



中国科学院
CHINESE ACADEMY OF SCIENCES



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

Research Data Infrastructure of CAS :Practices and Challenges

Dr. Jianhui LI(lijh@cnic.cn)

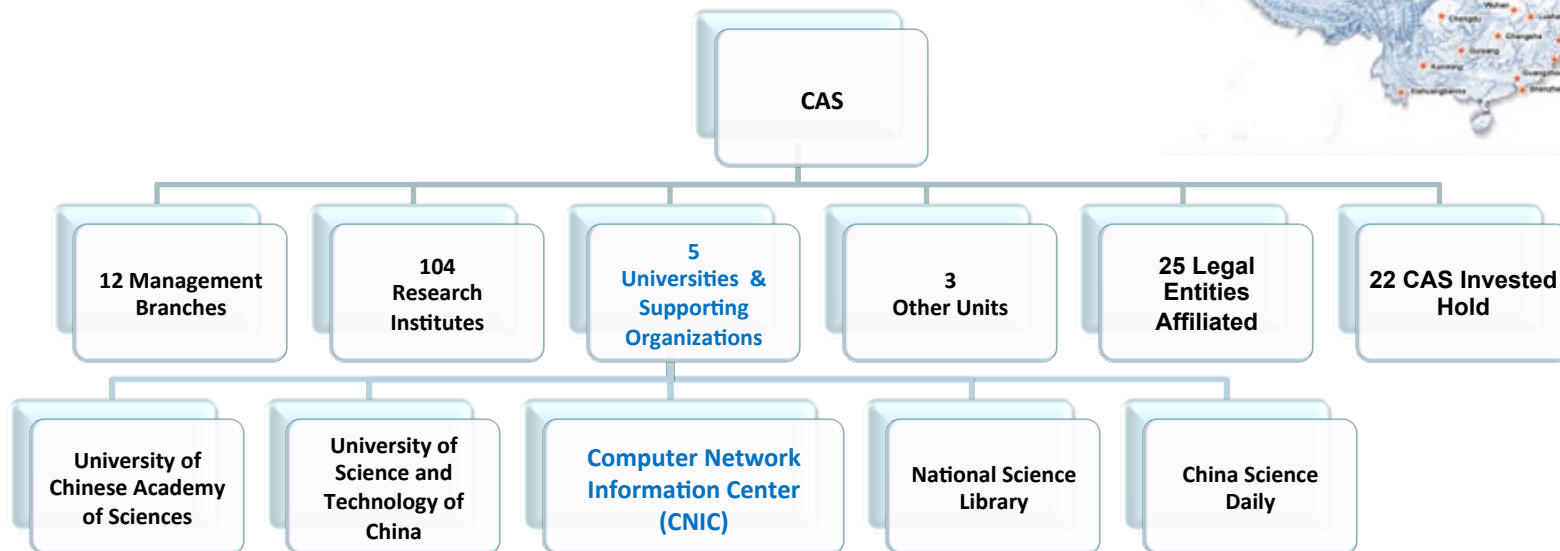
Computer Network Information Center

Chinese Academy of Sciences



CNIC in CAS

Founded in November 1949, **Chinese Academy of Sciences (CAS)** is the primary academic institution in China in the natural sciences. It is also China's largest comprehensive R&D organization in the natural sciences and high technology as well as the country's foremost science and technology advisory body.



Founded in April 1995, **Computer Network Information Center (CNIC)** is supporting institute involved in constructing and operating IT infrastructure and providing IT-related services. In addition, it serves as an R&D and demonstration base for IT technology applications.



中国科学院
CHINESE ACADEMY OF SCIENCES

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

CNIC Located in zhongguancun, Beijing



Zhongguancun, Haidian District , Beijing



Data Center located in
HuaiRou District, Beijing





CNIC's History and Mission



**First router
developed in China**



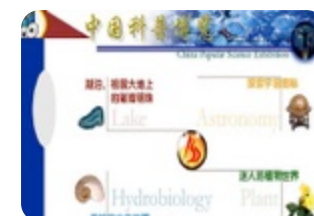
**China's first .CN
Domain Name
Server**



**China's earliest
HPC service**



**Earliest CAS
Scientific
Database**



**Earliest CAS Virtual
Science Museums**

The State Planning Commission approved the initiation of Scientific Database Project.

N C F C achieved full-functional connection to the Internet.

.CN Domain Name Server was established and began to provide registration service.

The State Science and Technology Commission approved the founding of CNIC.

CNIC started the earliest scientific computing application and service in China.

C N I C initiated Virtual Science Museums of China.

CNIC constructed the early version of the CAS website system.

CNIC provided operation support to Academia Resource Planning system of the CAS.

The State Development and Reform Commission approved the initiation of China Internet of Things Name Service Platform.

1986 1994.4 1994.5 1995.3 1996 1999 2001 2002 2008 2013 2015



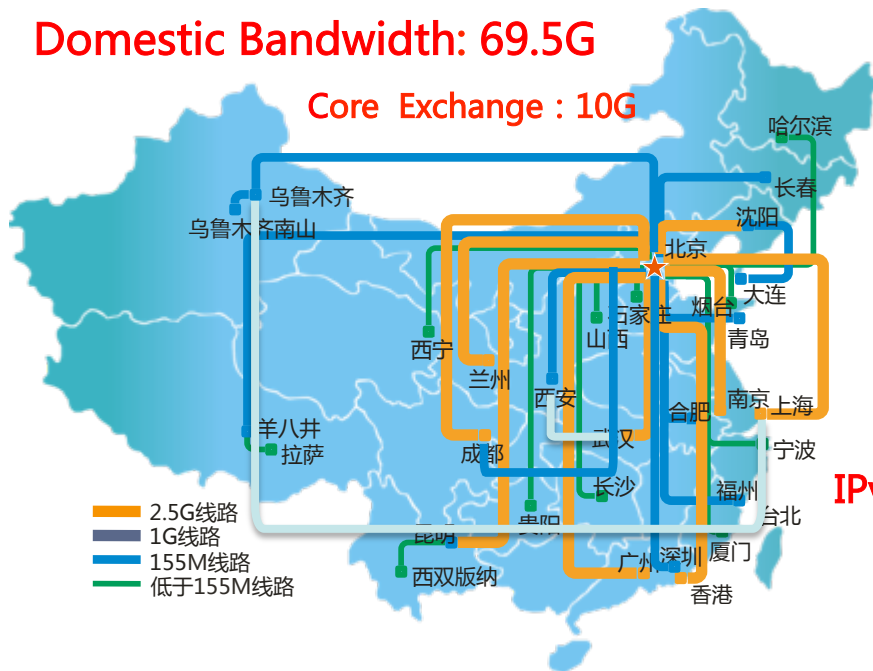
Internet



CSTNet

Domestic Bandwidth: 69.5G

Core Exchange : 10G



International Bandwidth : 27.5G



USA : 10G
EU : 10G

IPv4/IPv6

OrientPlus: 中国与欧洲科研网络连接的桥梁
(GLORIAD)



