员、企事业单位等各行业人士提供具有特色的中国植物数据服务。

中国植物主题数据库坚持以植物基础数据建设为核心。2021年，通过类群和分布地两个维度进一步填充植物数据空白区建设。

27个
在线子库

1013.58万条
总记录数

7个
特色软件工具

8个
支持科研项目

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	植物彩色照片库	1092651张	20000张
2	古植物标本照片库	9000张	4000张
3	省级植物名录数据库	9个	4个
4	文献志书数据库	102911本	100本

肖翠
xiaocui@ibcas.ac.cn
中国科学院植物研究所

参建单位
中国科学院文献情报中心

服务成效

中国植物主题数据库在建设植物数据的同时，通过线上、线下等不同形式开展服务。2021年度植物主题数据库支撑了京津冀地区物种全覆盖调查项目的开展，为植物相关的同行培训、研讨会提供数据支持，为科普活动提供知识依据。并支撑专著2部、学术论文6篇。

支撑的论文列表

[1]陈晓玲,唐丽玉,胡颖,等.基于ALBERT模型的园林植物知识实体与关系抽取方法[J].地球信息科学学报,2021,23(07):1208-1220.

[2]张斯琪,杨光.立足校园资源的地理项目式学习初探——以“校园植物”主题为例[J].地理教学,2021(13):33-36.

[3]徐振华,田玉童,王学勇,等.河北乡土树种应用现状与发展建议[J].河北林业科技,2021(02):40-43.

[4]钱灵颖,黄智洵,杨盛昌,等.厦门市重点保护植物空间优先保护格局研究[J].生态学报,2021,41(11):4367-4378.

[5]张秀华,齐海鹰,刘建,等.山东省珍稀濒危植物分布及其影响因素[J].山东大学学报(理学版),2021,56(05):42-50.

[6]樊信,盘金文,何嵩涛.气候变化背景下基于MaxEnt模型的刺梨潜在适生区分布预测[J].西北植物学报,2021,41(01):159-167.

典型案例

参与雄安新区植被特征与植物多样性调查服务国家建设需求

中国植物主题数据库经过一年的实地调查，全面梳理雄安新区植物多样性情况。雄安新区共有维管植物128科471属729种18杂种13亚种20变种1变型35品种。这些数据为以后本区域的生物多样性监测做好了基础数据库，同时给出了雄安新区建设中在植物造林、园林绿化、珍稀物种保护、外来入侵植物防治等方面需要重点关注的物种清单。为雄安新区的规划者和建设人员提供了物种参考。基于实地调查，出版《雄安新区植物图鉴》和《雄安草木行》两本专著。



支持京津冀地区物种全覆盖调查

植物主题数据库基于地区植物志和实际的野外调查，对京津冀地区实行全面的物种全覆盖调查，建立京津冀常见野生植物数据库，共收录京津冀地区植物7126种，是迄今为止收录植物种类最多的京津冀植物名录。调查发现，京津冀地区本土野生维管植物记录有2292条，对应143科703属2024个物种，其中：2072种及变种得到确认；增加了2个新种；2种确认为引入植物，而非原产植物；39种确认不产京津冀地区。为国家公园布局、生态环境保护工程的实施提供数据依据。



国家基础学科公共科学数据中心
National Basic Science Data Center

中国植物DNA条形码数据库

庄会富
zhuanghuifu@mail.kib.ac.cn
中国科学院昆明植物研究所

中国植物DNA条形码数据库是我国植物DNA条形码科学数据的汇聚和共享服务平台。其依托中国西南野生生物种质资源库的持续积累，开展以中国植物资源为核心的DNA条形码加工、采集和整合，共有中国植物DNA条形码数据17.7万条，覆盖中国植物2584属11911种，所有数据资源面向公众开放共享。平台提供植物DNA条形码检索和在线比对服务，基于DNA条形码推动新一代智慧植物志(iFlora)建设，为国家战略生物资源的精准鉴定、综合管理和相关科学数据服务提供基础数据和技术服务。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	中国植物DNA条形码数据库	65.91GB	0.7GB

服务成效

中国植物DNA条形码数据库建立了植物各大类群的DNA条形码技术规范体系，构建了中国最大的野生植物DNA条形码参考数据库和物种鉴定平台，相关工作显著提高了物种鉴定功能准确率，提升了中国西南野生生物种质资源库科学研究和公共服务的能力，成果丰硕。共支撑发表论文138篇（总影响因子210.5），其中SCI收录93篇，其中8篇代表性SCI论文被引409次，相关成果受到国内外同行的广泛关注和正面评价。先后培养DNA条形码专业技术人员30余人，持续为我国植物DNA条形码研究输送和培养骨干力量。

65.91GB
资源总量

4,849,435人次
在线访问总人次

64913.28MB
资源下载量

主要成果列表

论文

Fu CN， Mo ZQ， Yang JB , et al. Testing genome skimming for species discrimination in the large and taxonomically difficult genus Rhododendron[J]. Molecular Ecology Resources.　7.090/Q2

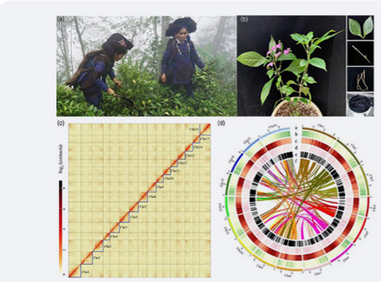
社会成效

助力实现物种的简易和准确鉴定，开发了多类群的DNA条形码技术，牵头组织我国生命条形码计划，使我国成为国际生命条形码计划的四个中心节点之一。

典型案例

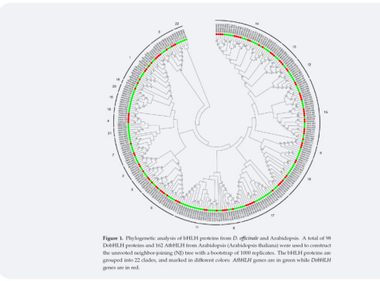
为靛蓝的生物合成分子基础及板蓝的驯化利用提供理论指导

靛蓝具有多种药物活性也是传统中药中的重要方剂，但业界关于其中重要组成靛蓝色素（indigo）的生理和分子基础知之甚少，解析其物质生物合成的遗传基础对板蓝的高效利用具有重要的指导意义。中国植物DNA条形码数据库开展了板蓝的精细基因组测序并揭示靛蓝生物合成的分子机制，该研究成果以“Blue genome: Chromosome-scale genome reveals the evolutionary and molecular basis of indigo biosynthesis in Strobilanthes cusia”为题发表在《The plant journal》杂志上，该杂志资深编辑Alisdair R. Fernie对该项研究给予高度评价。



支持植物医生研发中心铁皮石斛黄酮类化合物累积的基因家族分析

铁皮石斛含有多种次生代谢产物，研究发现具有抗氧化、抗衰老、增强免疫力等功效。中国植物DNA条形码数据库根据公开的铁皮石斛基因组数据，通过HMM比对的方法鉴定铁皮石斛中的bHLH家族基因，为后续解析铁皮石斛中黄酮类化合物累积的机制奠定基础，也为中国科学院昆明植物研究所与植物医生联合建立的植物医生研发中心开发含黄酮类化合物的铁皮石斛护肤品提供理论基础。



国家基础学科公共科学数据中心
National Basic Science Data Center

化学主题数据库

赵月红
yhzhao@home.ipe.ac.cn
中国科学院过程工程研究所

参建单位

中国科学院上海有机化学研究所
中国科学院长春应用化学研究所

化学主题数据库是国内目前规模最大、稳定在线服务的化学数据库，汇聚了中国科学院化学化工学科领域80%以上的基础数据资源，在线服务数据子库40余个，数据覆盖化学化工研究主要的二级学科；自主开发建设国内唯一的化合物登录系统，当前登录化合物860万余种，并以此为基础实现了基于化合物唯一标识的异地异构化学数据资源整合与共享；面向全球提供化合物全景式的化学科学数据检索展示、化学数据加工及面向项目的专题数据服务，国外用户的数量长期保持在总用户数的10%以上。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	药物天然产物数据库	623.1 MB	2千条晶体数据
2	化合物登录数据库	8,175MB	5万个化合物
3	化合物生物活性数据库	20MB	5万个化合物
4	中药代谢组学数据库	250MB	500条记录
5	废锂电池图片库	24MB	2000张图片

服务成效

数据总量为53.7GB，支撑论文37篇，对外服务支撑的各级各类科研项目/课题100余个，其中包括973、863及重大研发计划等国家重大项目总计14项，国家自然科学基金（重大研究计划、重点项目、杰青及面上项目）42个，省部级、地方和企业项目50余个。

主要成果列表

[1] Y.H. Ren, Q.Z.Han, Y.H.Zhao, et al. The exploration of metal–free catalyst g–C3N4 for NO degradation[J]. Journal of Hazardous Materials, 2021, 404: 124153.

[2] W.Y. Yan, Z.L. Wang, H.B. Cao, et al. Criticality assessment of metal resources in China[J]. iScience, 2021, 24(6): 102524.

[3] T.J. Xu, W.M. Chen, J.H. Zhou, et al. Computational Analysis of Naturally Occurring Aristolochic Acid Analogues and Their Biological Sources[J]. Biomolecules, 2021, 11(9): 1344.

[4] H.Y. He, F. Xu, Q.L. Li, et al. Separation of hafnium from zirconium in HNO3 solution by solvent extraction with Cyanex–572[J]. Hydrometallurgy, 2021, 202: 105600.

[5] X.L. Xu, Q.L. Xiao, Z.H. Pang, et al. Molecular mechanisms of An–Chuan Granule for the treatment of asthma based on a network pharmacology approach and experimental validation[J]. Bioscience Reports, 2021, 41(3): BSR20204247.

[6] F. Wang, J. Liu, L.P. Zhong, et al. Theoretical Analysis of the Decomposition Pathways and Species of Environmentally Friendly Insulation Gas C6F12O Based on the DFT and TST[J]. Plasma Chemistry and Plassma Processing, 2021, 41: 133.

[7] Y.L. Fan, X.Y. Yin. Potential therapeutic targets and biological mechanisms of Centella asiatica on hepatic fibrosis: a study of network pharmacology[J] Annals of Translational Medicine, 2021, 9(11): 932.

[8] Y.H. Ren, Q.Z.Han, Q.Z. Su, et al. Effects of 4d transition metals doping on the photocatalytic activities of anatase TiO2 (101) surface[J]. International Journal of Quantum Chemistry, 2021, 121(16): e26683.

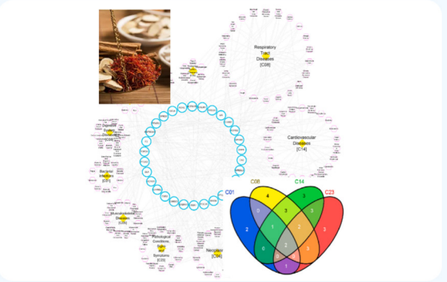
[9] 赵月红，谢勇冰，曹宏斌，等. 基于全过程污染控制策略的钢铁工业园区水网络全局优化基于全过程污染控制策略的钢铁工业园区水网络全局优化[J]. 过程工程学报, 2021, DOI: 10.12034/j.issn.1009 606X.220373.

