



第二届GECAM研讨会

面向“多信使、多波段、多设备”联合观测的 数据发布与服务探讨

许琦、胡晓彦、邹自明、佟继周
中国科学院国家空间科学中心
国家空间科学数据中心

Outline

- 一．国内外数据管理与共享新趋势
- 二．当前联合观测项目数据共享方式
- 三．面向“多信使、多波段、多设备”联合观测的数据发布与服务探讨

科学数据国内外数据管理与共享的趋势

国际 发展新的数据评价准则 与科研社区联合建设良性科学创新生态



Findable **A**ccessible **I**nteroperable **R**eusable

Position Statement:



- ❑ ENABLING FAIR DATA IN THE EARTH SPACE AND ENVIRONMENTAL SCIENCES
- ❑ PRINCIPLES TOWARDS A DATA CULTURE OF USE, SHARING, CURATION, AND ATTRIBUTION

可发现 (Findable)	F1 数据(元数据)被分配全球唯一且持久的标识符
	F2 使用丰富的元数据描述数据
	F3 元数据清楚地包含所描述的数据的标识符
	F4 数据(元数据)在搜索应用服务中注册或索引
可访问 (Accessible)	A1 数据(元数据)可以使用标准化通信协议,通过其标识符进行检索
	A1.1 该协议是开放的、免费的,并且可普遍实施
	A1.2 协议允许在必要时进行身份验证和授权过程
可互操作 (Interoperable)	A2 即使数据不再可用,也可以访问其元数据
	I1 数据(元数据)使用正式的、可访问的、可共享的和广泛适用的语言来进行表示。
	I2 数据(元数据)使用遵循 FAIR 原则的词汇表
可重用(Reusable)	I3 数据(元数据)包括对其他数据(元数据)的合法引用
	R1 数据(元数据)用多个准确且相关的属性进行充分描述
	R1.1 发布的数据(元数据)包含清晰且可访问的数据使用协议
	R1.2 数据(元数据)包含详细的出处信息
	R1.3 数据(元数据)符合领域相关标准

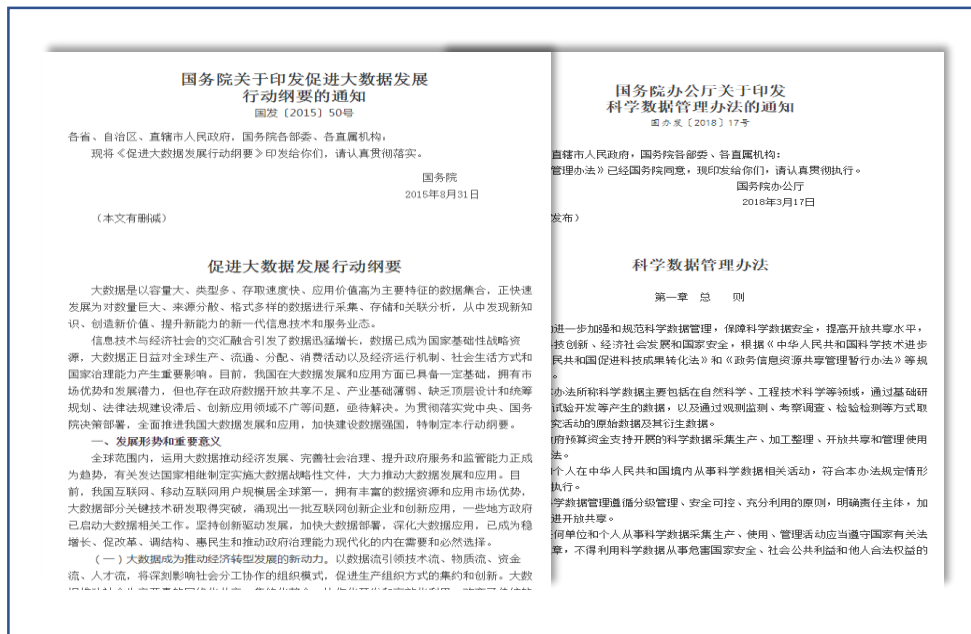


Enabling FAIR Data Project : <http://aims.fao.org/activity/blog/enabling-fair-data-across-earth-space-and-environmental-sciences-eses>

一、国内外数据管理与共享的趋势

我国

《大数据发展行动纲要》、《科学数据管理办法》
20个国家科学数据中心成立



序号	国家科学数据中心名称
1	国家高能物理科学数据中心
2	国家基因组科学数据中心
3	国家微生物科学数据中心
4	国家空间科学数据中心
5	国家天文科学数据中心
6	国家对地观测科学数据中心
7	国家极地科学数据中心
8	国家青藏高原科学数据中心
9	国家生态科学数据中心
...	