4G



5G



中国移动通信
CHINA MOBILE

2.5G



中国教育网
www.edu.cn

32G



10G



10G



10G

- 30 provinces
- 370 + Organizations
- Above 1M Users



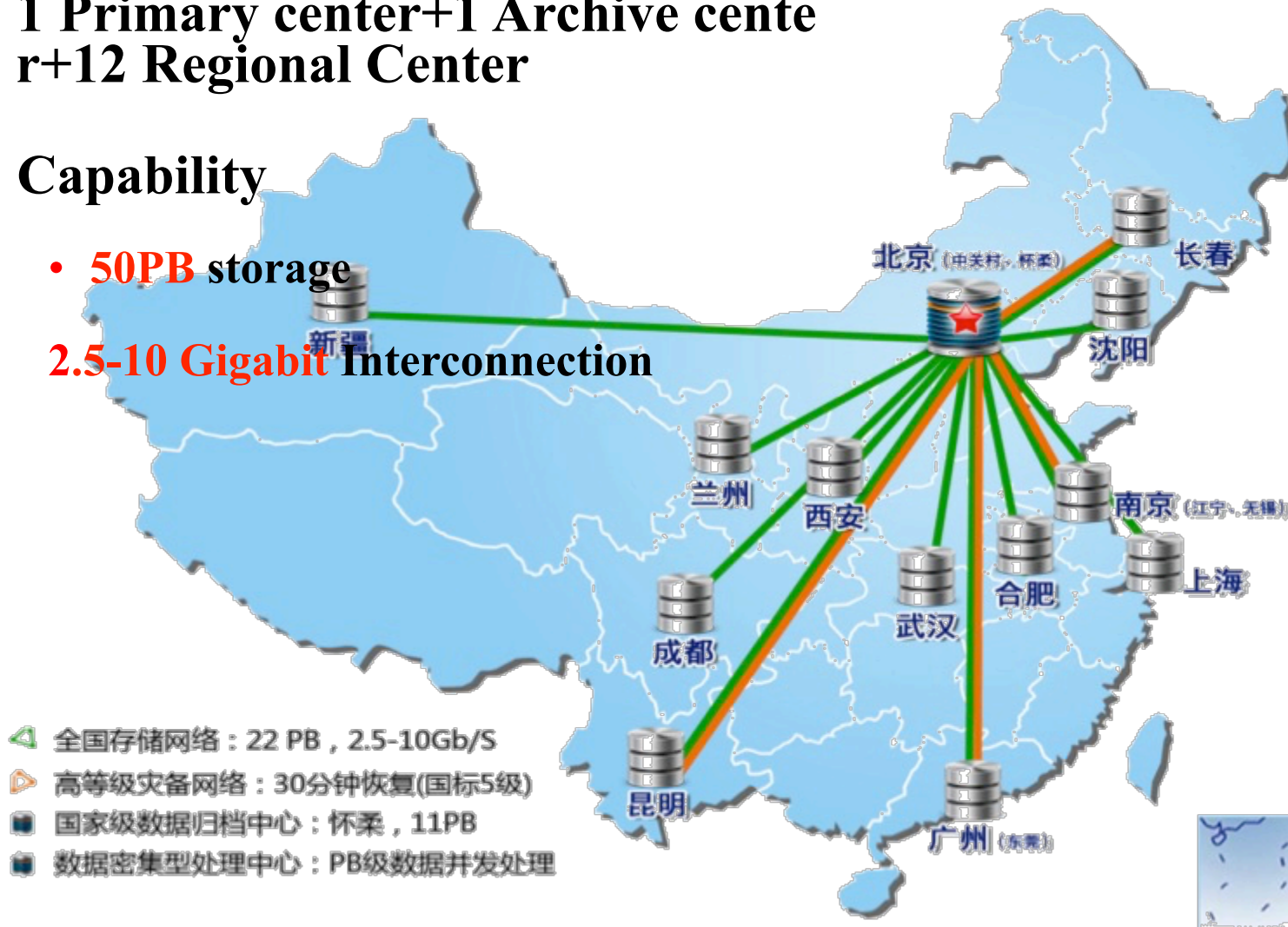
Massive Storage System

– 1 Primary center+1 Archive center
r+12 Regional Center

– Capability

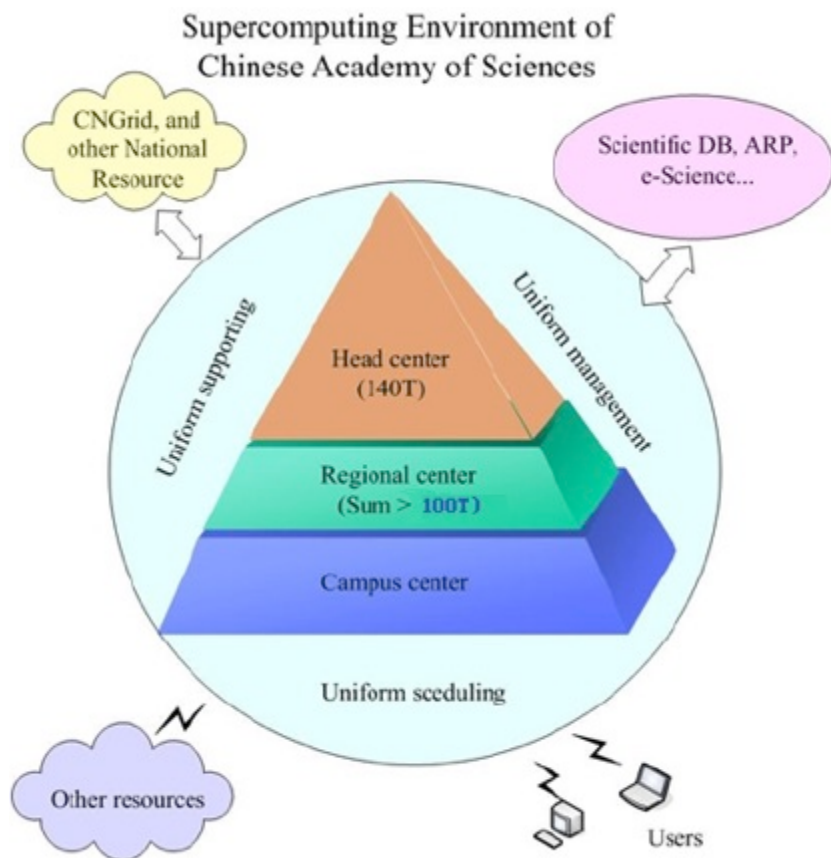
• 50PB storage

2.5-10 Gigabit Interconnection





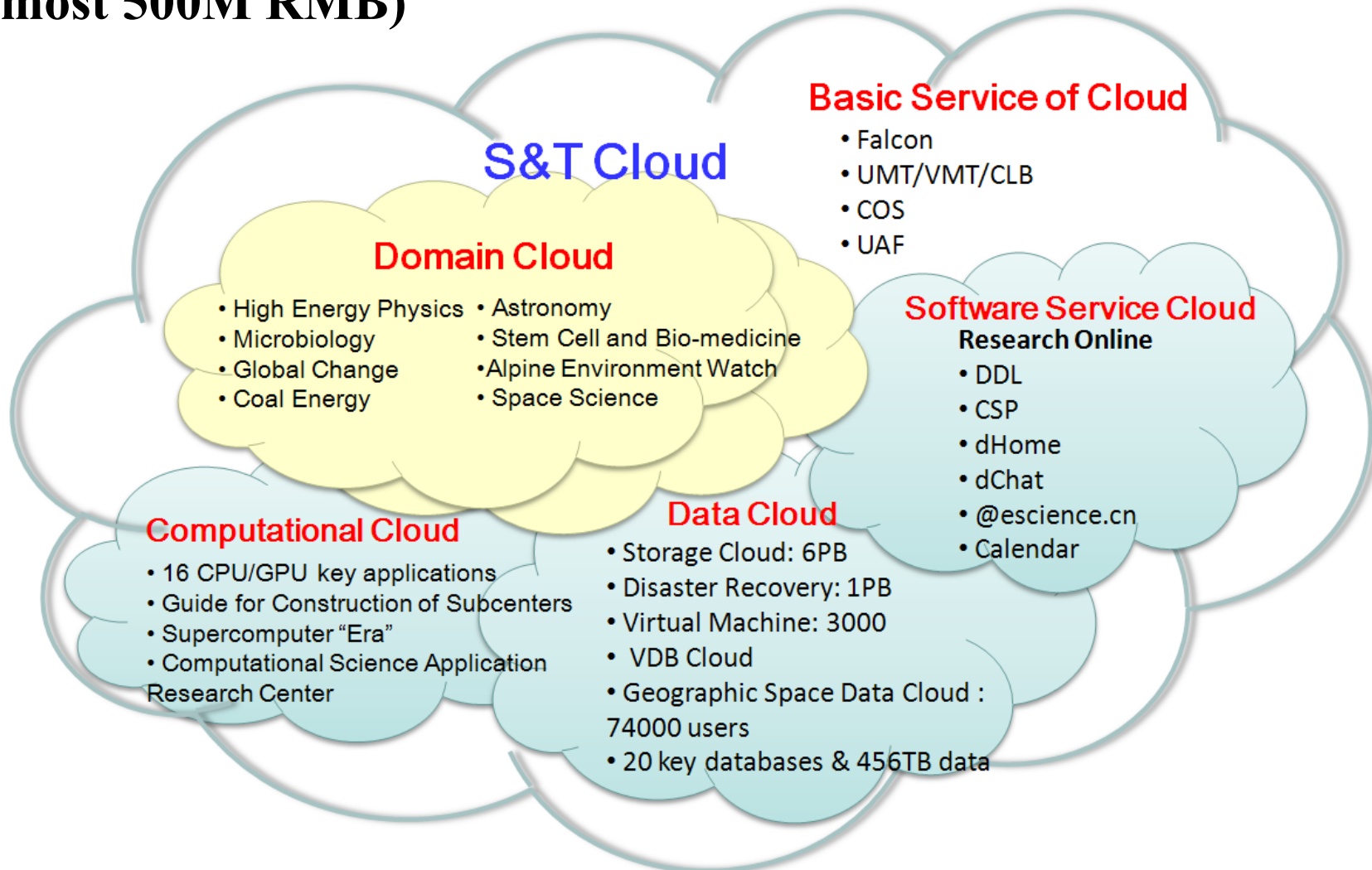
Computing Grid in CAS



- Aggregated Computing capacity: 1.3Pflops(CPU)+3Pfops (GPU)
- Tier0: 1. 2Pflops
- Tier1: >100Tflops
 - 10 nodes
 - general or specific Purpose
- Tier2: 50Tflops
 - ~20 institutes of CAS



Informatization Program of CAS from 2010-2015 (budget is almost 500M RMB)





Why We Need Research data Infrastructure

- Large scientific facilities produce huge data
 - +20 being operation
 - +20 under construction
- Long-Term field observation stations
 - +100 stations including Ecology, Environment, Space, etc.
- Long-Term Research data need to be archived and curation and sharing
 - 100+ institutes



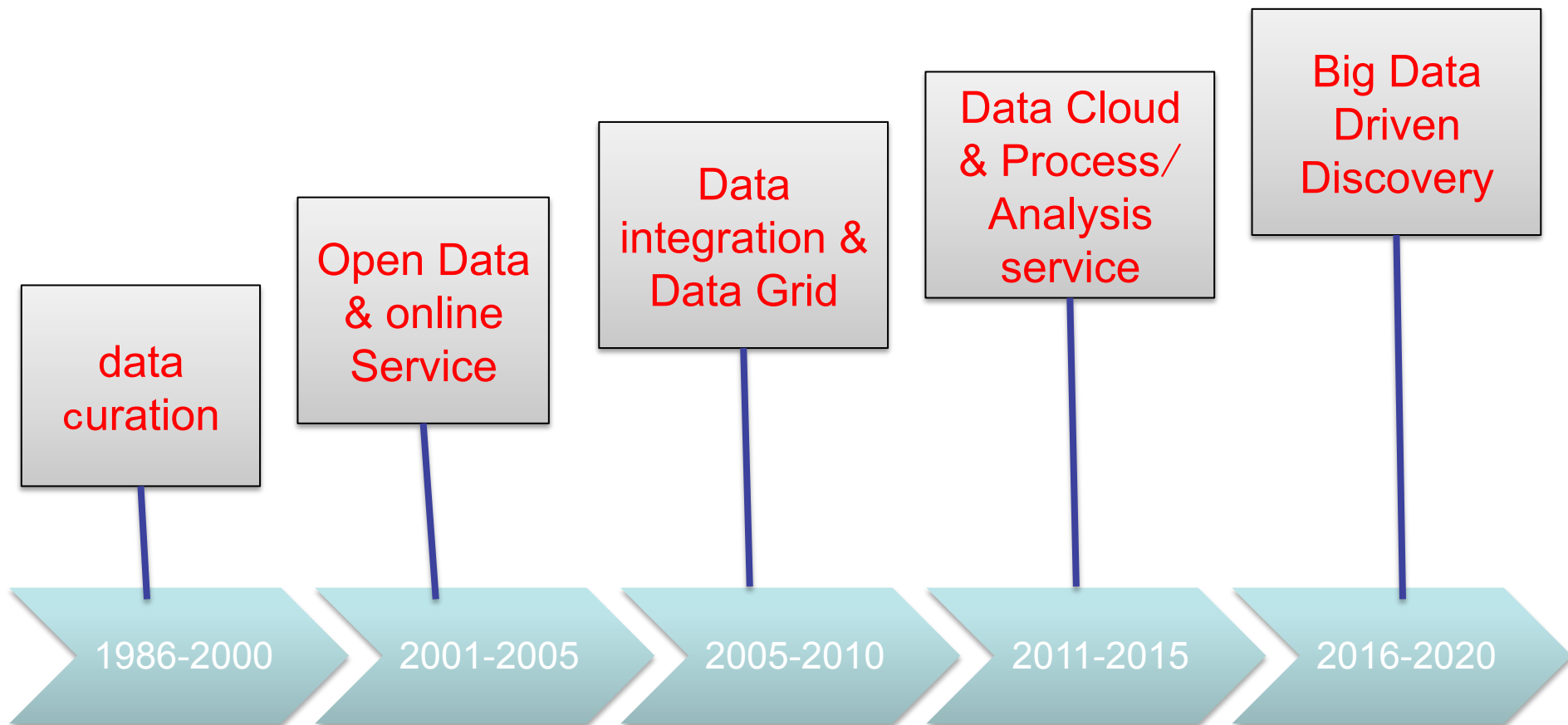
Field observation stations



Large Scientific facilities

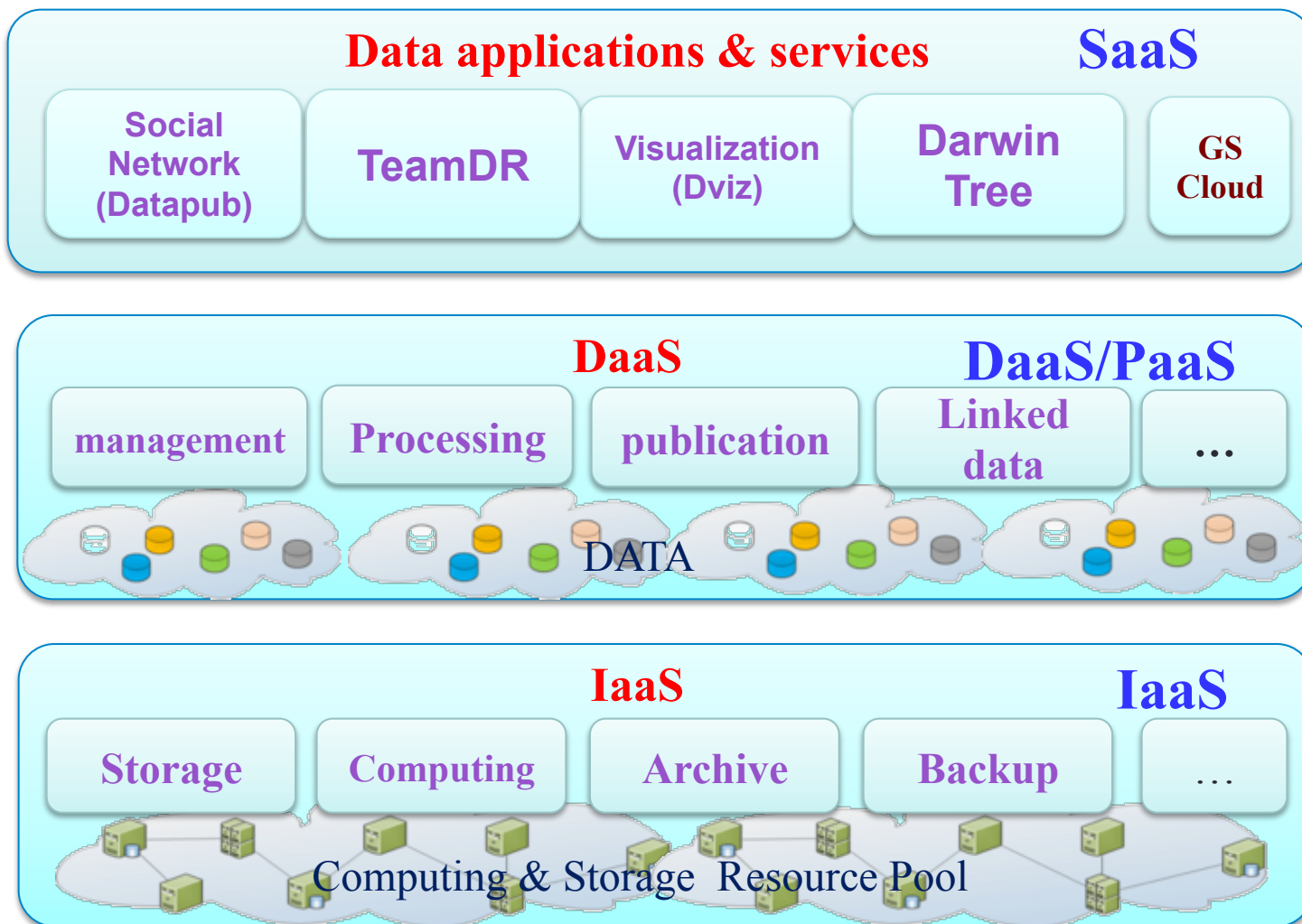


Evolution of Research Data Infrastructure in CAS





Data Cloud of CAS





Data Cloud Portal



www.csdb.cn

[登录](#) [注册](#)

您好！欢迎您来到中国科学院数据云

[首页](#)

[基础设施云服务](#)

[科学数据云服务](#)

[数据应用云服务](#)

[关于我们](#)



数据应用云服务

[服务资源](#)

[服务状况](#)

[快速查找数据库](#)

[数据服务接口检索](#)

[查找论文](#)



DVIz可视化发布平台
所见即所得的方式实现数据的
可视化快速生成



科技在线推送
自动、准确推送最新论文、专
利和项目信息



DataPub数据共享社区
数据社交平台，加快数据共享
与交流



Darwin Tree
系统生物进化的数据环境，模
型分析、生命之树、分子鉴定

服务公告

新闻动态

- > 毫米波射电天文数据库开放20...
- > 黄土高原生态环境数据库新...
- > 关于科技数据资源整合与共享工...
- > 中科院数据工程科学数据库课...
- > 关于举办2015年数据工程技...