[10] 李英勇. 中药材化学成分的晶体结构数据集[J]. 中国科学数据, 2021, DOI: 10.11922/11–6035.csd.2021.0052.zh

典型案例

为传统蒙药有效成分与作用机制的研究提供数据支撑

传统蒙药七味葡萄散广泛用于缓解咳嗽和呼吸困难，研究发现对人II型肺泡腺癌细胞的死亡具有较强的协同保护作用，通过对化合物登录库、植物化学成分数据库和化合物生物活性数据库的数据挖掘和跨库检索，获得了七味葡萄散的主要化学成分数据，并明确了芪甲苷治疗慢性阻塞性肺病的药理作用机制。研究成果发表在《JOURNAL OF ETHNOPHARMACOLOGY》，DOI:10.1016/j.jep.2020.113385。



服务中国金属资源关键性评价

基于收集的中国金属资源分布集中度、可替代性、进口依赖性、二次资源利用率、产量占比等基础数据，以及金属元素自身特有毒性、在采选及加工过程中释放到环境中的三废比重等专题数据，建立中国金属资源关键性评价数据集。基于此数据集提出并研发了中国金属关键性量化评价方法和软件，实现对于现有某个行业的关键元素从供给安全指数、国民经济指数和环境风险指数三个方面进行判定，综合分析元素对国家工业发展的重要意义，同时为政府及企业在发展策略的制定上提供理论指导。



国家基础学科公共科学数据中心
National Basic Science Data Center

原子分子物理主题数据库

原子分子物理主题数据库由北京应用物理与计算数学研究所建立，是目前国内原子分子领域唯一的数据库。原子分子数据在天体物理、武器物理、聚变能源（ICF、Tokomak）和工农业生产等领域都有重要应用，与国际上的IAEA原子分子数据组、俄罗斯spectr-w3数据库建有长期合作关系。该数据库现有原子光谱、能级、电离势、电子与原子分子碰撞截面等多个方面的科学数据200万余条，可全部在线检索，实现数据免费共享服务。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	原子光谱数据库	187万记录数	3.0万记录数

刘春雷
liu_chunlei@iapcm.ac.cn
北京应用物理与计算数学研究所

服务成效

原子分子数据在各领域应用广泛，尤其是天体物理等基础科学研究领域、聚变科学等高科技研发领域、照明科技和EUV光源等相关国民经济技术研发领域、以及战略武器研发等相关国防建设领域。

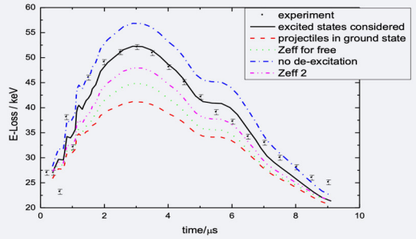
支撑发表了10余篇高质量学术论文，发表于The Astrophysical Journal Supplement Series (IF:12.203)、Atomic Data and Nuclear Data Tables (IF:2.165)等期刊。

支撑了国家重大科技专项“X基本参数研究”、科技部大科学装置前沿研究（2017YFA0403200）等重大科研项目的申请和研究。

典型案例

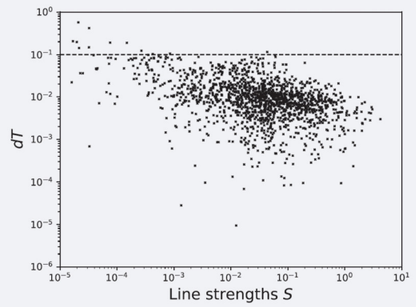
为中国工程物理研究院TB2403课题组提供碰撞数据支撑

带电粒子在等离子体中的能量沉积是原子物理与等离子体物理交叉学科中重要而又基本的研究课题，原子分子物理主题数据库研究H++Li、He2++H(nl)、 H+H、H++H(1s)等多个过程的碰撞激发、电离、电荷转移过程。原子分子物理主题数据库应用于α粒子的能损模拟，支持得到与实验相吻合的结果，为课题组进行更高精度的聚变理论模拟和分析、得到更精准的等离子体状态诊断结果做出了贡献。



为国家重点研发计划"高温高压高密度物质的压缩性质、不透明度和输运系数研究"提供数据支撑

Cr-IX~CUXIV离子在等离子体诊断中有着重要的应用，FeXI离子的谱线在日冕的研究中有重要应用，其发出的谱线在日冕的活动中被多次发现。原子分子物理主题数据库采用相对论多组态相互作用方法，研制并提供了Fe、Co等元素的类S原子数据，以及Ni、Cu等元素的类P原子数据，将这些数据应用于等离子体状态诊断程序，得到了更可靠的等离子体状态诊断结果，为项目课题更高精度的理论模拟和分析，得到更可靠的等离子体状态诊断结果做出了贡献。



国家基础学科公共科学数据中心
National Basic Science Data Center

光学技术数据库

光学技术数据库是目前国内唯一的光学领域数据库，包含有光学镜头数据、光学材料数据、光学系统常用国家标准和光电元器件库等公开数据。该库始于1987年中国科学院长春光学精密机械与物理研究所牵头建立的光学专业库（原名光学镜头数据库）。目前已建成5个主题库、21个子库、数据资源量640GB，下载量为120GB，包含结构数据以及各种像质评价结果数据和图形，所有数据均提供站点查询与下载。与光学设计软件集成两种方式提供数据共享服务，已为国内外上千家行业单位、数十万用户提供大量的科学数据，帮助科研人员取得了多项重要科研成果。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	光学玻璃数据库	10GB	5GB
2	光学晶体数据库	5GB	10MB
3	光源数据库	2GB	100MB
4	光电探测器数据库	5GB	205MB
5	光学镜头数据库	10GB	100MB
6	光学标准库	1GB	200MB
7	光学成像数据库	380GB	1GB
8	光学检测数据库	80G	500MB
9	特色光谱信息数据库	120G	4.1GB
10	激光器件数据库	2GB	109MB
11	激光传输特性数据库	1GB	104MB
12	激光光学系统数据库	1GB	100MB
13	短波光学数据库	1GB	60MB

特色光谱数据库

物质特征谱线数据集

- 378条 930MB
  松材变色立木样例数据
- 540条 1.3GB
  河道样例数据
- 384条 945MB
  果园全息监测原始数据
- 402条 980MB
  城市绿化率监测

娄洪伟
louhw@ciomp.ac.cn
中国科学院长春光学精密机械与物理研究所

服务成效

20项+
承担国家光学系统相关设计项目

5项+
支持国家重大项目

19篇
支撑论文

44个
支撑项目

875个
在线计算任务

185个
在线数据预定

50TB
累计下载量

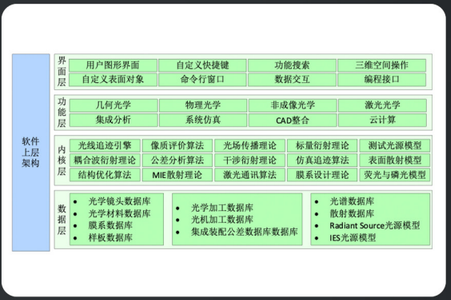
2021年度支撑项目

国产软件“卡脖子”工程项目	“国产化光学设计分析软件研制 ”
国家重点研发计划	“黑土地保护与利用科技创新”
国家重大项目	“激光创新中心”

典型案例

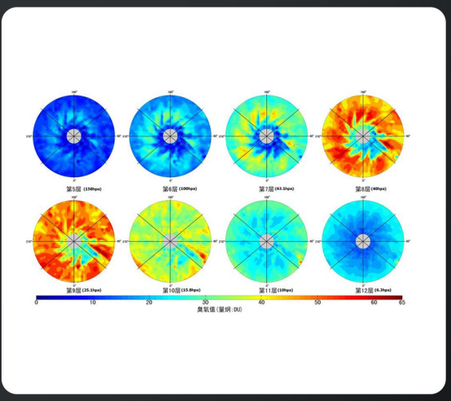
为国产化光学设计分析软件的研制提供光学基础数据

2020–2021年，我国在国产化软件工程项目中持续支持国产化光学设计分析软件建设，光学技术数据库为该软件研制提供了各类光学基础数据，如光学镜头数据、光学材料数据、膜系数据、样板数据、光学加工数据、装配工差数据、光源数据、散射数据等，同时提供了光学相关数学计算模型，解决了软件使用初始值问题，支持了软件的研制、推广，对我国核心软件自主化生产和生态形成做出了重要贡献。



支持大变倍比高清电动长焦监控用镜头产业化研究

目前国内监控企业生产的镜头以低端短焦产品为主，高清长焦镜头长期依赖进口。该项目目标是将大变倍比高清电动长焦监控镜头研制成果产业化，解决高清电动长焦镜头国产化问题。光学技术数据库支持研究的关键技术包括高清、大变倍比、宽光谱、体积小、重量轻、低成本、适合批量生产的镜头设计技术；长焦镜头加工、装调与检测等批产工艺技术。目前，“大变倍比高清电动长焦监控用镜头产业化研究”项目已经完成原理样机研制，后续将重点解决适合批产的设计和工艺等问题。

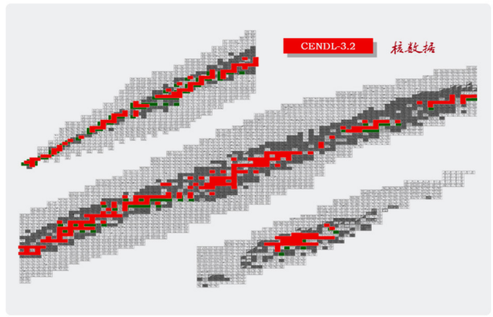


国家基础学科公共科学数据中心
National Basic Science Data Center

核物理主题数据库

金永利
jyl@ciae.ac.cn
中国原子能科学研究院

核物理主题数据库是核基础数据信息化合集，收录数据来自我国或者国际权威机构，主要包括中子、带电粒子等引起的反应数据、裂变产额、核结构和衰变数据等。这些数据是表征原子核自身特性和原子核与其它粒子发生反应的重要数据，是核基础研究、核装置设计与运行、核技术应用的重要数据。中国评价核数据库是我国自主研发、国际五大主流核数据库之一，具有较大的国际影响力。新版中国评价核数据库CENDL-3.2发布后，中国评价核数据库的核素数量由240种增加至272种，数据质量、数据种类均有大幅度提升，为我国重大核科学装置的建设提供了数据支撑。



2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	评价核数据库	834GB	0.5GB
2	实验核数据库	952MB	0.1GB