开放vs安全

新的挑战

智能化分析挖掘技术

跨领域数据发现与应用

可持续发展生态



二、当前联合观测项目数据共享方式

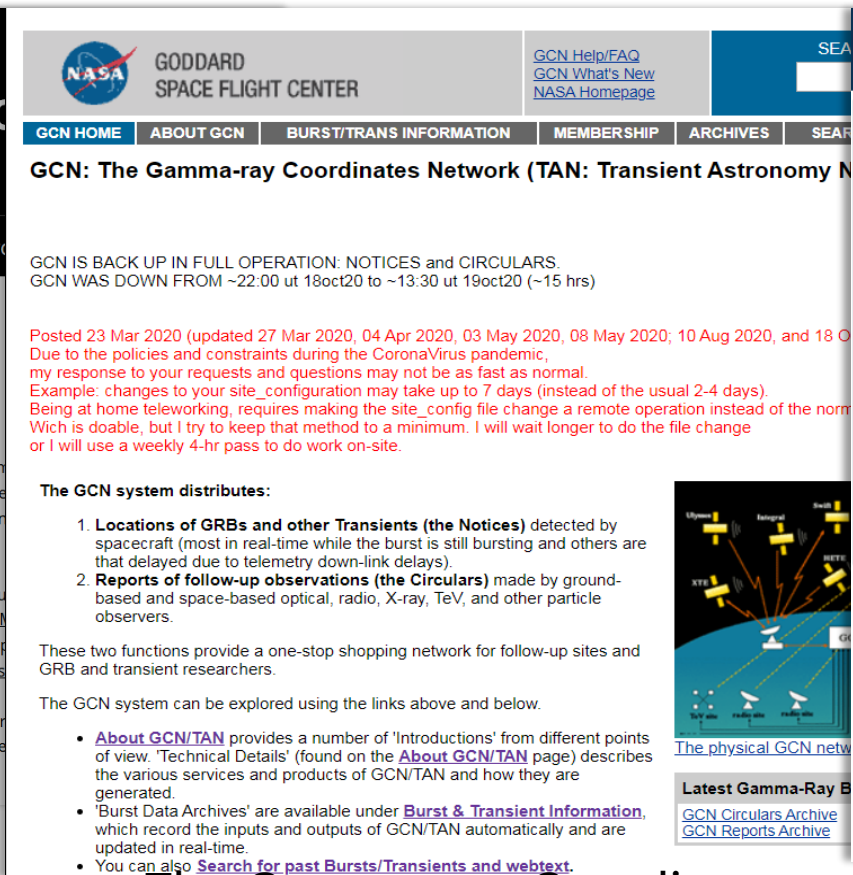
□ 从科学提案提出到科学观测计划的组织、联合观测等方面相对成熟



Event Horizon Telescope

网站截图

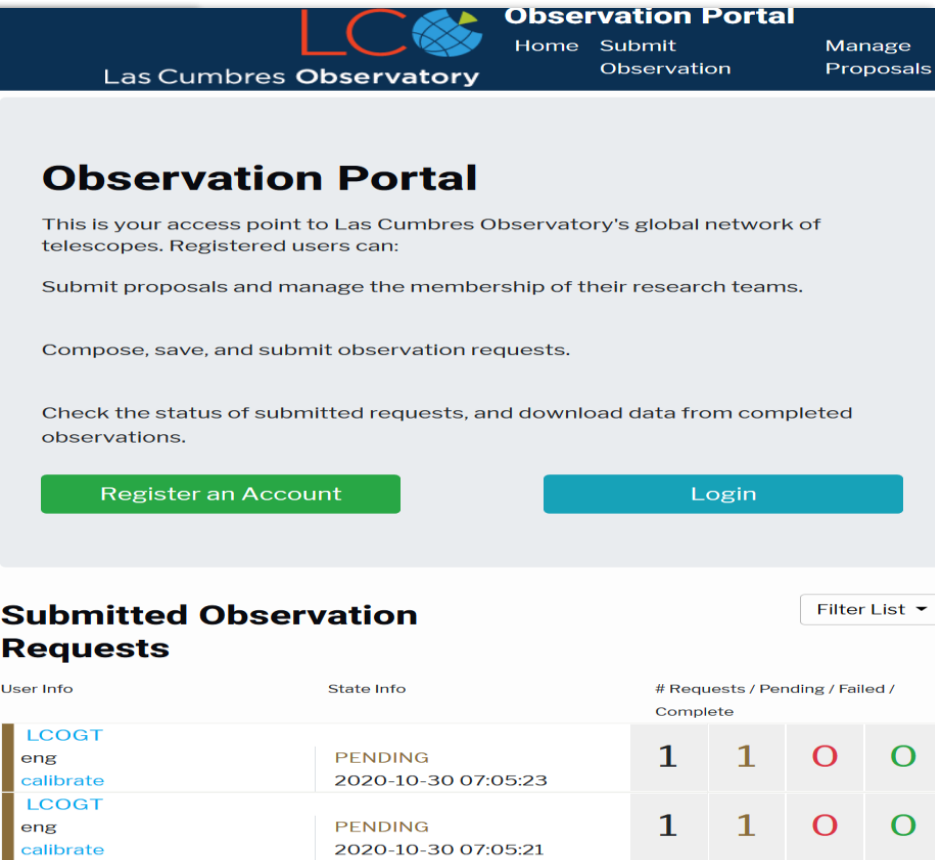
<https://eventhorizontelescope.org/>



The Gamma-ray Coordinates Network

网站截图

<https://gcn.gsfc.nasa.gov/>



Las Cumbres Observatory 's global network of telescopes

网站截图

<https://lco.global/>

二、当前联合观测项目数据共享方式

□数据共享方式相对传统，按照卫星任务型号、数据产品类型等分类查找方式进行数据发现和共享。

The screenshot displays the NASA LAMBDA Data Products website. The header includes the NASA logo, the text "National Aeronautics and Space Administration", "Goddard Space Flight Center", and "Sciences and Exploration". A search bar and social media links are also present. The main navigation bar includes "Home", "Data", "Simulations", "Tools", "Papers", "Education/Graphics", and "News". The "Data" section is expanded, showing "Data Hosted", "Experiment Tables", "Space-Based", "Suborbital", "Astrophysical", and "About Products". The "Data Hosted" section is selected, showing a list of experiments: ABS, ACT, BICEP2/Keck, CLASS, MIPOL, Planck, POLARBEAR, QUIET, SPIDER, SPT, and WMAP. A red box highlights this list, and a red arrow points to the "Overview" section. The "Overview" section is titled "ABS Data Products on LAMBDA" and includes a note: "If you use these data products please acknowledge Kusaka et al." Below this is a table titled "ABS 2018 Data Release".

Product Download Page	Description	No. Of Files/Size
Field A Maps	These are the coadded maps of Q and U for Field A shown in the bottom two panels of Figure 4 in Kusaka et al. The HEALPix FITS file contains the following fields in order: More...	1 file 3.6 MB
Field A Mask	The ABS Field A mask excludes sources with flux density greater than 1000 mJy and applies a taper at the field edge boundary. This mask is applied to the maps before taking the power spectra.	1 file 3.2 MB
Field A Spatial Window Function	The ABS Field A Spatial Window Function includes the mask and inverse noise weighting for the Field A map.	1 file 3.2 MB
Window Function	This is the Hankel transform of the convolution of the beam and the pointing error (or jitter), effectively representing the effect of smearing of the map due to these effects. The first column is ℓ and the second column is the Legendre coefficient in the expansion. The square of the second column is the window function. The window function is shown in Figure 2 of Kusaka et al.	1 file 18 KB