[更多>>](#)

服务案例

- > 多功能景观 [2015-11-03]
- > 湖泊对长江中下游地区... [2015-11-03]
- > 冰冻圈变化及其影响研... [2015-11-03]
- > 江西省洛溪口水闸枢纽... [2015-11-03]
- > 旅游地聚居空间演化过... [2015-11-02]

[更多>>](#)

科学数据云服务

[服务资源](#)

[服务状况](#)

[科学数据库资源](#)

[十二五数据库项目资源](#)




**VDBCloud数据管
理平台**
在线数据管理, 定制
化数据发布平台

[简介](#) [立即体验](#)



地理空间数据云
数据、存储、处理模
型, 一站式数据服务
平台

[简介](#) [立即体验](#)



科学数据出版
在线数据论文期刊,
致力于科学数据的
快速出版与传播

[简介](#) [立即体验](#)



科学文献与科学...
科学文献与科学数
据关联检索示范平
台

[简介](#) [立即体验](#)

十二五数据库服务监控与统计

累计访问人次: 3672082


累计网页访问量: 75631656

累计下载量: 23.49TB

基础设施云服务

[服务资源](#)

[服务状况](#)




云存储
提供海量、弹性、高可用、高性价比的云存储服务, 任何地方任何时间可以通过网络访问存储数据
[立即体验](#)



云计算
面向政府、科研单位、企业、门户网站提供安全可靠、成熟稳定、高性价比的私有云计算解决方案
[立即体验](#)



云归档
将用户不再需要被常规访问的数据移到一个单独的私有、专用、安全的云存储环境来进行长期保存的
服务



云灾备
针对生产、管理和关键业务系统提供先进的灾备平台和多种灾备模式为用户提供安全、可靠、稳定性高的“两地三中心”数据灾备服务



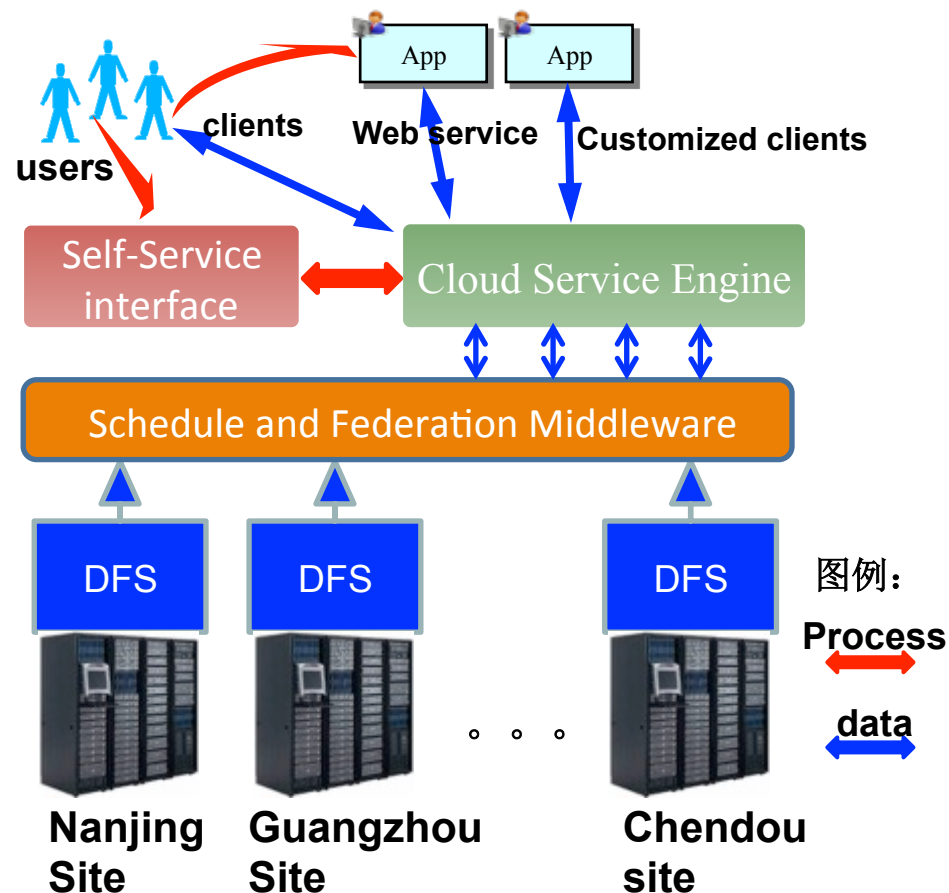
“课题数据宝上线”
欢迎使用!!

[分布式存储架构](#)



IaaS

- Computing Cloud
 - ZeStack based on OpenStack
 - Multi-domain resource management and Federation service
- Storage Cloud
 - Distributed File System
 - Providing one-surface federation service

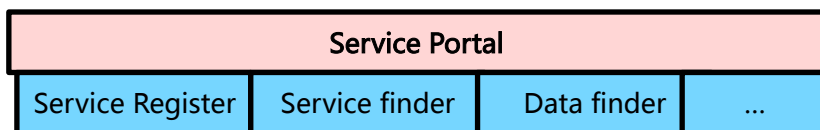




DaaS

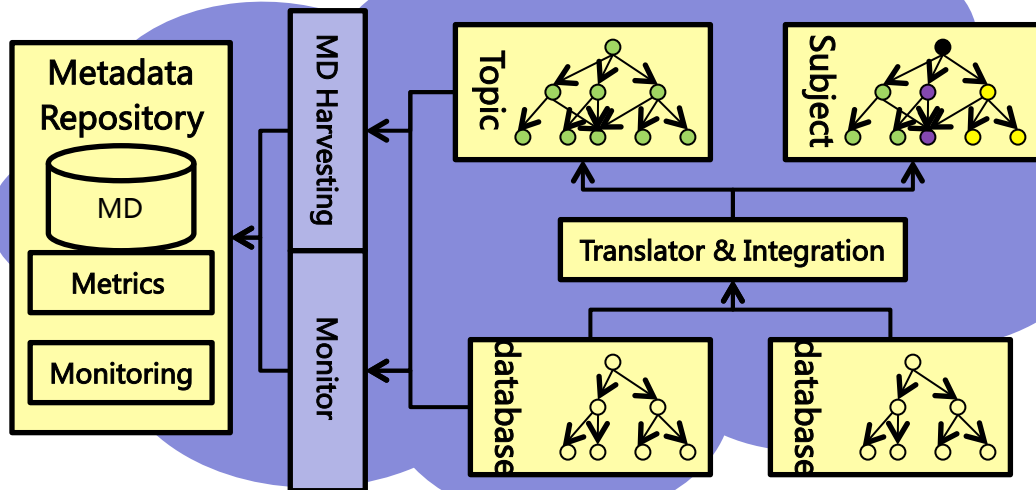


Data services



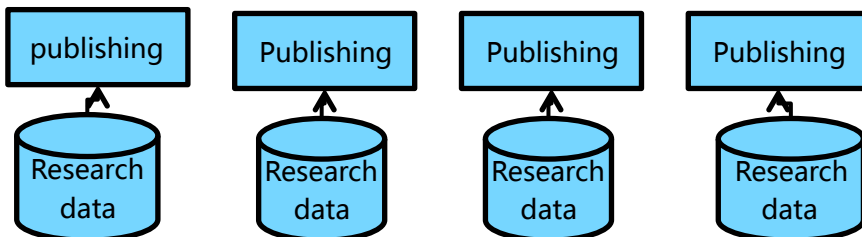
Service access and federating Middleware

Data link & Integrating



Access Brokers

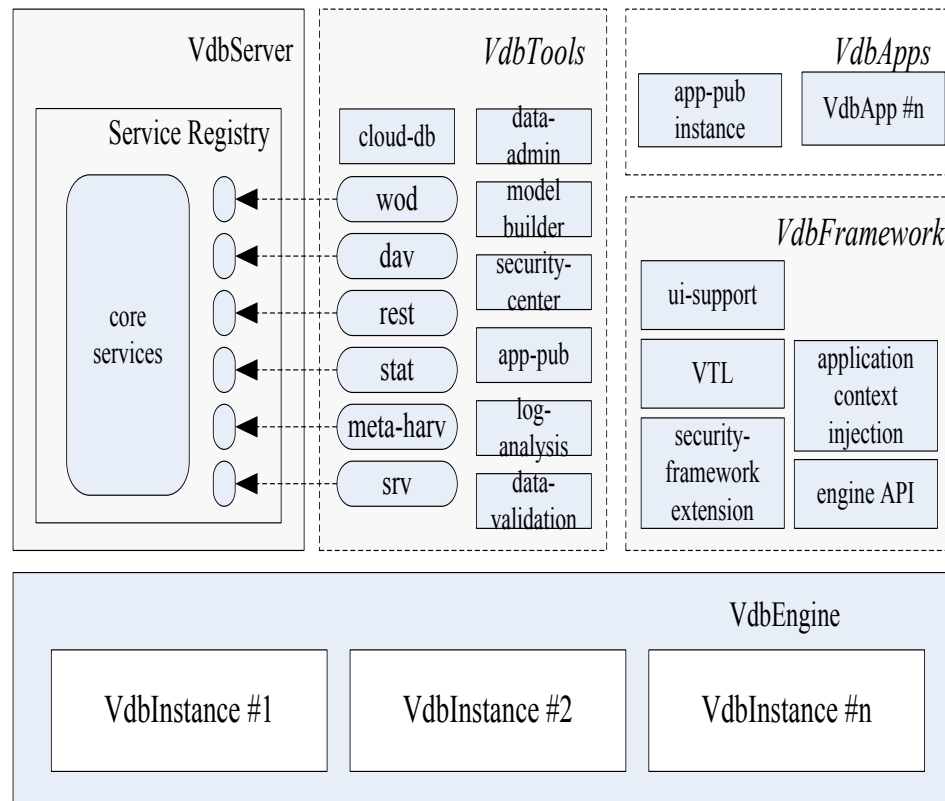
Distributed Data Sites





VDB/VDBCloud

- ① • An invisible **VdbEngine** which organizes all data into a whole database
- ② • A set of **VdbTools** which help data managers manage and publish their data
- ③ • A **VdbServer** which serves local database as an accessible data source on the Web
- ④ • A **VdbFramework** which help developers build applications (VdbApps) based on VisualDB





• VdbTools - ModelBuilder

新增/编辑字段

字段类型选择

- 12 整数类型
- 1.2 实数类型
- Az 字符类型
- 日期类型
- 化学结构
- 联合字段
- 文件类型
- 枚举类型
- Gps类型
- 集合类型
- 引用类型

字段配置

名称:

标题:

强关联: 是 否

是否是主键:

是否作为标题字段:

请选择映射类型: 关联表集合字

字段映射配置

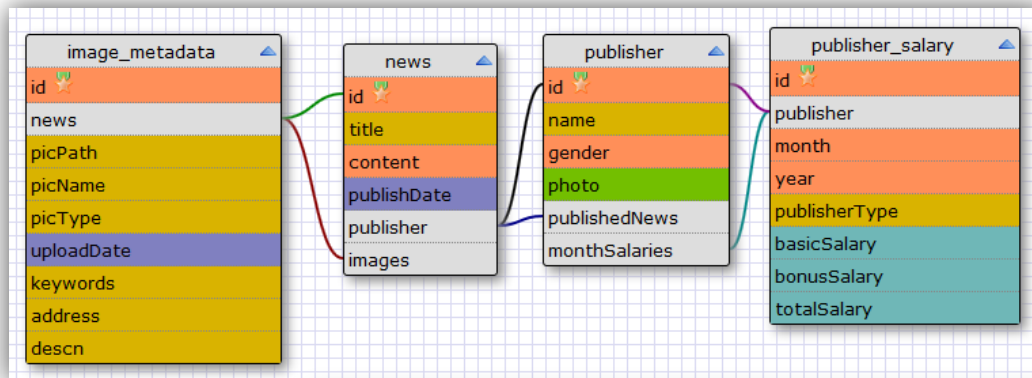
目标实体:

关联表所在存储位置:

关联表:

本端关联列:

目标端关联列:



构建用户导航目录树

- 数据集
 - 新闻表
- 位置根路径
- 显示文件
- conference

关系实体

- news
- image_metadata
- publisher
- conference
- publisher_salary
- contact_basic
- contact_phone_number
- users
- student
- course
- 班级
- 课程
- 学生
- 文件实体
 - D:/
 - 元数据表
 - sql_image_metadata

新建表

数据库物理建模-建表

表名:

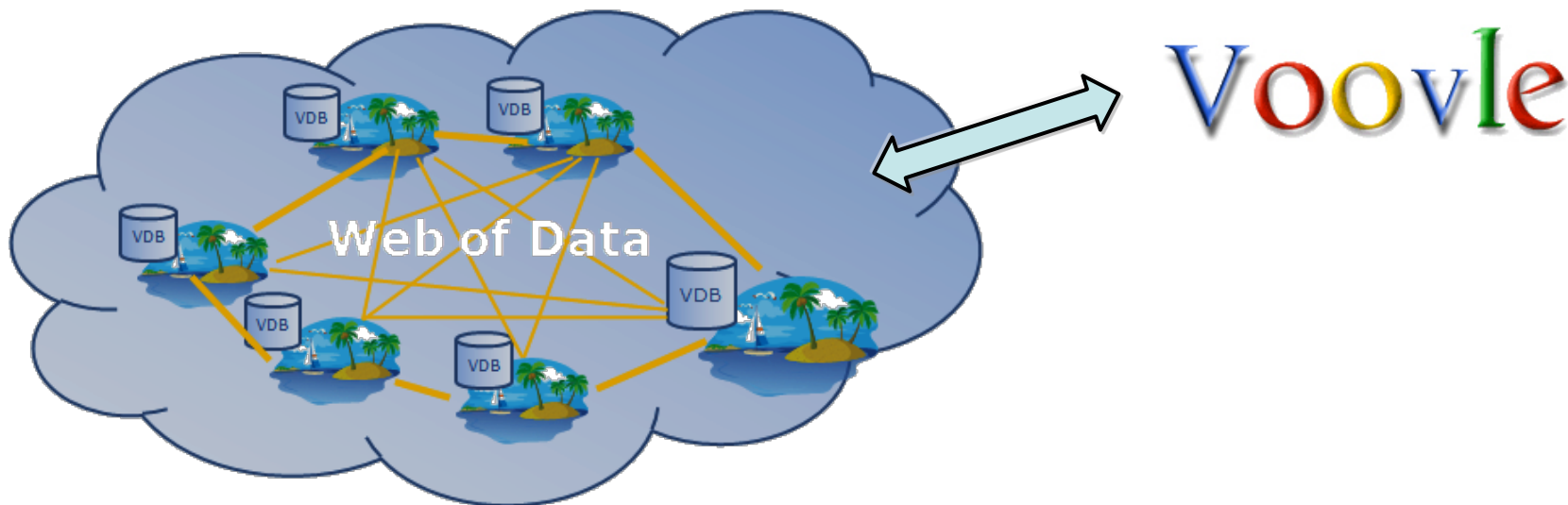
增加字段

	字段名	注释	数据类型	长度	允许为空
1	newField	新字段	CHAR		<input checked="" type="checkbox"/>
2	newField	新字段	CHAR		<input checked="" type="checkbox"/>
3	newField	新字段	CHAR		<input checked="" type="checkbox"/>
4	newField	新字段	CHAR		<input checked="" type="checkbox"/>
5	newField	新字段	CHAR		<input checked="" type="checkbox"/>



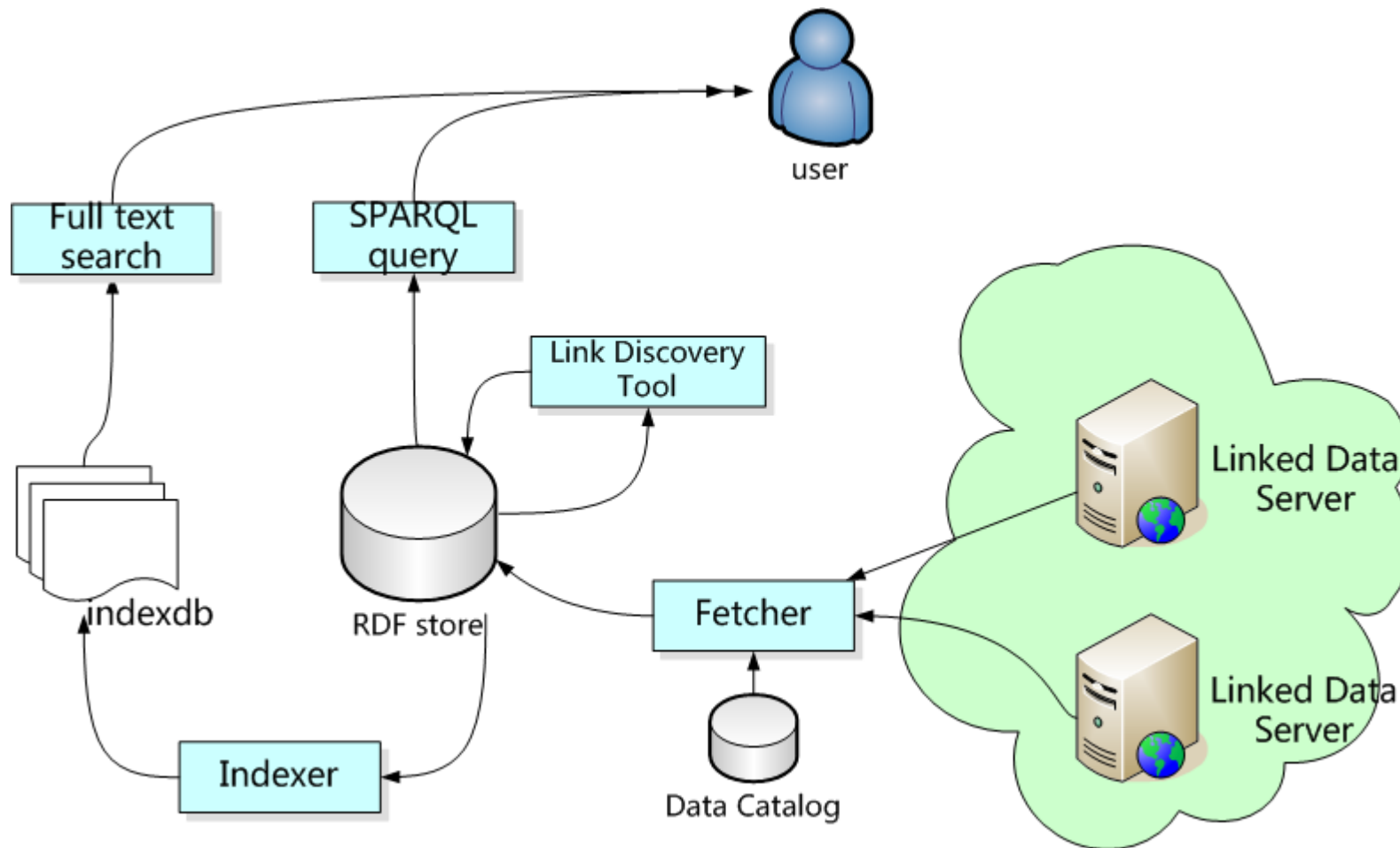
Finding Data Services

- Voovle fetches all scientific data from distributed databases via interfaces exposed by middleware, builds a large data store, creates index, and provides search interfaces for end users.



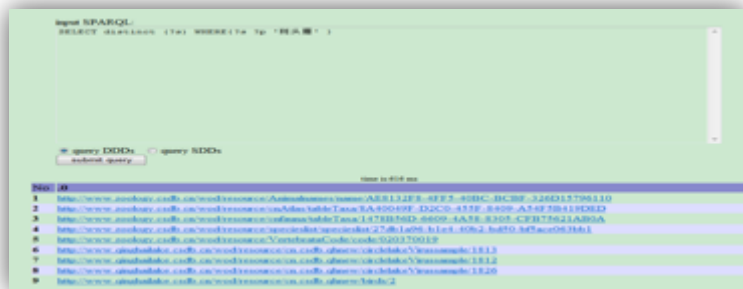


Architecture of Voovle





Finding Data



- SPARQL query interface
- free text search
 - Voovle use Jena + Lucene to index and search data descriptions

Voovle

斑头雁

检索关键字示例: 斑头雁 棉花 红土





linking data

标题	红土
form	chemphy2_2
class_id	8
valume_id	valum2_2
土种名称	红土
soil_id	20422
一般性描述	1. 归属与分布 红土, 属红粘土亚类红土土属。零星分布于辽宁省丘陵地区, 以大连、丹东、营口、朝阳市。现全部开垦为耕地, 面积66.7万亩。2. 主要性状 该土种母质为红色粘土堆积物, 土体深厚, 层次分明, 红色率由表层向下逐渐增加, 平均红色率3.94; A11层质地砂质粘壤土, C层壤质粘土, 粘粒含量A11率2.68, 硅铝率3.35, 铁的游离度36.8%。土壤多呈中性反应, pH6.5—7.8, 阳离子交换量通体较高, 在17个农化样分析: 有机质含量1.12%, 全氮0.067%, 速效磷6ppm, 速效钾123ppm。有效微量元素硼0.64ppm, 铁4ppm。3. 典型剖面 采自大连市甘井子区苏家村校前。低丘缓坡中部, 坡度10°, 海拔56m。中度片蚀, 水量640mm, ≥10℃积温3620℃, 无霜期201天。种植玉米。A11层: 0—20cm, 油棕色(湿, 7.5YR5/3), 多, pH7.5。C1层: 20—50cm, 油红棕色(湿, 5YR5/3), 砂质粘壤土, 块状结构, 较紧, 有植物根系, pH5YR4/3, 砂质粘壤土, 块状结构, 紧实, 植物根系极少, 有少量铁锰胶膜, pH7.8。C3层: 85—130cm, 构, 极紧, 无植物根系, 中量铁锰胶膜, pH7.8。4. 生产性能综述 该土种土体深厚, 质地上壤下粘, 结构出苗。另外耕作层薄, 养分含量缺乏, 尤其磷素极缺。适种作物窄, 产量不高。目前种植高粱、谷子、措施: 应加强深翻、深松、耙地、压地、铲等管理措施, 增厚耕作层, 改良土壤物理性状; 合理施用有机质含量; 因地制宜, 发展果树生产, 一般以苹果和山楂为宜。

Descriptive information of red soil, including soil class no., soil class name, general description, etc.

关联数据

- [1] 亚类: 红粘土. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/soilSubclass/70>. [2011-01-23 17:14:33]
- [2] 县市名: 大连. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/sublocation/587>. [2011-01-24 05:02:33]
- [3] 县市名: 丹东. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/sublocation/589>. [2011-01-24 05:02:37]
- [4] 县市名: 营口. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/sublocation/590>. [2011-01-24 05:02:37]
- [5] 县市名: 朝阳. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/sublocation/1675>. [2011-01-24 05:24:44]
- [6] 理化性质表: 2435. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/valum1/2435>. [2011-01-24 06:13:48]
- [7] 理化性质表: 2433. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/valum1/2433>. [2011-01-24 06:13:44]
- [8] 理化性质表: 2436. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/valum1/2436>. [2011-01-24 06:13:53]
- [9] 理化性质表: 2434. [EB/OL] <http://www.soil.csd.cn/wod/resource/cn.csd.soil.soiltype/valum1/2434>. [2011-01-24 06:13:47]

subclass

geographical distribution

physical and chemical properties



Scientific Databases (SDB)

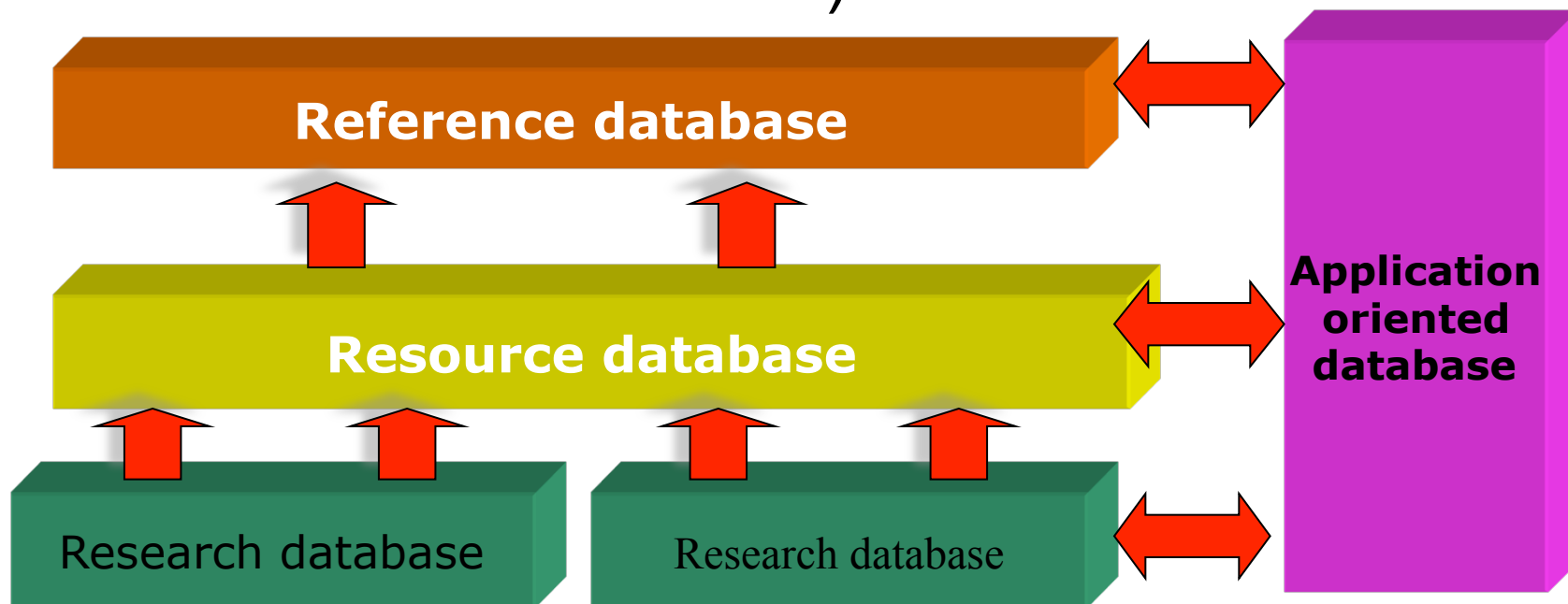
- A Long-term mission started in 1986 which funded by CAS
 - many institutes involved
 - long-term, large-scale collaboration
 - data from research, for research
- Collecting multi-discipline research data and promoting data sharing
 - More than **350** research databases and **1350** datasets by 61 institutes
 - Over **60TB** data available to open access and download





Scientific Databases (cont.)

- focusing on data integration and improving research database to be resource database and even reference database)



Scientific Databases (cont.)