服务成效

837GB
数据资源量

7万
在线计算和绘图

347.32GB
累计数据下载量

3,506,872个
累计访问文件数

36个
科研院校申请专用账号

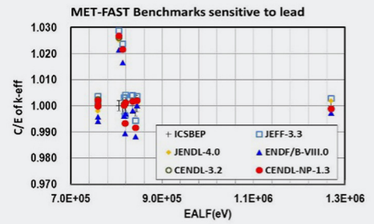
主要成果列表

序号	成果名称	引用
1	核反应实验数据库(EXFOR)	王记民, 陶曦, 金永利, 刘丽乐, 陈国长, 葛智刚, 大塚直彦. 核反应实验数据库(EXFOR)[J]. 原子核物理评论, 2021, 38(1): 107–115.
2	板状Bi样品泄漏中子和γ谱实验测量与模拟计算	赵齐;聂阳波;丁琰琰;吴海成;张环宇;任杰;阮锡超.板状Bi样品泄漏中子和γ谱实验测量与模拟计算[J].原子能科学技术2021, 55(4): 584–593
3	Measurements of neutron energy spectra of 9Be(d, n)10B reaction with a thick beryllium target	Zhang, Shuangjiao & Hu, Zhijie & Hu, Zhiming & Han, Chao & Bai, Xiaohou & Liu, Chang-Qi & Huang, Chang & Xie, Qin & Huo, Dongying & Wu, Kang & Nie, Yangbo & Ding, Yanyan & Kai, Zhang & ZAHANG, Yu & Deng, Zhi-Yong & Guo, Rui & WEI, Zheng & Yao, Ze-En. (2020). Measurement of neutron energy spectra of 9Be(d,n)10B reaction with a thick beryllium target[J]. Chinese Physics C,2020 ,45 , 101088
4	Understanding the isoscaling relationship in the fissioning system with evaluated data	Yong-Jing Chen et al, Understanding the isoscaling relationship in the fissioning system with evaluated data[J]. Chinese Phys. C, 2021 ,45, 084101
5	Analysis of nuclear fission properties with the Langevin approach in Fourier shape parametrization	Li-Le Liu, Yong-Jing Chen, Xi-Zhen Wu, Zhu-Xia Li, Zhi-Gang Ge, and Krzysztof Pomorski. Analysis of nuclear fission properties with the Langevin approach in Fourier shape parametrization[J].Phys. Rev. C,2021, 103, 044601

典型案例

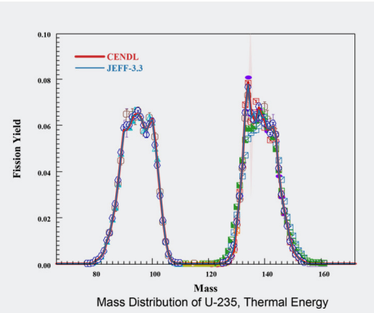
为自主研发项目“铅铋合金冷却反应堆技术”提供数据支持

中国核工业集团有限公司自主研发项目“铅铋合金冷却反应堆技术”需要为铅铋反应堆设计推荐最佳的铅、铋核反应数据。基于国际上发布的最新评价核数据库，中国核数据中心推荐了CENDL-3.2铅同位素、JEFF-3.3库Bi-209核反应数据用于铅铋快装置计算，为铅铋合金冷却反应堆技术的下一步研究工作提供了数据支持，其中核物理主题数据库作为重要的评价数据来源被该项目应用。



为核能开发科研项目“燃耗分析基础数据库开发技术研究”提供数据支持

国家国防科技工业局核能开发科研项目“燃耗分析基础数据库开发技术研究”评价、推荐燃耗信用制感兴趣的锕系和裂变产物核素的裂变产额、衰变数据和中子活化数据，用于改进燃耗数据库研制所需的基础评价数据。核物理主题数据库提供了项目研究中燃耗分析需要的精确裂变产额数据和衰变数据，为燃耗数据库的研制奠定了数据和技术基础，有助于提升燃耗分析技术准确度。



国家基础学科公共科学数据中心
National Basic Science Data Center

宇宙学数值模拟数据库

宇宙学数值模拟数据库是中国科学院紫金山天文台星系宇宙学与暗能量研究团组承建，以中国科学院紫金山天文台的“宇宙大尺度结构与星系形成高性能计算平台”为硬件依托的天文数据库，该库是目前国内体量大而且种类全的宇宙学模拟数据库之一，科学数据总量400TB左右，并以每年约30%的速度增加。其中“弱引力透镜全天球光线追踪数值模拟数据库”与“中国空间站工程巡天望远镜CSST的观测数据仿真数据库”（仅限国内参与CSST项目的科学家获取）是全球独有的两个数值模拟数据库，已广泛为科研人员和工程人员提供科研数据服务。截止目前，该库已为国内外70余位天文专业注册用户提供数据服务，平均每年支撑论文数十篇，已申请国家发明专利10项。

1.3PB
资源量

102TB
年度下载量

2021年度新增数据资源清单表

序号	数据库/集名称	数据总量	年度增量
1	SDSS,WISE,ROSAT星系，星系团的星表库观测数据库	12GB	2GB
2	FAST观测图像	104.1TB	101TB
3	CSST光路模拟	87GB	66GB
4	中国空间站工程巡天望远镜CSST的观测数据仿真	41TB	21TB
5	孪生宇宙模拟数据集	6.3TB	6.3TB
6	银河系星流数值模拟	1.6GB	1.6GB

王蕾
leiwang@pmo.ac.cn
中国科学院紫金山天文台

服务成效

宇宙学数值模拟数据库覆盖了从数值模拟到观测，从N体粒子到流体，从宇宙大尺度结构到银河系小尺度星流，从射电波段到X射线波段，从图像处理技术开发到中国空间站工程巡天望远镜CSST的科学数据图像仿真等数据产品资源。该数值模拟数据库为全球12家科研单位，110位科学家提供科研数据服务，并已为大量高影响力的科研工作（论文，专利等）提供主要数据支持。

具有国内独有的5个权威数据子库：中国空间站工程巡天望远镜CSST图像数值模拟、Elucid半解析模型等宇宙大尺度结构数值模拟、NIHAO星系流体数值模拟、弱引力透镜光线追踪数值模拟、银河系动力学模拟，星流模拟。

2021年发表数据库相关SCI论文9篇（包括Nature、天体物理杂志、英国皇家天文学报等高影响力权威杂志），申请发明专利2项，支撑科研项目11个。

中国空间站工程巡天望远镜
CSST图像数值模拟

Elucid半解析模型等
宇宙大尺度结构数值模拟

NIHAO星系
流体数值模拟

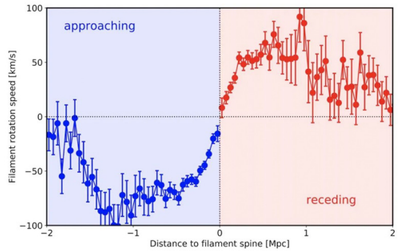
弱引力透镜
光线追踪数值模拟

银河系动力学模拟
星流模拟

典型案例

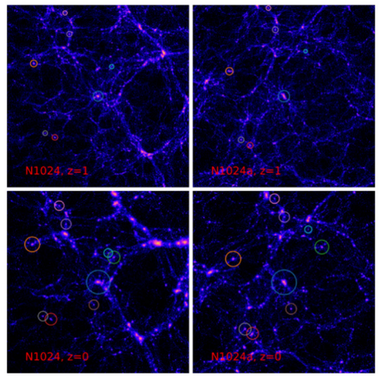
助力宇宙大尺度纤维状结构旋转信号的发现

宇宙大尺度结构中一个显著结构是纤维状结构，是一种由众多星系和暗物质组成细长结构，可以看作是星系掉落入团块结构的“高速公路”，通过使用斯隆星系巡天（SDSS）数据，首次发现了宇宙大尺度纤维状结构的旋转信号，结果表明角动量可以在前所未有的大尺度上产生。“宇宙学数值模拟数据库”为该成果提供完全无信号的随机数值模拟样本与观测样本数据，经对比发现观测样本偏离随机样本最多4.2个标准差，说明观测样本中有明显的旋转信号，从而使研究人员和审稿专家确信这一发现。



助力宇宙大尺度结构与暗物质晕关联性研究

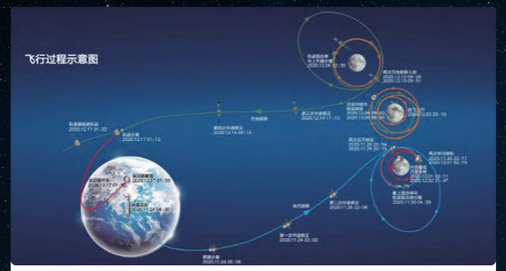
由中国科学院紫金山天文台、浙江大学-中国科学院紫金山天文台联合天文学研究中心、德国波茨坦天体物理研究所三家单位的科研团队合作完成的最新研究成果“Cosmic Web-Halo Connection Between Twin Universes”，即将发表于国际天文学顶级期刊《天体物理杂志》(Astrophysical Journal)。研究人员开发了一种所谓的“孪生模拟”（twins simulation）技术使得跨模拟的单个暗物质晕间比较成为可能，从而在研究暗物质晕演化与大尺度环境间关系的时候，达到了“控制变量”的效果。宇宙学数值模拟数据库在其中发挥重要支撑作用。



国家基础学科公共科学数据中心
National Basic Science Data Center

射电天文与深空探测数据库

射电天文与深空探测数据库是基于中国VLBI网，利用甚长基线干涉测量（Very Long Baseline Interferometry，VLBI）进行天文测地和月球与行星探测所获观测数据建立的公共数据库。VLBI数据库通过服务网站提供数据公开检索服务，并通过工程任务内网建立镜像站点，直接提供便捷的数据服务。VLBI数据库直接服务于月球及火星探测器的工程测控，探月工程历次任务VLBI测量数据是射电天文观测网服务于国家月球与深空探测重大专项过程中积累的重要数据集，具有唯一性和独有性特征，有力支撑了探月工程嫦娥五号和火星探测天问一号任务的顺利开展。



嫦娥五号月面采样返回任务各阶段示意图

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	嫦娥五号数据集	4100GB	4100GB
2	天问一号数据集	2500GB	2500GB

陈中
zhchen@shao.ac.cn
中国科学院上海天文台

服务成效

44篇
支撑论文

304个
探月工程内网在线计算任务

2个
服务国家科技重大专项

2个
支持国家重点研发项目

26个
服务国家自然科学基金项目

10个
服务省部级项目

4个
服务其他科技计划项目

服务成效列表

序号	成果名称	类型
1	中国VLBI网和e-VLBI技术在探月与深空探测工程中的应用与展望	论文
2	Fast radio burst search: cross spectrum vs. auto spectrum method	论文
3	Space Low-Frequency Radio Observatory and the Earth-Moon VLBI Experiment	论文
4	VLBI 观测双线极化条纹拟合方法	论文
5	空间VLBI数据记录格式 RDF解析与数据解码	论文
6	实时VLBI处理机技术	论文
7	Chinese VLBI network and e-VLBI technology applications in Chinese Lunar Exploration Program	论文
8	The Localization of the Single Pulse in VLBI Observation	论文
9	Fast radio burst search: cross spectrum vs. auto spectrum method	论文
10	VLBI相关处理机快视软件	软件著作权