Legacy Archive for Microwave Background Data Analysis

数据共享页面截图

二、当前联合观测项目数据共享平台现状

发布在了
DATA
COMMONS
开放式的仓储
中心

The screenshot displays the EHT Data Commons interface. At the top, a navigation bar includes links for 'About', 'Data Store', 'Discovery Environment', and 'Atmosphere', along with a search bar and a 'Submit' button. The 'Data Commons' logo is highlighted with a red box and an arrow pointing to the text '发布在了 DATA COMMONS 开放式的仓储中心'. Below the navigation bar, the breadcrumb path is 'shared / commons_repo / curated / EHTC_MonitoringM87_Sep2020'. The main heading is 'EHT Data Products'. A table lists data products with columns: 'EHT Data Product Code', 'Title', 'Author(s)', 'Last Updated', 'DOI', and 'Reference'. The first entry is '2020-D03-01' for 'EHT Visibility Amplitude Data of M87* in 2011-2013' by 'The EHT Collaboration et al.', updated on '23 Sep 2020', with DOI 'https://doi.org/10.25739/daft-jz11' and reference 'Wielgus, M., et al. 2020, ApJ'. Below the table, the 'Publisher' is 'CyVerse Data Commons' and the 'Publication Year' is '2020'. The 'Rights' section states: 'This data is made available under the Public Domain Dedication and License v1.0 whose full text can be found at <http://www.opendatacommons.org/licenses/pddl/1.0/>'. This sentence is highlighted with a red box and an arrow pointing to the text '公共领域专用许可'. The 'Citation' section includes 'Export Formats: BibTeX Endnote'. Below this, there are buttons for 'show more' and 'Download Metadata'. A file listing table follows with columns: 'Name', 'Size', 'Created', and 'Last Modified'. The files listed are 'dat' (highlighted with a red box), 'EHTC_MonitoringM87_Sep2020.tgz', 'INVENTORY.txt', 'LICENSE.txt', and 'README.md' (highlighted with a red box). The 'Event Horizon Telescope 数据共享页面截图' text is overlaid on the bottom right of the file listing table.

EHT Data Product Code	Title	Author(s)	Last Updated	DOI	Reference
2020-D03-01	EHT Visibility Amplitude Data of M87* in 2011-2013	The EHT Collaboration et al.	23 Sep 2020	https://doi.org/10.25739/daft-jz11	Wielgus, M., et al. 2020, ApJ

Publisher: CyVerse Data Commons
Publication Year: 2020
Rights: This data is made available under the Public Domain Dedication and License v1.0 whose full text can be found at <http://www.opendatacommons.org/licenses/pddl/1.0/>
Citation: Export Formats: [BibTeX](#) [Endnote](#)

Name	Size	Created	Last Modified
dat	--	Sep 22, 2020 6:32:55 AM	Sep 22, 2020 6:32:55 AM
EHTC_MonitoringM87_Sep2020.tgz	24.0 kB	Sep 22, 2020 6:32:58 AM	Sep 22, 2020 6:32:58 AM
INVENTORY.txt	149.0 bytes	Sep 22, 2020 6:33:00 AM	Sep 22, 2020 6:33:00 AM
LICENSE.txt	34.5 kB	Sep 22, 2020 6:33:01 AM	Sep 22, 2020 6:33:01 AM
README.md	3.3 kB	Sep 22, 2020 6:33:00 AM	Sep 22, 2020 6:33:00 AM