- **Resource databases**
 - Geo-Science
 - Biodiversity
 - Chemistry
 - Astronomy
 - Space Science
 - Micro biology and virus
 - Material science
 - Environment
- **Reference databases**
 - China Species
 - Compound
- **Application-Oriented databases**
 - High Energy (ITER)
 - Western Environment Research
 - **Ecology research**
 - Qinghai Lake Resource databases
- **Research databases**
 - Physics & Chemistry, Geosciences, Biosciences, Atmospheric & Ocean Science, Energy Science, Material Science, Astronomy & Space Science



Typical Case1: Geospatial Data

- A open data cloud for geospatial data
 - Collecting Global Data sources dynamically
 - Coverage main focus on China and surrounding area
 - Make data Findable, Accessible , Interoperable Reusable (FAIR) and Citable





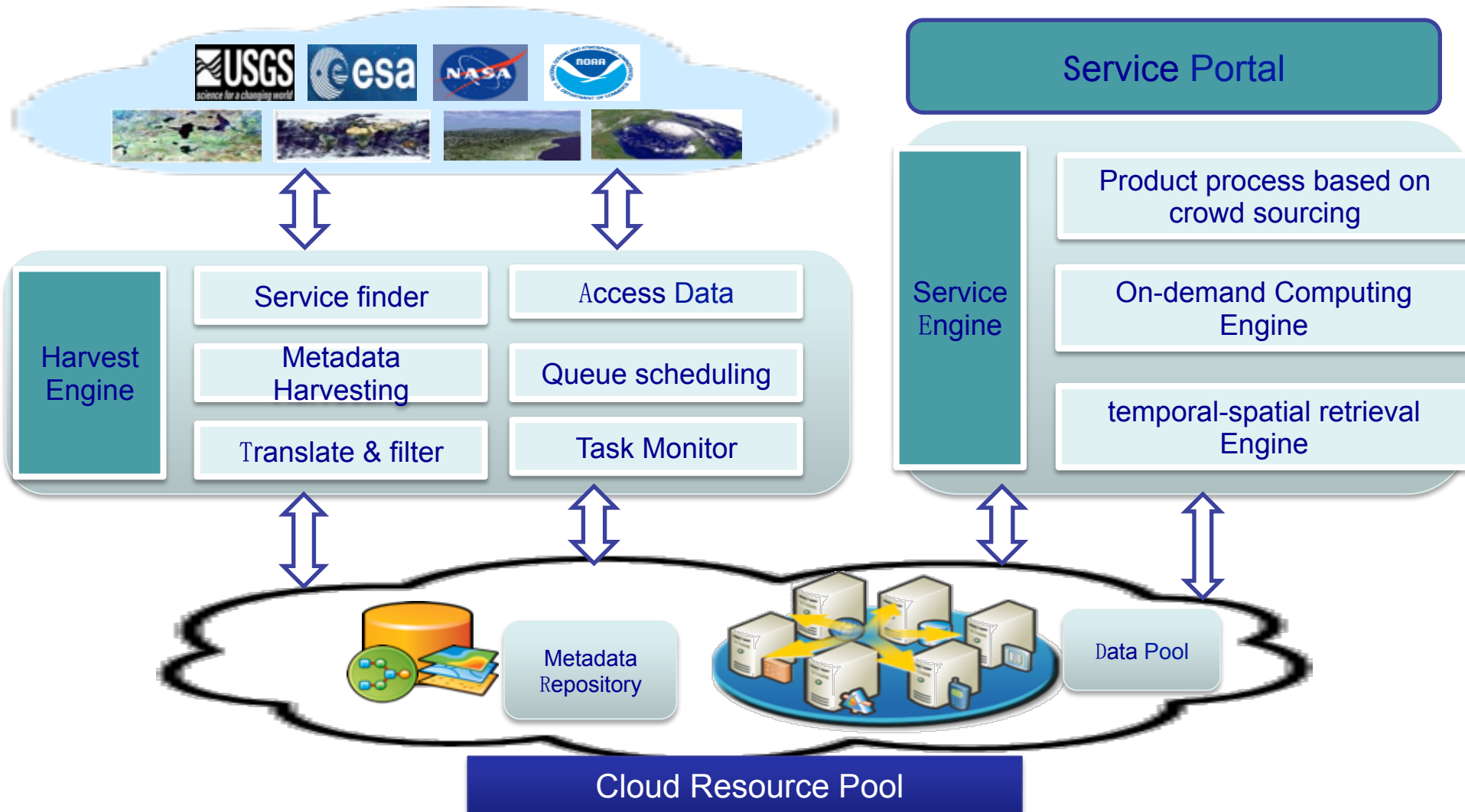
Data Sources



open Data which can be harvested online



GSCloud Architecture





[Http://www.gscloud.cn](http://www.gscloud.cn)

地理空间数据云 Geospatial Data Cloud 高级检索 数据资源 数据众包 在线计算 平台信息

GSCloud 退出

ICARTO 在线云制图

地图记录你的足迹、分享展现你的精彩
· 轻松成为地图制作专家!

用户总数：153066

今日注册：122

13:00-14:00注册：12

总数据量：438.7 TB

[更多>](#)

免费数据 免费数据，持续更新，触手可及

[更多>](#)



LANDSAT系列数据



MODIS系列数据



MODIS中国合成产品



MODIS L1B标准产品



DEM数字高程数据



EO-1系列数据

商业数据 商业卫星数据强势入驻，遥感数据一站式服务从这里开始

[更多>](#)



高分一号



高分二号



资源3号



更多数据




平台首页 / 数据资源

返回收藏: [收藏](#) [取消](#) [删除](#)


关键字搜索

显示第 1 到 12 条, 共 102 条记录



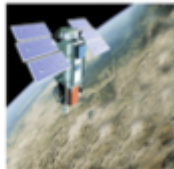
Landsat 8 OLI_TIRS 卫星数字产品

[详细信息](#) [数据列表](#)



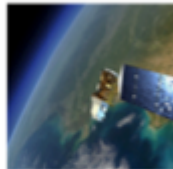
Landsat 7 ETM SLC-off 卫星数字产品(2000-)

[详细信息](#) [数据列表](#)



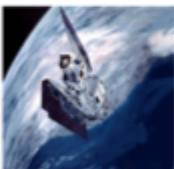
Landsat 7 ETM SLC-on 卫星数字产品(1999-2002)

[详细信息](#) [数据列表](#)




Landsat 4 TM 卫星数字产品

[详细信息](#) [数据列表](#)




Landsat 5 TM 卫星数字产品

[详细信息](#) [数据列表](#)




Landsat 3 MSS 卫星数字产品

[详细信息](#) [数据列表](#)



Landsat 中国内陆水体信息产品

[详细信息](#) [数据列表](#)



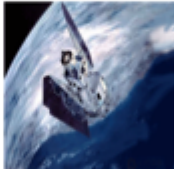
Landsat 全球合成数据(1999-2002)

[详细信息](#) [数据列表](#)

平台首页 / 数据资源 / 高分一号


关键字搜索

显示第 1 到 2 条, 共 2 条记录



高分一号 pms

[详细信息](#) [数据列表](#)



高分一号 wfv

[详细信息](#) [数据列表](#)

Data resources -- 438.7TB

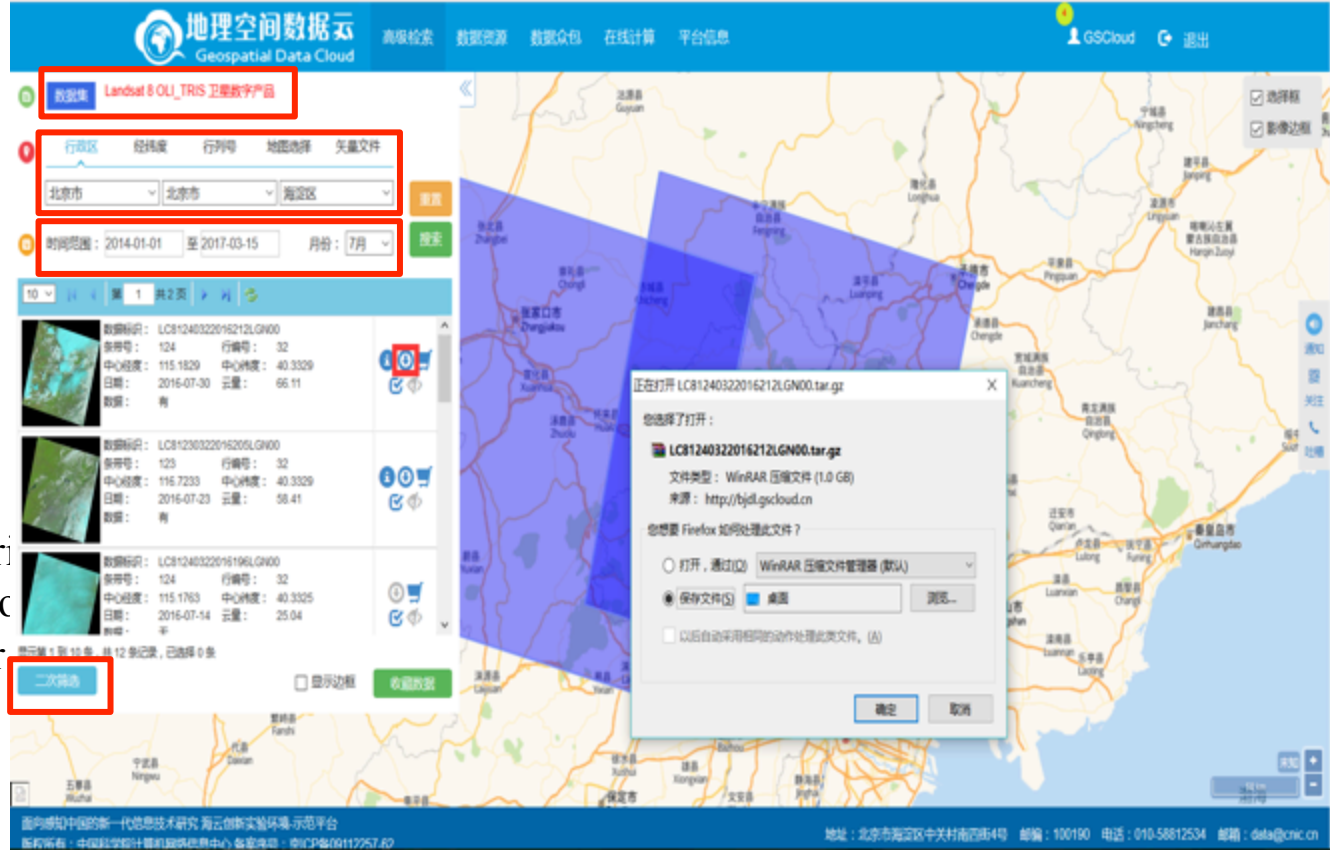
总数据量: **438.7 TB** 昨日下载: **475.6 GB**

#	产品名称	数据量	昨日下载
1	LANDSAT	75.3 TB	433.8 GB
2	SENTINELS	15.6 TB	18.2 GB
3	Mosaic_Landsat	305.4 GB	6.6 GB
4	MODIS_PRODUCT_CN	12.4 TB	6.3 GB
5	DEM	1.6 TB	5.3 GB
6	Mosaic_Landsat1984-1997	511.1 GB	2.8 GB
7	modis_land	96.9 TB	2.4 GB
8	MODIS_L1B	200.5 TB	235.2 MB
9	LANDSAT_WATER	2.5 GB	7.7 MB
10	EO_1	3.8 TB	0 bytes
11	ENVIRON_DATA	23.2 TB	0 bytes
12	GLS	7.1 TB	0 bytes
13	global_change	1.5 TB	0 bytes

Data retrieve

Query options:

- Dataset Filter
- Geographic Filter
 - Administration boundary
 - latitude and longitude coordinate
 - Defined regions by user
 - Path and row
 - Shapefile
- temporal range Filter
- Metadata Filter



The screenshot shows the Geospatial Data Cloud interface. At the top, there are navigation links for '高级检索', '数据资源', '数据众包', '在线计算', and '平台信息'. The search filters are as follows:

- 数据集: Landsat 8 OLI_TRIS 卫星数字产品
- 行政区: 北京市 (北京市) 海淀区
- 时间范围: 2014-01-01 至 2017-03-15 月份: 7月

The search results list three datasets with their respective metadata. A '二次筛选' (Secondary Filter) button is highlighted. A dialog box is open, showing the download of a file named 'LC8124032201621ZLGN00.tar.gz' (1.0 GB) to the desktop.