典型案例

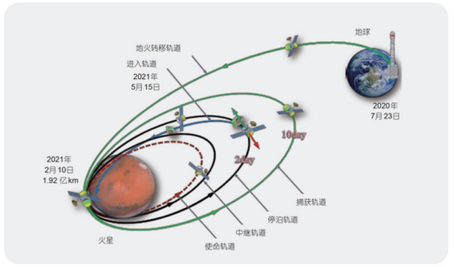
助力探月工程三期嫦娥五号任务轨道和定位快速测量

VLBI测轨分系统作为探月工程测控与回收系统的重要组成部分，与现有统一测控网和深空网一起，共同完成嫦娥五号探测器各飞行弧段的测定轨定位任务。在CE-5的11个飞行弧段中，VLBI参与了探测器9个飞行段的ΔDOR测量工作及探测器的轨道确定和预报，大大节约了项目时间，加快了技术方案验证。探月工程历史数据集提供了陆器月面着陆点及月面起飞点的精确坐标、深空网环月飞行段的三向测量标定、交会对接与样品转移段上升器的初轨确定，远程导引以及月地转移段的分离点预报等重要依据，为测控总体的工作提供有力保障。



服务中国首次火星探测器天问一号的高精度测轨

射电天文和深空测量数据库的使用大大节约了项目论证时间，加快了技术方案验证。目前天问一号探测器已在火星表面着陆，天问一号环绕器在轨运行超过600天，开展火星全球科学探测，同时为“祝融号“火星车提供数据通信与数据转发服务。射电天文和深空测量数据库实时跟踪和持续汇聚高质量的测量数据。根据近4个月的观测数据质量统计，时延测呈平均误差<0.5ns，时延率测量平均误差<0.5ps/s 。相关数据将用于后续高精度VLBI 测定轨研究，将进一步提升我国VLBI测定轨能力，达到国际前沿水平。



国家基础学科公共科学数据中心
National Basic Science Data Center

载人航天空间科学与应用数据库

载人航天空间科学与应用数据库由中国科学院空间应用工程与技术中心负责建设、运行、管理和维护的数据库，依托“载人航天工程空间科学与应用数据共享服务平台”（网址：www.msadc.cn）建设，是唯一公开的载人航天空间科学与应用数据权威发布网站。数据库在天宫一号、天宫二号任务阶段累计向各类用户提供19个有效载荷的29类数据产品，数据量达400TB，为载人航天空间科学与应用数据共享发挥着不可替代的作用。

目前，我国载人航天工程已全面进入空间站建造阶段，天和核心舱于2021年成功发射，正在轨开展无容器材料科学实验柜、高微重力科学实验柜，空间环境探测等科学实验。2022年将发射实验舱I，实验舱II，建成后将在空间生命生态、空间微重力基础物理、微重力流体物理等8大学科领域持续产生PB级的空间科学与应用数据。“载人航天空间科学与应用数据库”高效管理空间站各领域、各类型数据，为用户提供数据共享、成果管理、专题数据集等服务，助推具有国际影响力的前沿科学发现和应用研究成果产出。



402.8TB
数据总量

41.3万条
数据条目

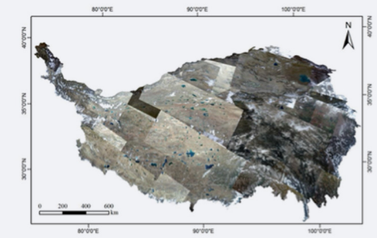
1025家
服务用户

44.6万次
累计下载

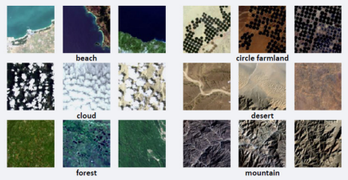
120.3万次
平台访问量

8个
学科领域

李轩
msadc@csu.ac.cn
中国科学院空间应用工程与技术中心



天宫二号宽波段成像仪青藏高原专题数据集



天宫二号自然场景数据集

服务成效

19篇
支撑论文

8项
专利

4项
标准

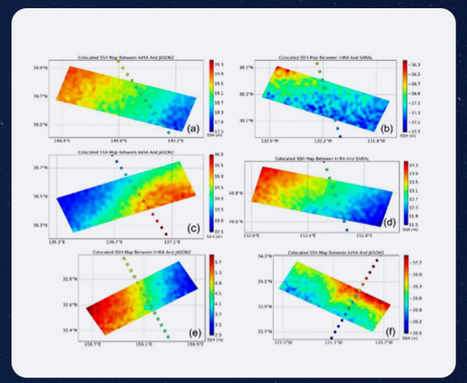
主要成果列表

序号	论文题目	期刊（会议）名称
1	Rotation adaptive correlation filter for moving object tracking in satellite videos	Neuro computing
2	Siamese networks with distractor-re-duction method for long-term visual object tracking	Pattern Recognition
3	NaSC-TG2:Natural Scene Classification With Tiangong-2 Remotely Sensed Im-agery	IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing
4	On Creating Benchmark Dataset for Aerial Image Interpretation Reviews, Guidances and Million-AID	IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing
5	A prioritization metric and modelling framework for fragmented saltmarsh patches restoration	Ecological Indicators
6	Classification of Urban Functional Areas From Remote Sensing Images and Time-Series User Behavior Data	IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing
7	A Lightweight CNN Model for Rffining Moving Vehicle Detection From Satellite Videos	IEEE Access
8	Air quality predictions with a semi-su-pervised bidirectional LSTM neural net-work	Atmospheric Pollution Re-search
9	Improved Interactive Color Visualization Approach for Hyperspectral Images	Information Visualization

典型案例

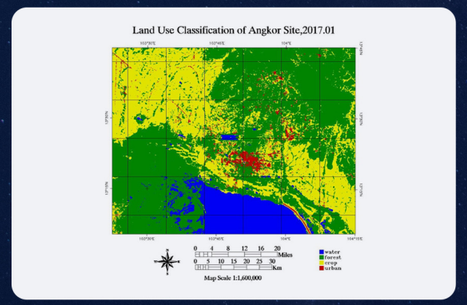
为二维海面高度数据初步评估和修正提供数据支撑

载人航天空间科学与应用数据库针对天宫二号三维成像微波高度计获取的二维海面高度数据，与来自inira和传统高度计jason-2，saral/altika和jason-3的一维海面高度数据进行比较，通过分析系统偏差和参数偏差，从而提高二维海面高度数据的准确性，提升反演精度。研究结果为进一步开展海洋工程环境和海洋测绘技术的应用基础研究，开展海洋测绘、工程勘察设计等提供重要数据支持。



支持柬埔寨吴哥窟古迹环境动态监测

在世界文化遗产柬埔寨吴哥窟古迹环境动态监测工作中，利用载人航天空间科学与应用数据库提供的天宫二号宽波段成像光谱仪和三维成像微波高度计数据产品，结合长时间序列其他遥感数据，监测柬埔寨吴哥窟世界文化遗产地本体及周边地区的环境要素动态变化（包括森林、水体、建筑等土地利用类型），并研究其变化的驱动因素，该项工作挖掘了天宫二号宽波段成像仪数据和三维成像微波高度计数据在文化遗产环境监测中的潜力价值，推动世界文化遗产的科学管理与可持续发展。



国家基础学科公共科学数据中心
National Basic Science Data Center

材料科学基础数据库

崔丽娜
clear@imr.ac.cn
中国科学院金属研究所

参建单位
中国科学院化学研究所
国家纳米科学中心

材料科学基础数据库是由中国科学院金属研究所承建，中国科学院化学所和国家纳米科学中心参加建设的科学数据库。该库涵盖金属材料、有机高分子材料及纳米材料数据，数据来源主要包括实验数据、手册数据以及文献数据。数据采用了普通分类方法，为基础研究人员、研发工程师和企业界之间建立联系。目前总数据量超过20万条，服务网站的累计数据下载量已达154.88GB，访问量达到369万次。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	金属材料节点	11.5万条	5000条
2	纳米材料数据资源节点	1.18GB	0.17 GB
3	高分子材料形态结构	2G	2G

服务成效

服务国家科技重大专项、重大工程清单

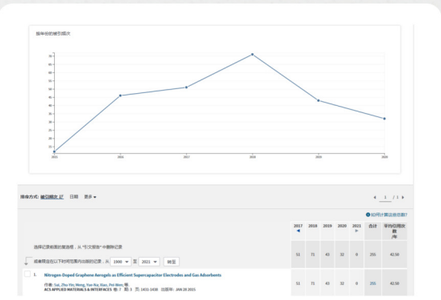
领域名称	支持项目名称	支持项目来源
国家重点研发计划	国家重点研发计划“材料基因工程关键技术与支撑平台”中“材料基因工程专用数据库架构、标准规范与数据库”（2016YFB0700500）项目	科技部
国家重点研发计划	材料环境失效过程的高通量计算模拟与高通量实验技术（2017YFB0702100）项目	科技部
国家重点研发计划	为科技部“重点基础材料技术提升与产业化”重点专项“高品质阻燃纤维及制品关键技术”（2017YFB0309300）项目的研制，提供树脂材料性能评估标准及前沿研究进展、发展趋势。为该项目顺利推进提供了支持。	科技部

主要成果列表

序号	成果名称	类型
1	A room-temperature four-terminal spin field effect transistor(自旋场效应晶体管),Nano Today, 38 (2021) 101138	重要期刊论文
2	X.F. Sun, H.T. Wang, E.-H. Han, Effect of Cr doping on the surface characteristics of Ni metal studied with first-principles calculation, Acta Metallurgica Sinica (English letters), 32 (2019) 461–470.	重要期刊论文
3	Amplitude response of conical multiwalled carbon nanotube probes for atomic force microscopy, RSC Advances, 9(2019) 429–434	重要期刊论文
4	Nitrogen-Doped Graphene Aerogels as Efficient Supercapacitor Electrodes and Gas Adsorbents, ACS APPLIED MATERIALS & INTERFACES, Vol.7(3) P1431–1438	重要期刊论文

典型案例

为国家重大项目"利用纳米材料与技术提高农药有效性与安全性的基础研究"提供数据服务支撑



材料科学基础数据库为"纳米载药系统的毒理学效应与安全性评价"(2014CB932204) 课题组提供纳米材料数据资源节点石墨烯相关数值数据检索整理，辅助课题组对石墨烯材料的掺氮合成及其物理化学性质进行研究，传统化学农药残留超标、污染环境等弊端有望通过纳米科技得以改观。根据Web of Science数据库（SCI）的检索数据，相关研究论文“Nitrogen-Doped Graphene Aerogels as Efficient Supercapacitor Electrodes and Gas Adsorbents”［ACS APPLIED MATERIALS & INTERFACES Vol.7(3) P1431–1438］，总计被引用255次, 研究结果受到广泛关注。

聚丙烯性能数据集助力企业研发抗菌材料口罩



新冠疫情期间医用口罩需求猛涨，减少因佩戴口罩滋生细菌的同时快速研制具有抗菌纳米材料口罩是企业占领市场份额的关键问题。研发抗菌口罩时，口罩制造商需要获得纤维口罩基材及改性后性能数据，然后将企业自主研发的抗菌纳米材料改性添加到聚丙烯或聚酯纤维材料中，赋予口罩纤维基材抗菌能力。有机高分子材料数据库为企业提供了纤维基体材料聚丙烯、聚酯的性能数据及改性助剂，增加纤维材料和抗菌纳米材料吸附作用。有效协助企业提升经济效益，助力中小企业发展。