三、面向联合观测的数据发布与服务探讨

针对国内外数据共享发展新趋势，面向“**多信使、多波段、多设备**”联合观测数据发布与服务需求，进一步完善服务方式，建立新型的数据共享服务平台，以解决：

- ① 国内多科天文卫星数据资源互连互通；
- ② 数据资源共享方面兼顾领域专业性，同时拓展共享范围广度；
- ③ 数据溯源和应用效益追踪。



三、联合观测的数据发布与服务探讨

1. 国内多科天文卫星数据资源互通

- 基于平台实现覆盖各卫星数据和联合观测产生的数据资源的检索发现
- 基于统一的协议实现底层数据互访，用户一站式获取数据实体资源。

HXMT卫星数据发布系统

Obs Num	Pno	Pno Num	Source Name	Longitude	Latitude	Obs Time from	Obs Time Len	Obs Pattern	Data Status	Offset	Option
P0114764013	P0114764	Too	4U 1608-52	243.1789	-52.4229	2018/7/5 12:52:06	17334	Pointing	Not Publish		Login
P0114764011	P0114764	Too	4U 1608-52	243.1789	-52.4229	2018/7/3 17:54:39	17338	Pointing	Not Publish		Login
P0114764010	P0114764	Too	4U 1608-52	243.1789	-52.4229	2018/7/2 08:29:52	17339	Pointing	Not Publish		Login

GECAM 卫星数据发布系统

The Nell Gehrels Gecam Observat...

GECAM is a constellation composed of two identical tiny satellite as (~150 kg/piece). The payload of each satellite includes a gamma ray detector (GRD), a charged particle detector (CPD) and a payload processor (EBOX), and 25 GRD detectors are used to monitor all the sky except the earth's occlusion. The GECAM satellite has an orbit height of 600 km and an inclination angle of 29 degrees. The two satellites operate in the same orbital plane with opposite orbital phases, thus forming a complete coverage of the entire sky.

Product Introduction

L1 Space Satellite
This data is the C2 image data of Soho Wide coron...

L1 Space Satellite
This data is the C2 image data of Soho Wide coron...

L1 Space Satellite
This data is the C2 image data of Soho Wide coron...

Data parameter

- Data set number: 0
- file number: 0
- Burst event number: 3
- Trigger event number: 3
- Downloads: 15
- Views: 9

Quick Access

BURST SEARCH

DAILY SEARCH

TRIG SEARCH

L3 SEARCH

Publish Trends

Links

- TEST
- NSSCC
- PAAS

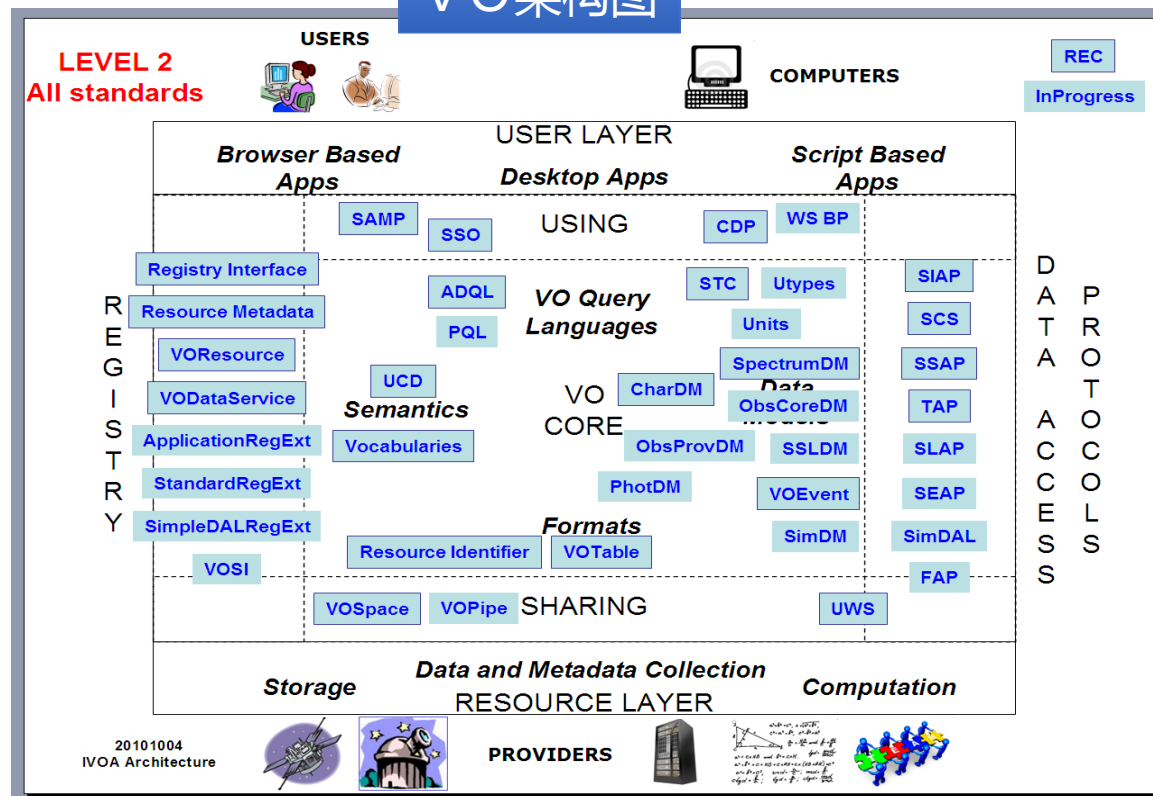
当前GECAM、SVOM、EP等卫星数据发布共享系统以满足任务需求为主，系统间相互独立。