Download options:

- Download: HTTP
- Bulk Download: FTP



The screenshot shows the FTP download options and a task list table. The FTP server information is as follows:

- FTP服务器: ftp://bjdl.gsccloud.cn
- FTP帐号: data@cnic.cn
- 推荐FTP客户端: Xftp 5(商业软件, 30天试用期), FileZilla(免费)

The task list table is as follows:

二次筛选	任务标识	任务名称	提交时间	操作
<input type="checkbox"/>	1	LA20170224151050301	20170205151047	2017-02-24 详情 删除
<input type="checkbox"/>	2	LA20160906092522925	20160902092521	2016-09-06 详情 删除



模型分类

遥感数据处理

地形信息提取

遥感数据应用

平台首页 / 在线计算

10 第 1 共 1 页

显示第 1 到 7 条, 共 7 条记录

编号	模型信息	说明	操作	缩略图
1	<p>地形坡向指数</p> <p>针对DEM数据的使用, 基于数字高程数据, 计算坡向 (Aspect)。</p> <p>评分: ★★★★★☆ 调用次数: 211次 创建时间: 2016-03-16 13:42</p>	详细	运行	
2	<p>地形阴影指数</p> <p>针对DEM数据的使用, 基于数字高程数据, 计算山体阴影 (Shaded relief)。</p> <p>评分: ★★★★★★ 调用次数: 280次 创建时间: 2016-03-16 13:42</p>	详细	运行	
3	<p>地形坡度指数</p> <p>针对DEM数据的使用, 基于数字高程数据, 计算坡度 (Slope)。坡度计算需要在投影坐标系统 (PCS) 下进行, 地理坐标系统(GCS)不支持坡度计算。</p> <p>评分: ★☆☆☆☆ 调用次数: 706次 创建时间: 2016-03-16 13:43</p>	详细	运行	
4	<p>地形坡位指数</p> <p>针对DEM数据的使用, 基于数字高程数据, 计算坡位 (Topographic Position Index)。坡位系数用来描述地形部位的一个地形参数, 在地貌分类中具有十分重要的意义, 用于确定研究目标点与其周围地形的位关系。</p> <p>评分: ★★★★★★ 调用次数: 546次 创建时间: 2016-03-16 13:44</p>	详细	运行	
5	<p>地形粗糙指数</p> <p>针对DEM数据的使用, 基于数字高程数据, 计算地形粗糙指数 (Terrain Ruggedness Index), 地表粗糙度是反映地表起伏变化与侵蚀程度的指标, 地面越粗糙, 地形粗糙指数越大。</p> <p>评分: ★★★★★☆ 调用次数: 737次 创建时间: 2016-03-16 13:46</p>	详细	运行	
	DEM高程数据切割			



任务名称

标题(必填): 北京市高程数据
您需要为云计算任务设定一个标题,以便今后查找!

模型参数

DEM数据源: GOTM 30米
目前提供90m分辨率SRM和30m分辨率GOTM数据

自由选择 行政区 手动输入 文件上传

北京市 地级市州 区县镇

空间范围(必填):
选择超出范围的空间范围, 当前下载大于 2.5万平方公里!

输出格式: GeoTIFF (*.tif)
目前系统支持GeoTIFF (.tif)和Erdas Imagine Images (.img)两种数据格式。

保存路径(必填): /test

1、 Set parameters of the model online:
For the DEM Clip Model, parameters include task name, data source (e.g. DEM of 30m and 90m resolution), spatial region, output file format, save path

2、 After executing the task, users can check the progress of the task in page “My tasks”.

3、 Results can be downloaded at last.

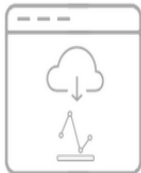
任务信息	
任务标识:	FA3589FA19B6638BCDECB38C296FED261CE8D9E
任务名称:	北京市高程数据
提交时间:	2017-01-17 10:44:31
开始时间:	2017-01-17 10:44:34
完成时间:	2017-01-17 10:45:25
任务状态:	结束
执行进度:	<div style="width: 100%; background-color: green;">100%</div>

结果列表 <small>(列表中文件可以在我的私人数据中找到并下载)</small>				
#	文件名	缩略图	大小	操作
1	201701171044340001.tif		83.9 MB	下载

数据众包 — 一平方公里精改专研 全球范围轻松玩转 您的需求就是我们的服务范围 众包众筹

[查看详情 >>](#)

需求汇



我有需求，GSCloud帮你定制

我需要遥感分类，我需要科学家手中的资源

我需要专题制图，我需要国际免费资源

我需要遥感反演计算，我需要高分遥感数据

我的需求更特殊，没关系，GSCloud帮你完成！

任务汇



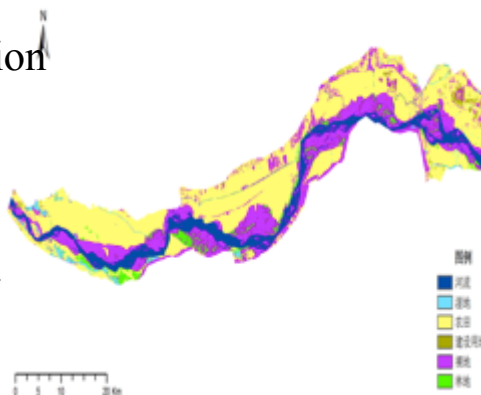
T2017030001: 基于Landsat影像的平台县植被覆盖度... [\[点击报名\]](#)

T2017020001: 基于遥感影像的杭州市萧山区部分区域土地... [\[报名结束\]](#)

T2017010001: 基于Landsat影像的崇礼县土地利用... [\[报名结束\]](#)

T2016120006: 基于Landsat影像的准格尔旗盐碱区域... [\[报名结束\]](#)

河南省黄河流域1988年土地利用类型分布图



数据汇



我有资源，GSCloud助你盘活

我有数据产品

我有地图产品

我有模型资源

资源是你的，收益也是你的，GSCloud只是帮你分发

Goals:

Extracting land use information from Remote Sensing images

Problem:

Difficult to do it automatically.
It is human computation task.

Steps:

- 1.Task Definition and Division
- 2.Recruitment and Talents Selection
- 3.Task Execution and Time Control
4. Quality Control and Result Aggregation

45 tasks have been published

平台首页 / 任务汇

任务类型: **不限** 数据预处理 遥感影像解译 专题图制作 大规模数据获取 地图矢量化 大数据计算 商业数据采购 其他

任务状态: **不限** 已发布 报名截止 任务分配 数据处理 任务完成

任务报酬: **不限** 100元以下 100-500元 500-1000元 1000-5000元 5000-1万元 1万元以上

发布时间: **不限** 最近一周 最近一月 最近半年 最近一年 至

10 1 共 5 页 显示第 1 到 10 条, 共 45 条记录, 已选择 0 条

编号	名称	任务类型	时间信息	报酬 (元)	操作
T2017030001 正在报名	基于Lansat影像的平谷县植被覆盖度指数专题图制作	专题图制作	报名开始:2017-03-10 报名结束:2017-03-15	1000.0	正在报名 已报14人 <input type="button" value="我要报名"/>
T2017020001 认领结束	基于遥感影像的杭州市萧山区部分区域土地利用类型解译	遥感影像解译	报名开始:2017-02-21 报名结束:2017-03-01	4000.0	报名结束 已报19人
T2017010001 认领结束	基于Landsat影像的泰和县土地利用类型解译	遥感影像解译	报名开始:2017-01-04 报名结束:2017-01-09	4000.0	报名结束 已报11人
T2016120006 认领结束	基于Landsat影像的贺兰山等区域土地利用类型解译	遥感影像解译	报名开始:2016-12-20 报名结束:2016-12-27	6500.0	报名结束 已报10人
T2016120005 认领结束	基于Landsat影像的乌审旗等区域土地利用类型解译	遥感影像解译	报名开始:2016-12-20 报名结束:2016-12-27	6500.0	报名结束 已报12人
T2016120004 认领结束	基于Landsat影像的定边县等区域土地利用类型解译	遥感影像解译	报名开始:2016-12-20 报名结束:2016-12-27	4000.0	报名结束 已报10人
T2016120003 认领结束	基于Landsat影像的平罗县等区域土地利用类型解译	遥感影像解译	报名开始:2016-12-08 报名结束:2016-12-18	7500.0	报名结束 已报9人
T2016120002 认领结束	基于Landsat影像的鄂托克旗土地利用类型解译	遥感影像解译	报名开始:2016-12-08 报名结束:2016-12-18	5000.0	报名结束 已报12人
T2016120001 认领结束	基于Landsat影像的杭锦旗土地利用类型解译	遥感影像解译	报名开始:2016-12-08 报名结束:2016-12-18	5000.0	报名结束 已报12人
T2016080002 任务结束	基于Landsat影像的山东省平度市土地利用类型解译	遥感影像解译	报名开始:2016-08-19 报名结束:2016-08-28	6000.0	报名结束 已报22人

10 1 共 5 页 显示第 1 到 10 条, 共 45 条记录, 已选择 0 条

编号	基本信息	毕业学校/工作单位	数据处理经验	解决方案摘要	状态(得分)	备注	操作
1	潘平(女) 硕士 157-3651		处理过landsat数据进行植被覆盖度以及相关数据的计算; 进行过野外数据采集及处理	利用arcgis软件,对相关数据进行计算,得出最终需求(详细方案)	未处理(0分)		报名表 删除报名 审核报名
2	马(男) 硕士 182-8530			基于NDVI利用像元二分模型反演植被覆盖度(详细方案)	未处理(0分)		报名表 成果示例 删除报名 审核报名
3	徐(男) 本科 135-178	鲁东大学(研究生) 河北工业大学城市学院(本科)		收集2007、2012和2016年三期Landsat影像,进行大气校正和几何校正,消除辐射误差和几何畸变;计算NDVI值,根据NDVI计算植被覆盖度指数,利用ArcGIS出图(详细方案)	未处理(0分)		报名表 成果示例 删除报名 审核报名
4	屈(男) 硕士 152-911	兰州大学(研究生) 西北师范大学(本科)	参与并负责《全国水土流失动态监测与公告项目》中多个区域的土地利用解译、植被覆盖度计算、土壤侵蚀计算项目,有良好的数据处理能力和项目组织协调能力。如,黄河流域内孤山川流域、皇甫川流域、秃尾河流域、石马川流域等六条流域内基于航空遥感影像的土地利用类型解译、植被覆盖度计算、土壤侵蚀强度计算;内蒙古准格尔旗、新疆尉犁县基于资源三号、高分一号遥感影像的地利用类型解译、植被覆盖度计算、土壤侵蚀强度计算等。影像预处理包括几何校正、大气校正、影像拼接、裁剪、地理配准等。	1、影像准备:下载2007年,2012年和2016年3个时相北京市平谷区的植被覆盖度; 2、数据预处理:根据下载影像,做必要的数据预处理,如几何校正、大气校正、缓冲区分析、裁剪等; 3、植被覆盖度计算:确定植被覆盖度遥感估算模型,分别计算植被指数和植被覆盖度; 4、制作植被覆盖度专题图; 5、撰写报告:针对植被覆盖度计算过程和结果,撰写技术报告。(详细方案)	未处理(0分)	报名表 成果示例 删除报名 审核报名	

Task progress and results uploaded by task executors



任务处理结果列表

文档标题	文档类型	上传时间	审核文件	备注	操作
数据处理说明书	结果报告	2017-01-26 19:44:04			<input type="button" value="下载"/> <input type="button" value="删除"/> <input type="button" value="审核"/>

5 1 共 1 页 显示第 1 到 1 条, 共 1 条记录, 已选择 0 条

registered users: 150,000+

Supporting published papers: 1783

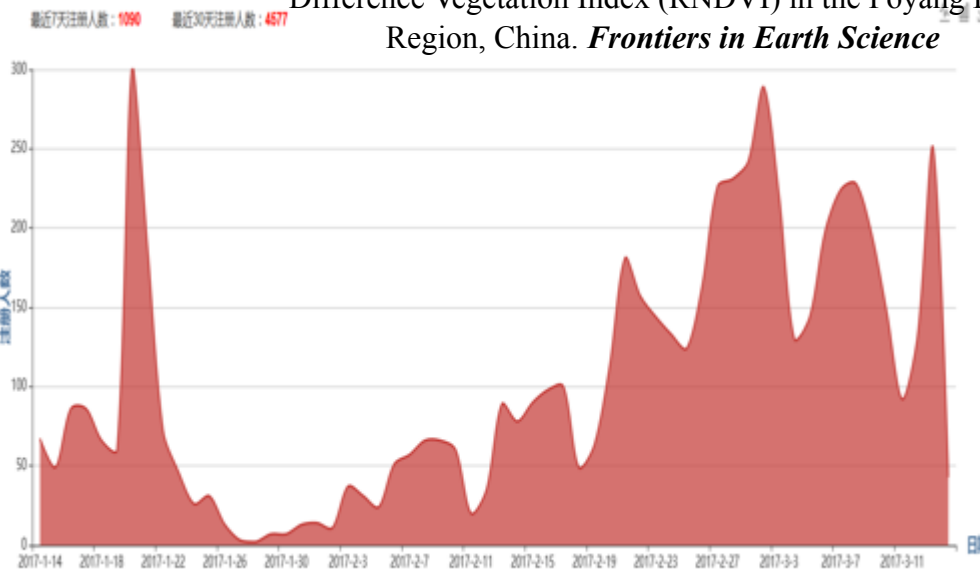
terrain-corrected (L1T) products. They have undergone systematic radiometric and geometric corrections and overall geometric fidelity with ground control points and a digital elevation model ensured (NASA Goddard Space Flight Center, 2011). In 2003, the Scan Line Corrector (SLC) mechanism on-board Landsat / permanently failed. Therefore, for the ETM+ SLC-off images used for rice cropping systems classification, multi-image adaptive local regression proposed by the International Scientific Data Service Platform. Computer Network Information Center, CAS (<http://datamirror.csdb.cn>) was applied to fill in gaps to improve data usability. This method generally used 1-2