国家基础学科公共科学数据中心
National Basic Science Data Center

干细胞与代谢疾病数据库

干细胞与代谢疾病数据库是以汇集干细胞基础研究数据、代谢疾病临床数据、细胞谱系资源和互联网公共数据为主要目标，基于高性能计算技术建立的集数据汇集、存储管理、挖掘分析、可视化呈现为一体的领域资源创新应用平台。

该库以中国科学院广州生物医药与健康研究院、广州再生医学与健康广东省实验室(生物岛实验室)、人类细胞谱系大科学设施为数据来源，同时汇聚高校、科研单位、医药企业多源异构数据，包括细胞生物学、免疫学、发育生物学、干细胞基础研究数据、代谢疾病临床数据和细胞谱系数据等。相关数据持续服务我国各级实验室及研究中心、科技计划项目和应用平台，为人类健康与“健康中国”战略实施提供强有力支撑。

赵娟娟
zhao_juanjuan@gibh.ac.cn
中国科学院广州生物医药与健康研究院

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	细胞多组学数据库	7000GB	6117.76GB

服务成效

2.78TB
2021年度数据增量

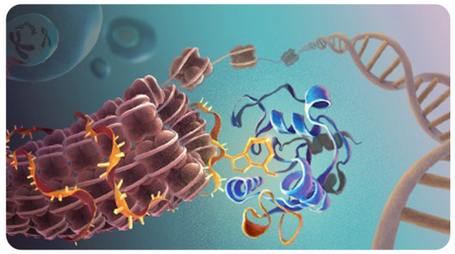
20项
撑科研项目数据汇交

23篇
论文关联数据汇交

典型案例

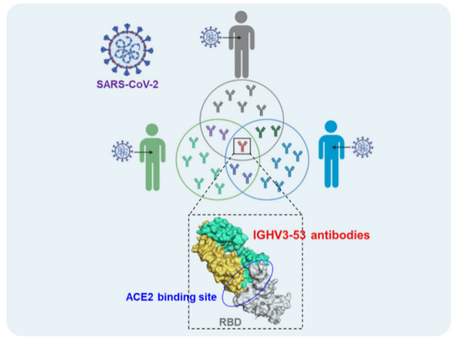
为RNA调控染色质新机制发现提供数据支撑

RNA 上的 N6- 腺苷酸甲基化（m^{6}A）是 mRNA 上最丰富的 RNA 修饰，RNA m^{6}A 可通过被含 YTH 结构域的阅读器蛋白（reader）识别进而参与各种生物学功能。干细胞与代谢疾病数据库为课题组提供了生物信息分析软件工具，用于高通量测序数据的一站式、流程化计算分析，助力课题组发现 RNA m^{6}A 修饰调控异染色质形成的新机制，阐明 RNA m^{6}A 阅读器 YTHDC1 在这一机制中的关键作用，相关研究以“The RNA m6A reader YTHDC1 silences retrotransposons and guards ES cell identity”为题发表在 Nature 上。



支持新冠患者抗体应答特征研究和抗体药物快速研发

新冠病毒严重威胁公共卫生安全，开发安全有效的疫苗是控制疫情的最有效途径。课题组对13名新冠患者发病后多个时间点的样本进行了高通量抗体组库测序，共计获得近2亿条抗体序列。干细胞与代谢疾病数据库为课题组提供了生物信息分析软件工具用于抗体序列全生命周期管理，为揭示SARS-CoV-2诱导的特异性抗体反应在感染人群中的共同特征，应对新发突发传染病抗体药物快速研发提供数据支持。



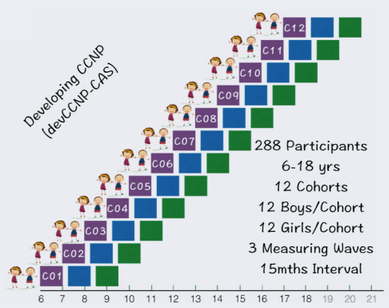
中国活体人脑成像共享数据库

National Basic Science Data Center 国家基础学科公共科学数据中心

中国活体人脑成像共享数据库（Chinese Data-sharing Warehouse for In-vivo Imaging Brain）是中国科学院心理研究所磁共振成像研究中心主导建立的包含广泛个体差异的人群的脑影像大数据库。该数据库依托于中国科学院心理研究所，包含中国儿童彩巢计划、中国重度抑郁症数据库、国际信度与可重复性（CoRR）大数据共享等数据资源，以及中国科学院心理研究所独立自主研发的人脑数据处理软件 CCS 和 DPABI，并经过该软件标准化流程处理的超过一万例的多模态人脑活体成像数据。国际人脑数据库自开放共享后，截至 2021 年底已有国内外超过万余名科研工作者下载了预处理过后的活体脑影像数据，成为国内外研究者最为常用的共享数据下载平台之一。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	2021年彩巢计划	800GB	250GB



2021年彩巢计划数据集

左西年
xinian.zuo@bnu.edu.cn
北京师范大学

参建单位

中国科学院心理研究所　北京大学
中国科学院生物物理所　西南大学

服务成效

3篇
数据论文

Dong, H. M., Margulies, D. S., Zuo, X. N., & Holmes, A. J. (2021). Shifting gradients of macroscale cortical organization mark the transition from childhood to adolescence. Proceedings of the National Academy of Sciences of the United States of America, 118(28), e2024448118.

Liu S, Wang YS, Zhang Q, Zhou Q, Cao LZ, Jiang C, Zhang Z, Yang N, Dong Q, Zuo XN; Chinese Color Nest Consortium (2021). Chinese Color Nest Project: An accelerated longitudinal brain-mind cohort. Developmental Cognitive Neuroscience, 52:101020.

Zhou Q, Liu S, Jiang C, He Y, Zuo XN; Chinese Color Nest Consortium (2021).Charting the human amygdala development across childhood and adolescence: Manual and automatic segmentation. Developmental Cognitive Neuroscience, 52:101028.

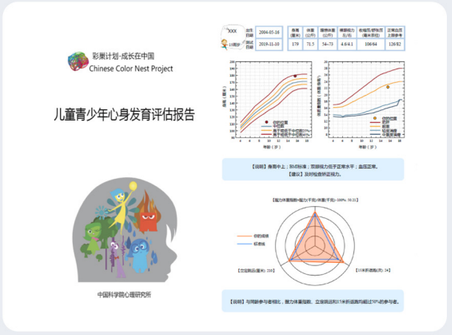
3项
支撑项目

国家社会科学基金重大项目：汉语自闭症儿童多模态交际障碍的多学科协同研究
北京师范大学领军人才项目：发展群体神经科学
中国科学技术协会 ：“发展群体神经科学”高层次专家战略研讨会

典型案例

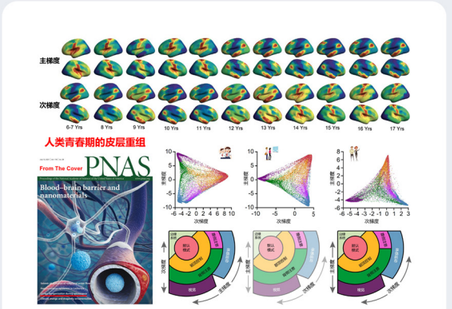
为中国儿童青少年身心发育评估报告提供数据支持

“学龄儿童青少年脑与心智发育研究”项目完成了521人次6-18岁学龄期儿童青少年脑影像及行为数据的采集，构成了我国首个学龄儿童脑与心智发育队列“彩巢计划-成长在中国”。中国活体人脑成像共享数据库通过整合儿童从身高体重到脑发育的诸多方面内容，绘制发育轨线，针对单个儿童提供个体化的身心发育评估报告，为家长提供该儿童各领域的生长发育情况提供科学的参照数据，并提出相应的科学评估建议，及时解决家长相关咨询，得到受众家长良好评价。



为青春期起源提供数据循证，支撑高影响力成果发表

中国活体人脑成像共享数据库基于彩巢计划（devCCNP）的追踪数据以及已经公开的其他中国儿童青少年形态学影像数据进行人脑功能连接梯度分析，发现人脑功能连接从小学阶段的基础功能网络为主导，在青春期（12-14岁）及以后快速转化为以高级认知为主的联合功能网络。该研究成果已在《Proceedings of the National Academy of Sciences of the United States of America》（IF:9.58）上发表，不仅为理解青春期的起源提供了新的思考视角，而且也为推进精准医学临床实践和个性化教育改革提供了循证科学的基础。



国家基础学科公共科学数据中心
National Basic Science Data Center

地理空间数据云

地理空间数据云是先进的“一站式”地学数据综合服务平台，通过自主研发可靠的大规模地理空间数据实时分析引擎，实现了亿级元数据高效管理和PB级实体数据分析服务。汇聚国际及国内开放地理空间数据资源，生产优质数据产品，创新模型在线、数据众包、在线FTP及在线解译等服务模式，提供数据和各类服务支持。数据量达160TB，服务用户50万，极大地推动了我国数据开放共享，节约大量科研成本，产生显著社会效益。

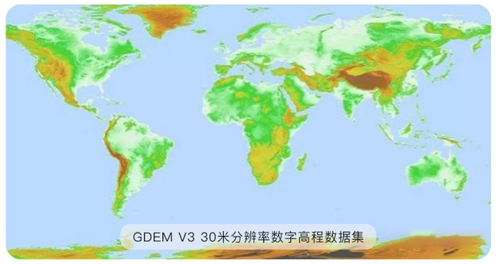
林青慧
lqh@cnic.cn
中国科学院计算机网络信息中心

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	高分一号WFV数据产品	2560GB	2560GB
2	LANDSAT系列数据	54272GB	3061.76GB
3	MODIS陆地标准产品	38912GB	1269.76GB
4	MODIS-L1B标准产品	48128GB	727.04GB
5	DEM数字高程数据	1228.8GB	286.72GB
6	Sentinel数据	18432GB	5324.8GB
7	“数据汇”分享数据产品	5GB	1GB

2021年度主要成果列表

序号	成果名称	类型
1	支撑CNKI学术期刊论文1183篇	论文
2	支撑CNKI学位论文340篇	论文
3	支撑SpringerLink 英文论文141篇	论文
4	支撑sciencedirect 英文论文325篇	论文


GDEM V3 30米分辨率数字高程数据集


高分一号WFV数据集

服务成效

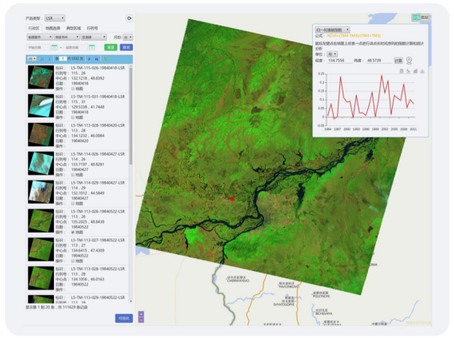
2021年地理空间数据云累计数据量达160TB，数据条目数达到800万条，服务用户50万人，总访问人次1000万次，数据在线服务量1PB，在线任务数10万个，支持科研项目1000余次，支持论文1500余篇。极大地推动了我国数据开放共享，节约大量科研成本，产生显著社会效益。