注册用户管理 / 用户列表

20 1 共 7656 页 显示第 1 到 20 条, 共 153118 条记录, 已选择 0 条

编号	邮箱	用户名	单位	学历	注册时间	上次登录
1	971612396@qq.com	l86412068		本科	2017-03-14 15:20	2017-03-14 15:22
2	zhaibing2007@163.com	uiui117		硕士	2016-06-15 09:37	2017-03-14 15:21
3	398875249@qq.com	康晓娟	上海同济城市规划设计研究院	硕士	2015-07-23 12:00	2017-03-14 15:20
4	123829335@qq.com	半久		博士及以上	2017-03-14 14:42	2017-03-14 15:20
5	1129971910@qq.com	张志贇	福州大学	硕士	2015-04-17 17:12	2017-03-14 15:19
6	4F972B5A31D6B7521514E9651C168BB4	娟子柚子			2017-03-14 15:19	2017-03-14 15:19
7	387715609@qq.com	龙泽鋈	哈尔滨师范大学	本科	2016-02-04 13:39	2017-03-14 15:17
8	910950311@qq.com	星汉 灿烂		本科	2016-03-31 09:44	2017-03-14 15:17
9	3149903302@qq.com	惜	池州学院	本科	2017-03-12 22:16	2017-03-14 15:17
10	790906399@qq.com	颗粒状		本科	2017-02-21 15:02	2017-03-14 15:16
11	755492936@qq.com	expo	辽宁工程技术大学	硕士	2016-09-12 20:45	2017-03-14 15:16
12	1525188076@qq.com	鱼儿		本科	2017-03-14 15:12	2017-03-14 15:14
13	793747426@QQ.COM	啾啾\<进螺		本科	2016-09-05 21:48	2017-03-14 15:13
14	253852778@qq.com	禹吉	郑州大学	硕士	2015-11-11 19:20	2017-03-14 15:12
15	jiang08@163.com	梁军	天津华北地质勘查总院	硕士	2015-03-02 15:59	2017-03-14 15:11
16	41359925@qq.com	休伯利安	生态所	硕士	2011-07-31 14:06	2017-03-14 15:11

Peng Li et al., 2015, Mapping rice cropping systems using Landsat-derived Renormalized Index of Normalized Difference Vegetation Index (RNDVI) in the Poyang Lake Region, China. *Frontiers in Earth Science*



Typical CASE: Research Data publication



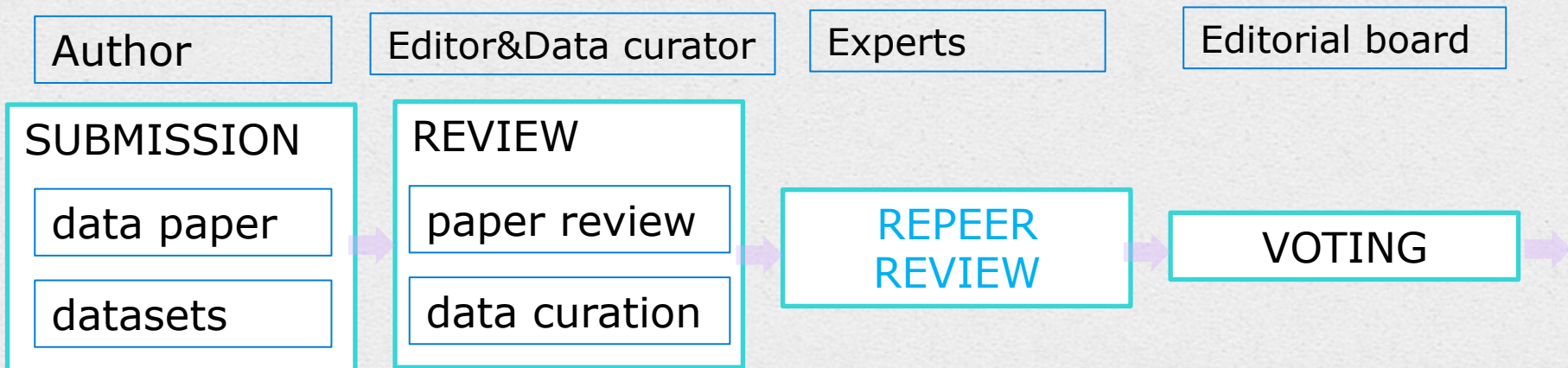
- Academic journal publishing **multidisciplinary data papers**
- **Quarterly** with **online** bilingual versions
- Open access, efficient processing, high exposure and rapid dissemination
- CN11-6035/N, ISSN 2096-2223



- SCOPE: Data papers describing (but not limited to) :
 - Datasets or data products generated from major scientific activities
 - Derived datasets or data products refined from raw data
 - Datasets linked to existing publications
- China Scientific Data does NOT publish new research findings, or techniques, methods and cases concerning data quality researches and data applications.

- Enhance data Quality
- Credit for data producers
- Incentive for open data

Publishing Process



OTHER
PROCESS
BEFORE
PUBLISHING

DATA PAPERS
DOI/Metrics

DATASETS
DOI/Metadata

Editorial office




- TITLE
- ABSTRACT
- KEYWORDS
- **Database/Dataset Profile**
- **Introduction/Overview**
- **Data collection and processing**
- **Sample description**
- **Quality control and assessment**
- Values and significance (optional)
- Usage notes (optional)
- Acknowledgments (optional)
- Authors and contributions
- References
- Data citation



Dataset Profile			
Chinese title	青藏高原 MODIS 逐日无云积雪面积数据集 (2002 – 2015 年)		
English title	MODIS daily cloud-free snow cover products over Tibetan Plateau (2002 – 2015)		
Corresponding author	Qiu Yubao (qiuyb@radi.ac.cn)		
Data author(s)	Qiu Yubao, Guo Huadong, Chu Duo, Zhang Huan, Shi Jiancheng, Shi Lijuan, Zheng Zhaojun, Laba Zhuoma		
Time range	July 2002 – April 2015		
Geographical scope	The study spans an area of 25° – 45°N and 67° – 107°E, including the entire Tibet Autonomous Region and Qinghai Province, parts of Sichuan, Yunnan, Xinjiang and Gansu provinces, as well as parts of foreign territories in southern and western Tibetan Plateau		
Spatial resolution	500-meter	Data volume	6.9 GB
Data format	Geotiff		
Data service system	< http://www.sciencedb.cn/dataSet/handle/55 >		
Source(s) of funding	Special Fund for Meteorological Scientific Research in the Public Interest “Constructing a Remote Sensing Product Dataset for Snow Pack over Tibetan Plateau” (No. GYHY201206040); State Key Program of National Natural Science Foundation of China (ABCC Grant No. 41120114001); the National Natural Foundation of China (No. 41371351); and “One-Three-Five” Planning Projects of the Chinese Academy of Sciences.		
Dataset composition	The dataset consists of two parts: the “MODIS daily cloud-free snow cover products over Tibetan Plateau from 2002 to 2015”, and the vector data of the research area. They are, respectively: (1) MODIS_Dysno_Cloud-free_2002-2015.zip (6.9 GB), i.e., the daily cloud-free snow data; (2) Tibet_Range.zip (24 KB), i.e., the auxiliary vector data storing the boundary of the study area over Tibetan Plateau.		

Zone II - Geoscience Advanced search




A dataset of No.5 large-scale integrated observation buoy on the Yangtze estuary (2014 – 2015) 2017年第1期
[Data Paper] January 25, 2017
浏览次数: 428

作者: Zhang Bin, Feng Lijiang, Wang Yanjun et al.
单位: Institute of Oceanology, Chinese Academy of Sciences, Qingdao, Shandong 266071, China
关键词: Yangtze estuary; observation buoy; observation data; meteorologic; hydrology; water quality
doi: 10.11922/csdata.170.2015.0037

摘要: For the purpose of monitoring the marine environment of high productivity area on the Yangtze estuary, No.5 large-scale integrated ocean observation buoy was laid at the Eastern of Chengshan Island (3 ...

[查看详情](#)




Soil type database of Tibet: a region-wide soil dataset based on the Second National Soil Survey 2017年第1期
[Data Paper] January 20, 2017
浏览次数: 709

作者: Gao Meirong, Shi Jianping, Pan Kai et al.
单位: Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Key Laboratory of Mountain
关键词: Tibet soil; soil type database; second national soil survey
doi: 10.11922/csdata.170.2016.0010

摘要: The data from Soil types database of Tibet is based on the Second National Soil Survey, China Science Publishing in 1994, is one of the summary of Tibet land resources survey. The primary coverage of ...

[查看详情](#)




CTD observation dataset of scientific investigation over the South China Sea (2009 – 2012) 2016年第3期 P 13-18
[Data Paper] December 31, 2016
浏览次数: 647

作者: Ju Chao, Li Sha, Chen Rongyu et al.
单位: South China Sea Institute of Oceanology, Chinese Academy of Sciences
关键词: CTD; temperature; salinity; the South China Sea; seabed
doi: 10.11922/csdata.170.2016.0013

摘要: In situ water depth, temperature, and salinity are basic items of sea observation. Conductivity-temperature-depth (CTD) is the self-recording instrument which is used to measure the depth, the vertical ...

[查看详情](#)




A dataset of land cover for Ganzhou region (1988 – 2014) 2016年第3期 P 1-7
[Data Paper] November 23, 2016
浏览次数: 1133

作者: Qiu Yubao, Ruan Yongqian, Li Hengfei et al.
单位: Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Acad
关键词: land cover; Ganzhou region; environment; landsat
doi: 10.11922/csdata.170.2015.0012

摘要: Land cover is a combination of different kinds of cover in the earth's surface, which is an important parameter to understand and obtain the regional natural resources and environmental changes. Acces ...

[查看详情](#)




Dataset of the soil sample archives of China (1989–2013) 2016年第3期 P 8-12
[Data Paper] October 10, 2016
浏览次数: 1201

作者: Pan Kai, Song Ge, Shi Jianping et al.
单位: Institute of Soil Science, Chinese Academy of Sciences,
关键词: dataset; soil sample; ecological change; data integration
doi: 10.11922/csdata.170.2015.0015

摘要: Soil samples record the history of soil evolution, and are extremely valuable to scientific research. The Chinese Ecosystem Research Network

[查看详情](#)

Zone II - Geoscience Advanced search




Water physical and chemical data of Taihu laboratory (2001 – 2006) 2016年第3期 P 13-20
[Data Paper] August 15, 2016
浏览次数: 1648

作者: Jinbo Xu, Ronghua Ma, Zhen Wang et al.
单位: Nanjing Institute of Geography & Limnology, Chinese Academy of Sciences
关键词: Taihu Lake; water quality; water chemistry; observation data
doi: 10.11922/csdata.170.2015.0014

摘要: This dataset is provided by TAILER (Taihu Laboratory for Lake Ecosystem Research) of Nanjing Institute of Geography & Limnology, Chinese Academy of Sciences. It contains 8 routine monitoring sites da ...

[查看详情](#)




Soil Type Database of China—the National Soil Dataset based on the Second National Soil Survey 2016年第3期 P 1-12
[Data Paper] August 5, 2016
浏览次数: 2317

作者: Shi Jianping, Song Ge,
单位: Institute of Soil Science, Chinese Academy of Sciences, Nanjing
关键词: soil; soil type database; second national soil survey; data integration
doi: 10.11922/csdata.170.2015.0033

摘要: Soil types database of China based on second national soil survey summary is the most comprehensive soil data resource over national scale. It covers the major kinds of soil type over nationwide in the ...

[查看详情](#)

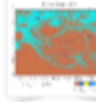


Atmospheric Physics Data of Field Observation Stations on the Tibetan Plateau 2016年第1期 P 12-22
[Data Paper] June 1, 2016
浏览次数: 1884

作者: Guo Xuejun, Wang Yongde, Zhang Guoshui et al.
单位: Institute of Tibetan Plateau Research, Chinese Academy of Sciences
关键词: Tibetan Plateau; atmospheric physics; boundary layer meteorology; field stations; observation data; data sharing
doi: 10.11922/csdata.170.2015.0003

摘要: The five field observation stations: Nam Co Station for Multisphere Observation and Research, Gornojiangna Atmospheric and Environmental Observation and Research Station, Southeast Tibet Station for Ap ...

[查看详情](#)



MODIS daily cloud-free snow cover products over Tibetan Plateau (2002 – 2015) 2016年第1期 P 1-11
[Data Paper] May 24, 2016
浏览次数: 3018

作者: Qiu Yubao, Guo Huadong, Chu Duo et al.
单位: Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Acad
关键词: Tibet Plateau; MODIS; MODIS daily snow product
doi: 10.11922/csdata.170.2015.0003

摘要: Snow cover over Tibetan Plateau plays an important role in regional water and energy circulation. Snow ablation also affects downstream rivers. Snow parameters and their long-term changes are sensit ...

[查看详情](#)



A dataset of suspended solids concentration inversion for Poyang Lake, China 2000–2013 2016年第1期 P 23-30
[Data Paper] March 31, 2016
浏览次数: 2183

作者: Wang Junjie, Chen Eiyang, Zhu Junxiang et al.
单位: -
关键词: suspended solids concentration; inversion model; Poyang Lake; environmental management; MODIS
doi: 10.11922/csdata.170.2015.0001

摘要: Suspended solids concentration is an important parameter to evaluate water quality and water environment. To obtain data on the spatio-temporal distribution of suspended solids concentration via remot ...