160TB	50万人	1PB	10万个	1000次+	1500篇+
累计数据量	服务用户	数据在线服务量	在线任务数	支持科研项目	支持论文

典型案例

为A类战略性先导科技专项“地球科学大数据”提供全面数据支撑

围绕科研人员在地理空间数据搜集、获取、存储、在线分析和可视化的迫切需求，地理空间数据云基于已汇聚的数十颗国际遥感卫星数据和国内陆地观测卫星数据，建立完善的数据储备库。地理空间数据云凭借长期的技术积累和丰富的数据资源，通过为专项提供Landsat、MODIS、DEM等重要数据资源支撑，为中国科学院A类战略性先导科技专项“地球大数据科学工程”基础地理空间数据资源及数据管理提供数据和技术保障。



支撑中国自然灾害数据库建设

基于地理空间数据云遥感数据积累及地理空间数据处理技术，综合分析了中国自史前起至清代共20个历史断代的自明代自然灾害宏观概况，并针对十余种典型自然灾害，如水灾、旱灾、蝗灾、地震灾害、地质灾害、疫灾等进行典型案例分析、府级空间分布及时间维度统计分析、展示。应用现代科技的基础和方法，探索历史时期自然灾害的特点和发生规律，对当今社会建设、防灾抗灾等具有重要意义。相关成果在中国社会科学院中国历史研究院七楼大屏系统开放展示。



国家基础学科公共科学数据中心
National Basic Science Data Center

青海湖区域综合研究专题数据库

青海湖区域综合研究专题数据库是集青海湖区域植被、兽类、水鸟、生态水文、遥感微生物等在内的基础数据资源收集、整合、关联性分析、集成、共享，融合数据、模型为一体的数据处理和应用环境，是基本完整的数据资源体系和一站式数据服务系统。该库共建设五大类44个子库，资源总量达20TB。服务成效方面，目前年平均下载量100GB，年平均访问量20万人次，为生物多样性保护、疫病防治、水环境监测、水华防治和保护区管理等提供了准确、及时的特色服务。

5类44个
子库

20TB
资源总量

100GB
年均下载量

20万
年平均访问人次

高瑜蔚
gaoyuwei@cnic.cn
中国科学院计算机网络信息中心

参建单位
中国科学院动物所
中国科学院武汉病毒研究所
中国科学院水生生物研究所
青海湖国家级自然保护区管理局

服务成效

SCI论文	20余篇	专利	3项
其他论文	70余篇	国家重大项目	10余项
软件著作权	30余项	科技计划项目	5项

特色软件

野外巡护移动终端、野生动物识别等软件已在车八岭、卧龙、呼伦湖、长白山、无量山等多个保护区进行了安装和试运行。

特色数据集

MODIS产品集、刚毛藻水华空间分布数据集等有效支持地方生态保护。

典型案例

为上海科技馆公众教育工作提供数据服务

青海湖野外鸟类监控视频自2015年7月正式接入上海科技馆至今已连续6年对公众展出，主要通过“文明的史诗”主题板块的常设主题展区“未来之路”进行展示。该展区以中国地图为背景，在不同地理位置放置了10余块用来展示保护区实时监控视频的显示屏，展示包括青海湖鸟类保护在内等工作情况，使得社会大众可以在科技馆中实时远程看到青海湖的鸟类真实生活状态，具有很好的生态保护宣传教育意义。



支持毗邻月牙湖变红事件调查与成因分析

2017年4月12日青海湖毗邻月牙湖突然变成粉红色，当地群众拍照并在网络上传播后，引起了社会各界广泛关注。青海湖数据库为本次调查事件报告的编写和成因分析、以及后续持续保护政策制定提供了重要数据支持，主要包括遥感数据、水体调查数据（水体调查植物群落数据、底栖动物数据、浮游动物数据、浮游植物数据、理化数据）、气象数据等，为该事件不同层面（管理局、地方政府、中科院）报告撰写提供了重要的依据，为青海湖流域2018–2021年持续生态保护工作决策提供了重要的支持。



国家基础学科公共科学数据中心
National Basic Science Data Center

干旱区药用资源数据库

干旱区药用资源数据库是涵盖新疆、中亚等干旱、半干旱地区的药用资源综合数据库，包含药用资源信息（2148条）、天然产物/标准提取物信息（1670条）、药材标本（292份）、基原植物标本（2063份）、标准品等实物信息。该库实现干旱区药用数据资源规范化线上发布，提供药材基原、有效成分化合物和标准提取物等数据查询和下载服务（网址：www.cam.nsdata.cn）。2021年度支撑发表论文30余篇、授权专利4个、支撑项目4个，项目经费约1000万，其中以中亚药用资源为基础，结合科研和生产，获批科技部“一带一路”联合实验室；中药1.1类创新药“艾拉片”获批新药证书。

2148条
药用资源信息

1670条
天然产物、标准提取物

292份
药材标本

2063份
基原植物标本

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	药用资源信息平台及标本库	2GB	1GB
2	药材标准信息平台及标本库	2GB	0.5GB
3	标准提取物数据库	2GB	0.5GB
4	天然产物资源库	2GB	0.5GB

刘戈宇
liugy@ms.xjb.ac.cn
中国科学院新疆理化技术研究所

参建单位
中国科学院新疆生态与地理研究所

服务成效

- 2021年度数据量1TB、下载量100余次，支撑论文30余篇；授权专利4个，支撑项目4个。
- 支持研发的中药1.1类创新药“艾拉片”，于2021年10月13日获得国家药品监督管理局颁发的药物临床试验批准通知书，获准开展临床研究。
- 以中亚药用资源为基础结合科研和生产，申报并获批科技部“一带一路”联合实验室。
- 为企业鉴定药材10余个，线上查询和线下咨询、下载。为医院制剂备案资料和线下咨询20余次。

典型案例

助力乌兹别克斯坦等中亚国家抗击新冠疫情



干旱区药用资源数据库先后在中亚建成了中乌天然药物实验室（与乌兹别克斯坦科学院植物化学研究所共建）。新冠疫情爆发以来，实验室联合中乌生物大分子实验室立即组建应急研发团队，合作发现天然抗病毒药物Rutan治疗新冠肺炎的新用途，在细胞水平证明其具有明显的抑制病毒复制作用。在乌兹别克斯坦103例中度新冠肺炎患者的临床研究表明，Rutan安全性良好，可显著降低炎症反应，缩短症状持续时间，已在乌兹别克斯坦生产并用于新冠疫情抗击一线，为保护人民生命健康做出贡献。

支撑中国科学院战略生物资源平台建设



依托干旱区药用资源数据库已收录的中亚地区药用数据资源，面向国家重大需求和国民经济主战场，获批中科院战略生物资源平台项目“中亚地区特色药用植物种质保藏培育与资源库建设”。作为中国科学院战略生物资源平台的重要组成部分，该项目将对具有重要科研、经济价值的药用资源，进行评价、挖掘；对干旱区野生和珍稀濒危植物进行生长习性、繁殖特征和生态适应性等科学问题研究；并对有重要经济价值的特有资源进行功能评价，制定种质保育、生产标准，培育新品种并加强示范推广。

国家基础学科公共科学数据中心
National Basic Science Data Center

黑龙江省特色基础科技数据库

黑龙江省特色基础科技数据库是汇集黑龙江省寒地特色物种、论文、专利及大型科学仪器的数据库。近五年来，通过数据库平台进行资源共享和服务，累计支持中小微企业及单位等使用上万次。2021年政策支持企业460家，涉及服务项目623项，从业人员总数9767人，企业及企业类服务机构新增就业人数151人，支持金额4092.30万元。2021年黑龙江省特色基础科技数据库新增数据资源：1979年–2000年特色黑土基要情况物理性质数据，为科研人员研究提供了黑土地数据支撑；智能心电数据2万条等数据资源，为相关企业开展评测服务。

2021年度新增数据资源清单

序号	数据库/集名称	数据总量	年度增量
1	智能心电数据	0.5GB	0.5GB

荀妍妍
191540343@qq.com
黑龙省科技资源中心

参建单位
东北农业大学
黑龙江省科信科技信息中心

服务成效

468828时
仪器库对外服务计时数

29,685万元
企业服务技术合同登记

4964台/套
仪器库对外服务仪器

20,448万元
仪器服务收入

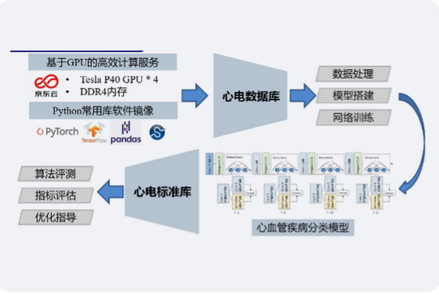
0.5GB/160次
智能心电数据数据量/下载量

2548次
心血管疾病相关的医学知识文献库在线访问量

典型案例

智能心电数据库支撑中小企业心电设备医疗领域应用

黑龙江地区的中小型企业的心电算法研究长期受限于软硬件评测经验以及数据短缺等问题，导致人工智能心血管研究难以开展。黑龙江省特色基础科技数据库通过规范心电标注术语、开发标注平台构建心电标准数据库，把心电数据进行人工智能分类，支持中小企业进行人工智能算法研究，该项研究补充了若干心电技术和临床输出要求，形成了一套硬件评测标准，促进人工智能技术在中小企业心电设备医疗领域的落地应用，为更多心电算法研究、数据测试和验证工作提供了数据支撑。



为哈尔滨第五医院提供语义关联的医药信息在线咨询服务

互联网医疗服务背景下，社会各界对在线医疗信息咨询服务需求量越来越大。黑龙江省特色基础科技数据库通过将不同来源数据统一化采集、清洗和智能标注整编成药品研发知识文献库，为药企、医院、科研院所等提供全面、权威、及时的医药资源，充分满足医务工作者对医药知识获取的需求。目前在哈尔滨第五医院进行试点提供医药信息在线咨询、查新服务。为医务工作者临床决策支持，如用药分析、药品不良反应、疾病并发症、治疗效果相关性分析、制定个性化治疗方案起到了辅助决策的作用，帮助规范化治疗。



附录

年度大事记

3月

院级科学数据中心考核获“优秀”

顺利通过中国科学院十三五信息化专项“中国科学院科学数据中心体系建设”项目验收，考核获评“优秀”。

6月

年度考核工作会顺利召开

年度考核工作会在天津顺利召开，来自40余家单位近50位数据专家和科研工作者参会。与会者顺利完成工作考核，并研讨数据整合模式、关键技术、数据应用和未来发展等。

6月

新增5个优秀科学数据库

6月18日，2021年度新建科学数据库课题专家评审会在北京顺利召开，世界山茶属植物品种数据库、机器学习基准图像集等5个优秀科学数据库成功入选。

3月

协办“第三届全国高校数据驱动创新研究大赛”

作为唯一国家科学数据中心协办由北京大学主办的“第三届全国高校数据驱动创新研究大赛”，鼓励全国高校学生利用数据科学技术和方法，开展基于高质量开放数据的研究，促进数据价值的挖掘和创造。

7月

国家科技基础条件平台中心等到访调研

国家科技基础条件平台中心会同国家科技资源共享服务工程技术研究中心一行到访，调研科技计划项目数据汇交工作，围绕科技项目汇交数据元数据方案、协同机制和技术方案等开展交流。

9月

科技部基础研究司等到访调研

科技部基础研究司领导一行到访调研科学数据工作情况，国家基础数据中心进行了专题工作汇报，详细介绍了成立两年以来在创新数据发展模式、服务国家战略方面的工作成效。

9月

王仲朗研究员当选国际山茶协会副主席

9月27日，“世界山茶属植物品种数据库”课题负责人王仲朗研究员当选国际山茶协会（International Camellia Society, ICS）副主席，是第三位获得这一职位的中国人。

9月

联合中国人工智能学会设立数据论文专刊

联合中国人工智能学会和《中国科学数据》组织设立多语种智能信息处理数据专刊，计划2022年正式出版，力争打造具有国际影响力的人工智能多语种数据集。

10月

入选国家“十三五”科技创新成就展

10月21–27日，成功亮相科技部会同有关部门举行的国家“十三五”科技创新成就展，展示了十三五期间在基础学科数据资源、服务、技术和应用方面取得的成就。

10月

成立交通运输分中心

10月11–12日，参与承办的第七届(2021)中国科学数据大会在内蒙古呼和浩特市顺利召开，来自全国的专家学者400余人参加。会上，国家基础数据中心交通运输分中心正式签约仪式成功举行。

10月

共同发起“开放科学国际创新联盟”

作为联合发起方，参与北京市科学技术研究院牵头成立的“开放科学国际创新联盟”以及联盟发起的“开放科学实践北京倡议”，收到中关村论坛执行委员会办公室发来的感谢信。

11月

服务800多个国家级项目数据汇交

服务支撑五个项目管理机构的国家重点研发计划项目科学数据汇交工作，总数突破800个，仅2021年度增量700余个。同时对外开展汇交培训服务工作60余次，覆盖项目人员6000人次以上。

11月

推出《中国活体人脑成像数据集专刊》

联合北京师范大学、《中国科学数据》推出《中国活体人脑成像数据集专刊》，以提升我国在神经影像领域的国际影响力，加速推动神经信息化科学研究的进步。

12月

入选中国科学院十三五信息化成果展

12月23日，中国科学院网络安全和信息化领导小组办公室经会议发布了首批通过认定的院科学数据中心名单，国家基础数据中心顺利通过学科中心认定，并成功入选中国科学院十三五信息化成果展。

序号	主题库名称	数据库/集名称	数据总量
1	交通运输分中心	抗滑性能数据集	0.5GB
2	交通运输分中心	结构强度数据集	15.1GB
3	交通运输分中心	动态响应数据集	600TB
4	交通运输分中心	路表状况数据集	7.7GB
5	交通运输分中心	环境数据集	1.47GB
6	交通运输分中心	桥梁服役性能数据	6.8GB
7	交通运输分中心	京津冀城市群多模式客运枢纽一体化运行数据集	12TB
8	交通运输分中心	高寒高海拔公路技术状况与气象环境数据集	1GB
9	交通运输分中心	全国道路货运市场运行监测数据集	218GB
10	交通运输分中心	道路冷链运输数据集	48GB
11	交通运输分中心	环境与设施融合的驾驶模拟试验数据集	20GB
12	交通运输分中心	危险情景模拟再现系统场景测试试验数据集	4.67GB
13	交通运输分中心	桥岛隧跨海集群环境驾驶模拟数据集	3.2GB
14	交通运输分中心	复合立体高速驾驶行为数据集	1.5GB
15	交通运输分中心	公路交通安全风险评估数据集	50TB
16	交通运输分中心	水运工程科学研究与应用数据库	190GB
17	世界山茶属植物品种数据库	世界山茶属植物品种名称数据集	75GB
18	机器学习基准图像集	农业病虫害研究图库	580GB
19	环境健康主题数据库	生物-污染物负荷数据集	7.07MB
20	中国海岸带植物资源数据库	海岸带陆生高等植物分类信息与图片数据集	5GB
21	中国海岸带植物资源数据库	海岸带底栖大型藻类图片与代表物种基因组数据集	580GB
22	中国海岸带植物资源数据库	海岸带底栖微藻分类信息与图片数据集	15GB
23	医学物理数据库	重离子肿瘤治疗数据集	300GB
24	医学物理数据库	同位素药物数据集	200GB
25	医学物理数据库	医疗辐射防护数据集	200GB
26	中国动物主题数据库	中国动物志数据库	7.84GB
27	中国动物主题数据库	中国动物图谱数据库	0.632GB
28	中国动物主题数据库	中国经济动物数据库	1.315GB
29	中国动物主题数据库	中国鸟类数据库	60GB
30	中国动物主题数据库	中国哺乳动物数据库	0.86GB
31	中国动物主题数据库	中国蝴蝶数据库	2.76GB
32	中国动物主题数据库	中国蜜蜂数据库	0.1GB
33	中国动物主题数据库	中国内陆水体鱼类数据库	1GB
34	中国动物主题数据库	中国两栖动物	0.47GB
35	中国动物主题数据库	中国爬行动物数据库	0.34GB
36	中国动物主题数据库	中国直翅目与革翅目昆虫数据库	0.38GB
37	中国动物主题数据库	中国蜚蠊目数据库	1MB
38	中国动物主题数据库	中国双尾纲与原尾纲数据库	53MB
39	中国动物主题数据库	中国甲虫名录数据库	2.26MB
40	中国动物主题数据库	中国动物物种编目数据库	111.54MB

序号	主题库名称	数据库/集名称	数据总量
41	中国动物主题数据库	中国灵长类物种文献及使用数据库	3.24TB
42	中国动物主题数据库	中国有毒动物数据库	200MB
43	中国动物主题数据库	西南县级脊椎动物分布名录	25MB
44	中国动物主题数据库	西南保护区脊椎动物分布名录	15MB
45	中国动物主题数据库	云南鸟类数据库	1.24TB
46	中国动物主题数据库	云南蝴蝶分布名录	15MB
47	中国动物主题数据库	云南森林昆虫分布名录	20MB
48	中国动物主题数据库	中国脊椎动物分类代码数据库	10MB
49	中国动物主题数据库	中国昆虫模式标本数据库	1.8GB
50	中国动物主题数据库	中国昆虫新种数据库	8MB
51	中国动物主题数据库	地方动物志数据库	1GB
52	中国动物主题数据库	实验动物数据库	100MB
53	中国动物主题数据库	动物图像标注数据库	1.58TB
54	中国动物主题数据库	中国脊椎动物红色名录	30MB
55	中国动物主题数据库	昆虫模式标本数据库v2	500MB
56	中国动物主题数据库	云南两栖类数据库	230MB
57	中国动物主题数据库	西南地区保护、濒危、特有种编目数据库	12MB
58	中国动物主题数据库	生物评估与保护	50MB
59	中国动物主题数据库	动物研究专家数据库	1MB
60	光学技术数据库	光学玻璃数据库	10GB
61	光学技术数据库	光学晶体数据库	5GB
62	光学技术数据库	光学塑料数据库	5GB
63	光学技术数据库	光源数据库	2GB
64	光学技术数据库	光电探测器数据库	5GB
65	光学技术数据库	光学传感器数据库	10GB
66	光学技术数据库	光学镜头数据库	10GB
67	光学技术数据库	光学棱镜数据库	1GB
68	光学技术数据库	光学标准库	1GB
69	光学技术数据库	镀膜材料及膜系数据库	1GB
70	光学技术数据库	光学成像数据库	380GB
71	光学技术数据库	光学检测数据库	80G
72	光学技术数据库	特色光谱信息数据库	120G
73	光学技术数据库	激光器件数据库	2GB
74	光学技术数据库	激光传输特性数据库	1GB
75	光学技术数据库	激光光学系统数据库	1GB
76	光学技术数据库	光纤光学数据库	2GB
77	光学技术数据库	微光与红外器件数据库	1GB
78	光学技术数据库	红外传输特性数据库	1GB
79	光学技术数据库	微光与红外光学系统数据库	1GB
80	光学技术数据库	短波光学数据库	1GB

序号	主题库名称	数据库/集名称	数据总量
81	地理空间数据云	LANDSAT系列数据	53TB
82	地理空间数据云	MODIS陆地标准产品	38TB
83	地理空间数据云	MODIS中国合成产品	4.5TB
84	地理空间数据云	MODIS-L1B标准产品	47TB
85	地理空间数据云	DEM数字高程数据	1.2TB
86	地理空间数据云	EO-1系列数据	3.24TB
87	地理空间数据云	NOAA AVHRR数据产品	1.56TB
88	地理空间数据云	大气污染插值数据	1.37TB
89	地理空间数据云	Sentinel数据	18TB
90	地理空间数据云	“数据汇”分享数据产品	5GB
91	地理空间数据云	高分一号WFV数据产品	2.56TB
92	化学主题数据库	化合物基本信息库	7,29MB
93	化学主题数据库	物性及热化学数据库	168MB
94	化学主题数据库	纯化合物相变数据库	18MB
95	化学主题数据库	化合物结构数据库	4,17MB
96	化学主题数据库	参考文献数据库	36MB
97	化学主题数据库	离子簇热化学数据库	44.5MB
98	化学主题数据库	化合物估算性质数据	50MB
99	化学主题数据库	化学核心期刊数据库	6,17MB
100	化学主题数据库	中国化学文献数据库	1,96MB
101	化学主题数据库	中国化学专利数据库	6,39MB
102	化学主题数据库	质谱谱图数据库	160.2MB
103	化学主题数据库	毒性化合物数据库	213.9MB
104	化学主题数据库	MSDS数据库	6MB
105	化学主题数据库	药物天然产物数据库	623.1MB
106	化学主题数据库	化学试剂数据库	1,61MB
107	化学主题数据库	精细化工产品数据库	16.6 MB
108	化学主题数据库	化合物活性数据库（HIV）	1,06MB
109	化学主题数据库	C-13核磁数据库	2,33MB
110	化学主题数据库	化合物登录数据库	8,175MB
111	化学主题数据库	量化计算结果数据库	4.6GB
112	化学主题数据库	稀土物性数据库	140MB
113	化学主题数据库	稀土萃取数据库	140MB
114	化学主题数据库	离子簇反应热化学数据库	0.98MB
115	化学主题数据库	植物化学成分数据库	5,835MB
116	化学主题数据库	中草药单一化学成分质谱数据库	600MB
117	化学主题数据库	化合物生物活性数据库	20MB
118	化学主题数据库	中药代谢组学数据库	250MB
119	化学主题数据库	化学反应数据库	2GB
120	化学主题数据库	红外谱图数据库	3.8GB