[查看详情](#)



中国科学院

CHINESE ACADEMY OF SCIENCES



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



Big Data for
Better Science



+ **BIG**
DATA =

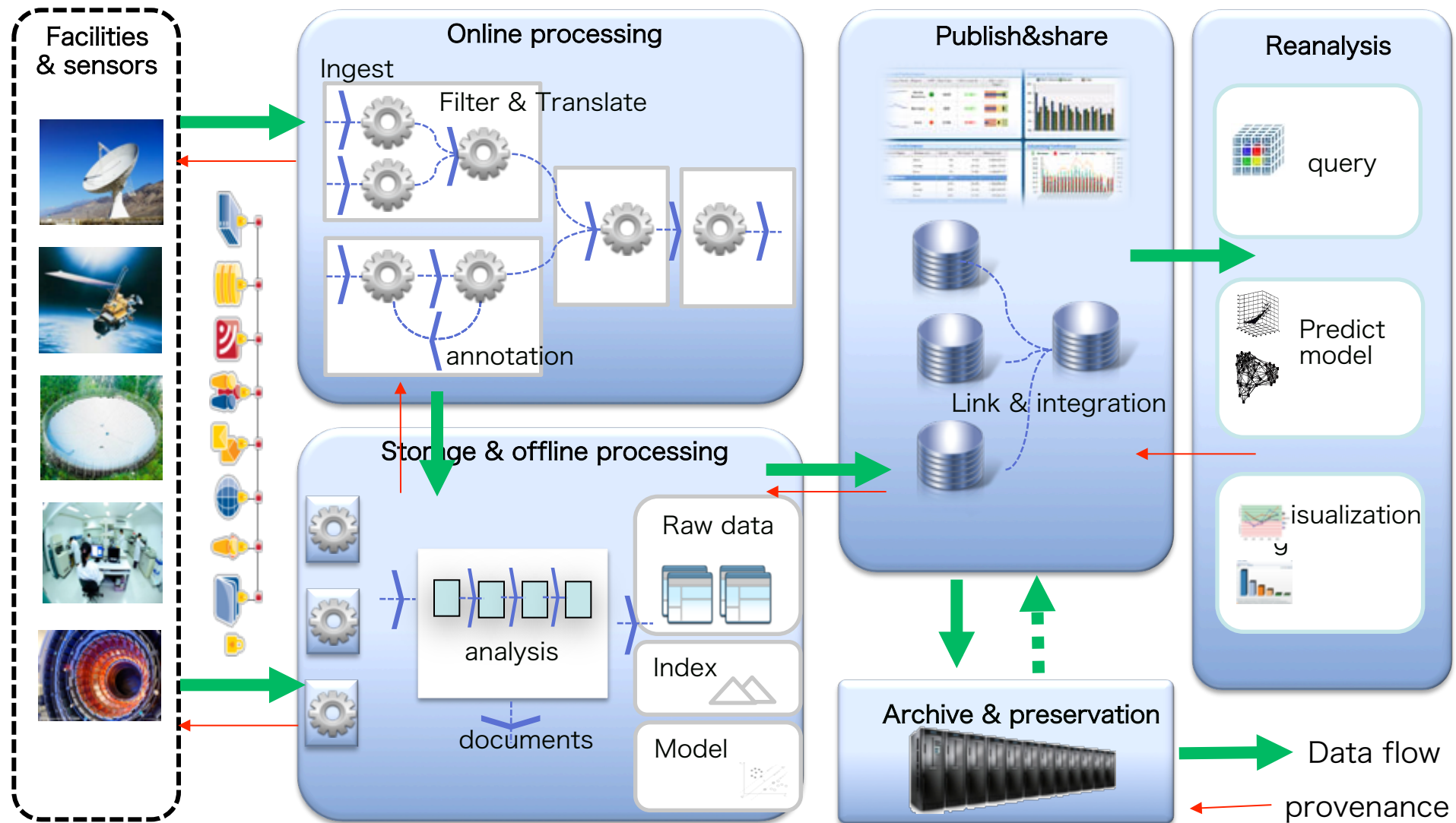




- Make data flowing from generator to applications (**Data Flow**)
 - without barriers, automatically or semi-automatically, on-demand
- Make huge multi-source data linked together as a data network (**knowledge graph**)
 - Based on semantics, automatically or semi-automatically find invisible relationship
 - Huge graph data management
- Semi-automatically find correlation from multi-source, high dimension data based on feature (**Data explore and data Intelligent**)
 - Feature selection, machine learning model selection , training and optimization
- **Make research data infrastructure scalable , evolution and invisible**



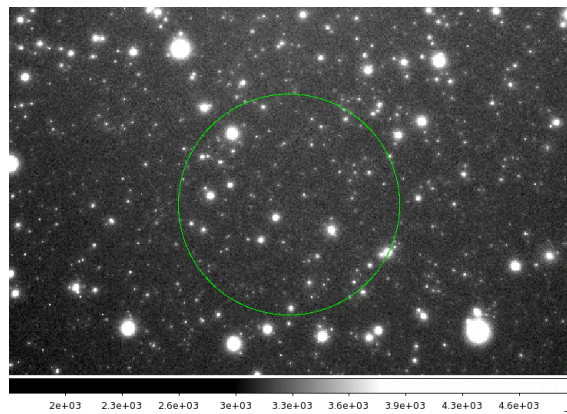
Data Flow Life Cycle



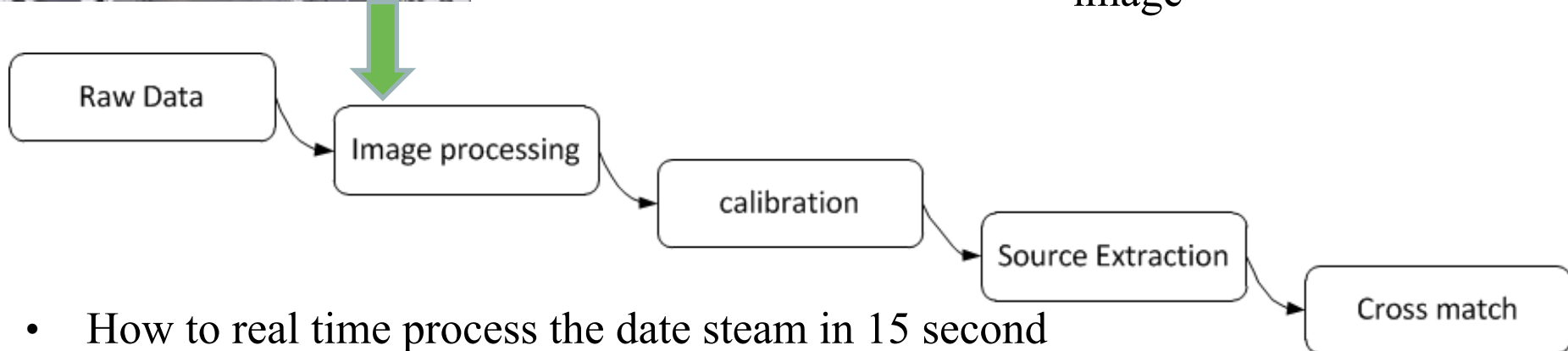


A typical use Case: GWAC

The Ground-based Wide-Angle Camera array, part of SVOM



- 40 telescopes
- 15 second : $40 * 32\text{MB}$ Image
- 1.7 Million object per image



- How to real time process the data stream in 15 second
- How to storage and query 100-1000 billion objects



One Size Does Not Fit All

Take away One Size Does Not Fit All

- Column store (stupid analytics)
- Array store (smart analytics)
- Streaming (velocity solution)
- New SQL (other velocity solution)
- No SQL (low end; semi-structured data)
- Legacy stuff (in place now - but obsolete)
- One or more curation systems (800 pound gorilla)

- Use the right tool for the job!!!!



Stonebraker, 2015年

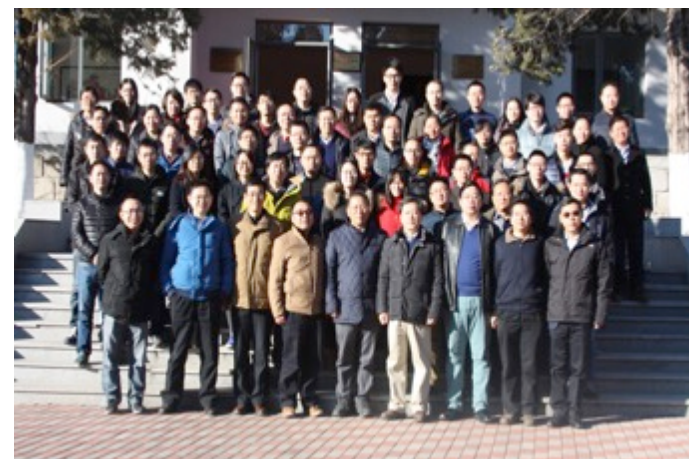
one surface fits all? !



Scientific Big Data Management system(PiStack)

one surface fits all

- Accelerate Data Flowing and data link
- A whole life cycle management
- Open architecture, Open source, and dynamic evolution
- Funded by National Key Research and Development Program(2017.7-2019.8)



Collaborators

Architecture



Relational data
and Astronomy
domain

PiStack

Graphic Data
and life science
domain



non structure
data and High
Physicals
domain

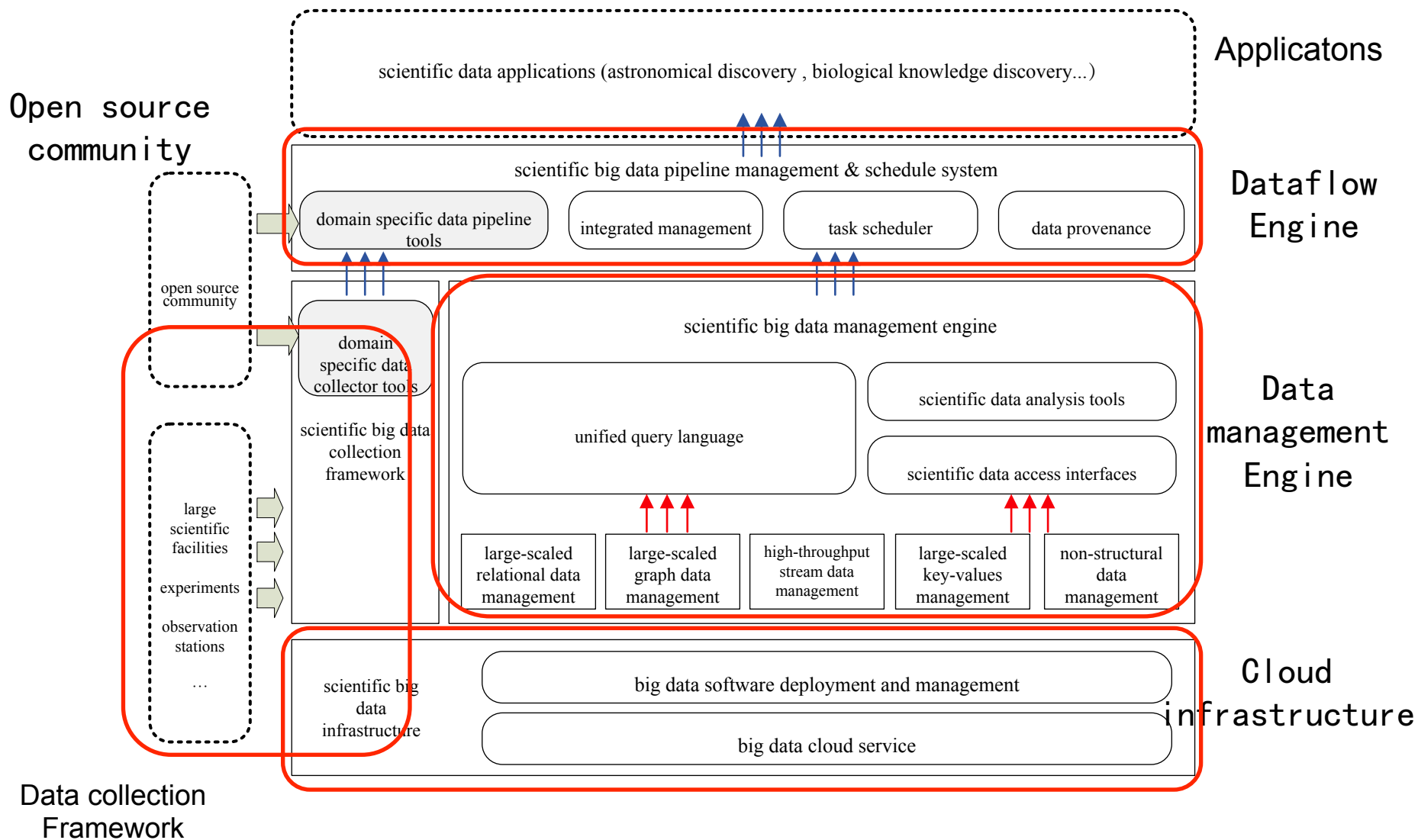


System
Development
and deployment





Architecture of PiStack





Progress and outlook

- 3 specific data management model
 - AstroDB(base on relational model)
 - GraphDB(based on graph model)
 - EventDB(based on no-SQL model)
- Simba: A Unified data query and access model
- PackOne: A Cloud resource scheduling and deployment toolkits
- A big data analysis framework for discovery
 - Methodology + specification + implementation + domain specific application
- CAS's Scientific big data program start, it will push PiStack development and deployment.



Conclusion

- Science discovery has increasingly become data intensive, and it calls for reliable and easily accessible scientific data infrastructure
- CAS is always promoting to build scientific data infrastructure to driven better science, especially in big data era
- Seeking potential cooperation with all international Colleagues in big data for science



中国科学院
CHINESE ACADEMY OF SCIENCES



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

Thank you!