序号	主题库名称	数据库/集名称	数据总量
121	化学主题数据库	化工产品数据	30MB
122	化学主题数据库	专业情报数据库	4MB
123	化学主题数据库	化学物质分析方法	10MB
124	化学主题数据库	药品名称数据库	5MB
125	化学主题数据库	化学配方数据库	50MB
126	化学主题数据库	工程塑料数据库	6MB
127	化学主题数据库	化学物质毒性数据库	300MB
128	中国植物主题数据库	植物彩色照片库	1092651张
129	中国植物主题数据库	古植物标本照片库	9000张
130	中国植物主题数据库	京津冀常见野生植物数据库	7126条
131	中国植物主题数据库	中国芳香植物	1377条
132	中国植物主题数据库	食用植物资源数据库	1133条
133	中国植物主题数据库	中国有毒植物物种库	1137条
134	中国植物主题数据库	省级植物名录数据库	9个
135	中国植物主题数据库	文献志书数据库	102911本
136	中国植物DNA条形码数据库	中国植物DNA条形码数据库	65.91GB
137	核物理主题数据库	评价核数据库	834GB
138	核物理主题数据库	原子核特性数据库	3MB
139	核物理主题数据库	实验核数据库	952MB
140	核物理主题数据库	常用核衰变数据库	24MB
141	核物理主题数据库	核天体数据库	24MB
142	核物理主题数据库	钍铀循环专用核数据库	800MB
143	核物理主题数据库	医用同位素数据库	7MB
144	原子分子物理主题数据库	原子光谱数据库	187万
145	原子分子物理主题数据库	原子能级数据库	9万
146	原子分子物理主题数据库	原子电离势数据库	0.2万
147	原子分子物理主题数据库	电子原子碰撞激发数据库	36万
148	原子分子物理主题数据库	电子原子碰撞电离数据库	1.4万
149	原子分子物理主题数据库	离子分子碰撞数据库	8.2万
150	原子分子物理主题数据库	电子分子碰撞数据	1.4万
151	宇宙学数值模拟数据库	SDSS,WISE,ROSAT星系，星系团的星表库观测数据库	12GB
152	宇宙学数值模拟数据库	NIHAO流体数值模拟数据库	16TB
153	宇宙学数值模拟数据库	弱引力透镜全天球光线追踪数值模拟数据库	28TB
154	宇宙学数值模拟数据库	ELUCID宇宙学数值模拟数据库	64TB
155	宇宙学数值模拟数据库	FAST观测图像	104TB
156	宇宙学数值模拟数据库	illustris TNG	64TB
157	宇宙学数值模拟数据库	CSST光路模拟	87GB
158	宇宙学数值模拟数据库	中国空间站工程巡天望远镜CSST高精度宇宙学数值模拟	955TB
159	宇宙学数值模拟数据库	中国空间站工程巡天望远镜CSST的观测数据仿真	40TB
160	宇宙学数值模拟数据库	孪生宇宙模拟数据集	6.3TB

序号	主题库名称	数据库/集名称	数据总量
161	宇宙学数值模拟数据库	银河系星流数值模拟	1.6GB
162	射电天文与深空探测数据库	嫦娥一号数据集	1.8TB
163	射电天文与深空探测数据库	嫦娥二号数据集	4.2TB
164	射电天文与深空探测数据库	嫦娥三号数据集	6.8TB
165	射电天文与深空探测数据库	嫦娥四号数据集	9.3TB
166	射电天文与深空探测数据库	嫦娥五号再入返回试验数据集	1.8TB
167	射电天文与深空探测数据库	嫦娥五号数据集	4.1TB
168	射电天文与深空探测数据库	天问一号数据集	2.5TB
169	黑龙江省特色基础科技数据库	论文数据库	1GB
170	黑龙江省特色基础科技数据库	专利库	0.5GB
171	黑龙江省特色基础科技数据库	仪器设置库	2GB
172	黑龙江省特色基础科技数据库	智能心电数据	0.5GB
173	中国活体人脑成像共享数据库	2021年彩巢计划	800GB
174	干细胞与代谢疾病数据库	代谢疾病数据资源库	1MB
175	干细胞与代谢疾病数据库	干细胞组学资源库	500GB
176	干细胞与代谢疾病数据库	临床医学知识系统	1TB
177	干细胞与代谢疾病数据库	健康科学数据库	20TB
178	干细胞与代谢疾病数据库	人类致癌基因数据库	400GB
179	干细胞与代谢疾病数据库	人类致癌转移数据库	10GB
180	干细胞与代谢疾病数据库	自身抗体生物标志物数据库	30MB
181	干细胞与代谢疾病数据库	单细胞肿瘤数据资源库	4.8GB
182	干细胞与代谢疾病数据库	小鼠细胞图谱	60TB
183	干细胞与代谢疾病数据库	人类细胞图谱	50TB
184	干细胞与代谢疾病数据库	细胞多组学数据库	7TB
185	干细胞与代谢疾病数据库	国际开放数据库	50GB
186	载人航天空间科学与应用数据库	“一带一路”沿线港口城市天宫二号多光谱影像数据集	28.3GB
187	载人航天空间科学与应用数据库	天宫二号宽波段成像仪青藏高原专题数据集	13.2GB
188	载人航天空间科学与应用数据库	天宫一号高光谱成像仪场景分类数据集	0.8GB
189	载人航天空间科学与应用数据库	天宫二号遥感图像自然场景分类数据集	8.8GB
190	材料科学基础数据库	金属材料节点	11.5万条
191	材料科学基础数据库	有机高分子材料节点	9.25万条
192	材料科学基础数据库	纳米材料数据资源节点	1.18GB
193	材料科学基础数据库	高分子材料形态结构	2GB
194	干旱区药用资源数据库	药用资源信息平台及标本库	2GB
195	干旱区药用资源数据库	药材标准信息平台及标本库	2GB
196	干旱区药用资源数据库	标准提取物数据库	2GB
197	干旱区药用资源数据库	天然产物资源库	2GB

2021年度代表性论文

（科学数据管理与技术）

序号	题目	单位
1	gcType: a high-quality type strain genome database for microbial phylogenetic and functional research	中国科学院计算机网络信息中心
2	VarEPS: an evaluation and prewarning system of known and virtual variations of SARS-CoV-2 genomes	中国科学院计算机网络信息中心
3	A Novel Image Dehazing Framework for Robust Vision-Based Intelligent Systems	中国科学院计算机网络信息中心
4	Machine Learning Techniques in China's National Foodborne Disease Surveillance System	中国科学院计算机网络信息中心
5	Foodborne Disease Risk Prediction using Multi-Graph Structural LSTM: Algorithm Design and Validation Study	中国科学院计算机网络信息中心
6	RSIMS：Large-Scale Heterogeneous Remote Sensing Image Management System	中国科学院计算机网络信息中心
7	Learning Adaptive Embedding Considering Incremental Class	中国科学院计算机网络信息中心
8	High-Efficiency Machine Learning Method for Identifying Foodborne Disease Outbreaks and Confounding Factors	中国科学院计算机网络信息中心
9	Prediction of Foodborne Diseases Pathogens: A Machine Learning Approach	中国科学院计算机网络信息中心
10	S2OSC: A Holistic Semi-Supervised Approach for Open Set Classification	中国科学院计算机网络信息中心
11	A deep hybrid neural network for single image dehazing via wavelet transform	中国科学院计算机网络信息中心
12	Modeling and Design of Enhanced All Optical Signal Regeneration Technique	中国科学院计算机网络信息中心
13	A Deep Hybrid Neural Network for Single Image Dehazing via Wavelet Transform	中国科学院计算机网络信息中心
14	Multi-scale Single Image Super-Resolution with Remote-Sensing Application using Transferred Wide Residual Network	中国科学院计算机网络信息中心
15	基于词嵌入的国家自然科学基金学科交叉知识发现方法	中国科学院计算机网络信息中心
16	科学数据安全标准规范关键问题探索	中国科学院计算机网络信息中心
17	面向烟草领域的科研知识图谱服务平台关键技术研究	中国科学院计算机网络信息中心
18	数据降维及聚类算法在烟叶相似性分析中的应用	中国科学院计算机网络信息中心
19	轻量级遥感数据分布式调度框架-DataboxMR	中国科学院计算机网络信息中心
20	全球科学数据出版发展态势分析	中国科学院计算机网络信息中心
21	LightCAKE: A Lightweight Framework for Context-Aware Knowledge Graph Embedding	中国科学院计算机网络信息中心
22	xpert Knowledge-Guided Length-Variant Hierarchical Label Generation for Proposal Classification	中国科学院计算机网络信息中心
23	Research on Embedding Unsupervised Learning and Application on Tobacco Leaf Data	中国科学院计算机网络信息中心
24	iAnalysis V1.0: An Interactive Analysis Service System	中国科学院计算机网络信息中心
25	Integrating Representation and Imitation in Reinforcement Learning for Mobile User Profiling: An Adversarial Training Perspective	中国科学院计算机网络信息中心
26	油气资源开发的大数据智能平台及应用分析	中国科学院计算机网络信息中心
27	基于词嵌入的国家自然科学基金学科交叉知识发现方法——以“人工智能”与“信息管理”为例	中国科学院计算机网络信息中心
28	Digital image watermarking in sparse domain	中国科学院计算机网络信息中心
29	数据视域下的布洛陀学术研究趋势与知识图谱分析	中国科学院计算机网络信息中心
30	探析大数据技术的发展及其在安防领域的作用	中国科学院计算机网络信息中心

运行保障

配套设施

国家基础数据中心依托中国科技云开展基础设施建设和数据服务。中国科技云是基于云计算的国家科研信息化基础设施，持续开放汇聚国内外优秀信息化基础设施和服务。基础科学数据中心通过与其在技术创新和集成创新等方面的合作，推动国家基础学科数据云建设，实现基础学科各类资源融合、智能运营和动态调度，提供安全、按需、智能化的云服务。

- 315PFlops 计算资源
- 150PB 存储资源
- 52个 综合平台
- 409个 科研软件
- 9类 典型云服务
- 840万+ 免费CPU计算机时
- 160万+ GPU计算机时
- 1000台 云服务器
- 10PB 云备份资源

服务

网络传输
云计算
云存储
通用性大数据处理环境
高性能计算网格
人工智能计算
数据服务

资源

关联搜索
共享文献
知识产权
领域云数据库

超融合通讯服务

电子邮件
即时通讯
在线会议

地域分布



资源安全保障

国家基础数据中心参照相关国家标准对平台的基础设施与网络环境、技术实现、数据保护、管理制度体系等方面进行整体设计。

执行国家标准：

《GB／T 35274–2017 信息安全技术 大数据服务安全能力要求》

《GB/T 22239–2008信息安全技术 信息系统安全等级保护基本要求》



- 建设数据全程可控的可信数据存储库
- 建立科学数据应急管理灾备和长期保存机制
- 积极开展科学数据安全框架和区块链技术研究