三、联合观测的数据发布与服务探讨

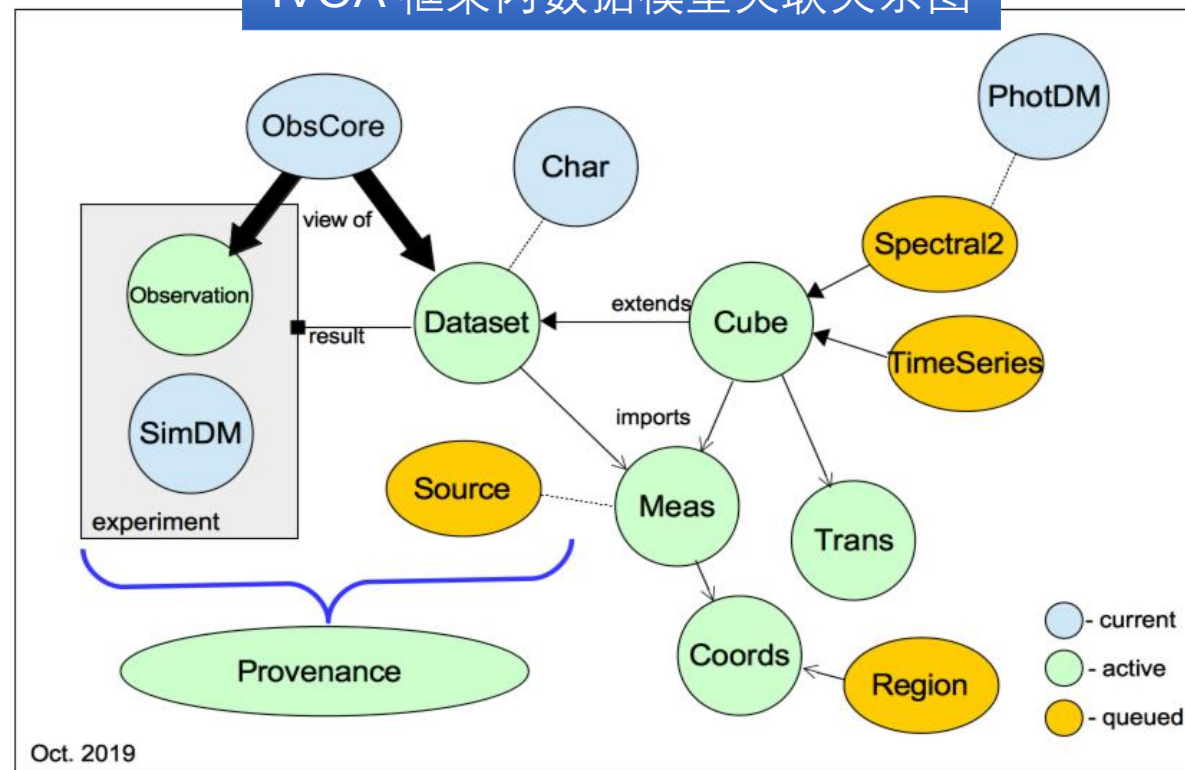
2. 数据资源共享方面兼顾领域专业性，同时拓展共享范围广度。

- 采用VO的元数据规范进行资源描述。

VO架构图



IVOA 框架内数据模型关联关系图

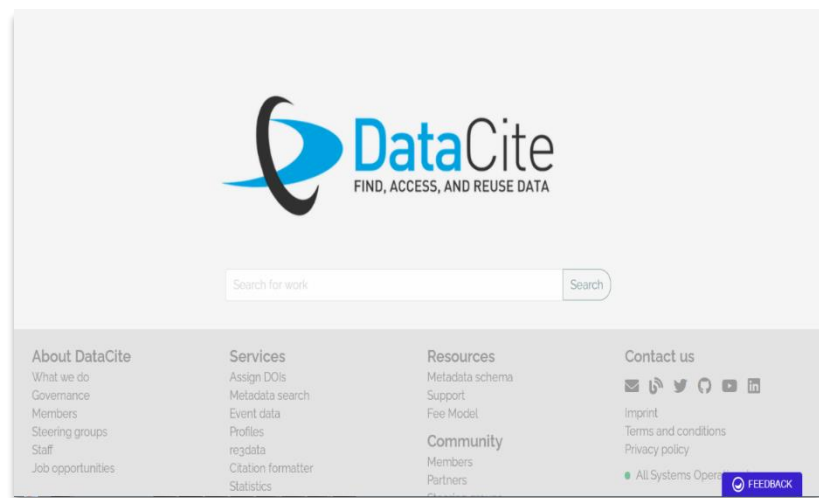


经过多年迭代发展，VO元数据规范覆盖了数据组织与描述、服务协议接口，框架体系成熟。

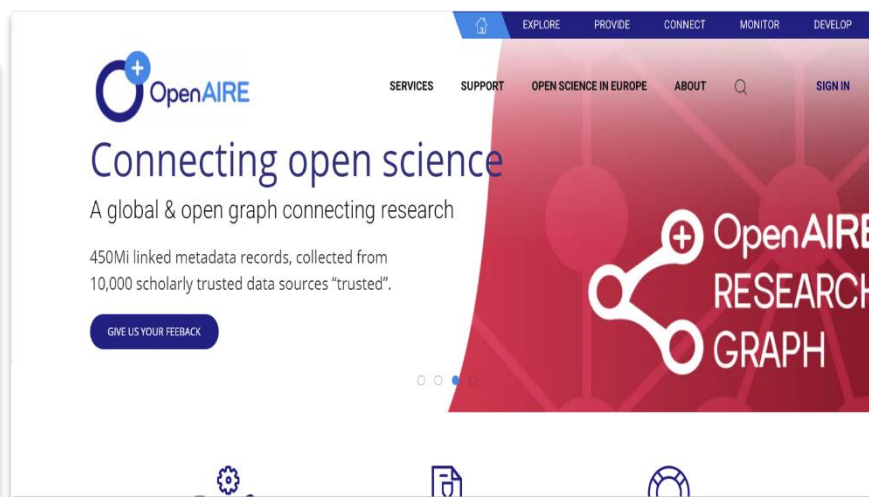
三、联合观测的数据发布与服务探讨

2. 数据资源共享方面兼顾领域专业性，同时拓展共享范围广度。

- 配置唯一标识符，在更广泛的范围内进行共享。
- 采用通用元数据规范，通过元数据收割等方式，拓展资源发现渠道。



DataCite 为研究数据 (research data) 提供 DOI。它收集每个有 DOI 的元数据，这些元数据用于大量研究数据索引，可直接查询和规范引用、获取统计数据并建立相关联系。所有元数据都可自由访问和查看。



Open Access Infrastructure Research for Europe 欧洲开放获取基础设施研究项目，提供研究成果记录的开放获取,以及由研究成果到作者、数据、软件、出版信息和资助信息等的可靠链接。



中国科技资源共享网，是我国跨部门、跨地区、跨领域的数据资源共享服务平台，平台拥有自然科技资源、科学数据、科学文献等海量信息资源，提供资源导航与检索、搜索引擎、绩效评估监测等服务。

三、联合观测的数据发布与服务设想探讨

A persistent identifier (PID) is a long-lasting reference to a digital resource.
数字资源包括：数字对象、数据集、数据论文、学术论文、研究人员等。



A well-known digital identifier for objects of any type.



[ISTIC & Wanfang Data](#)



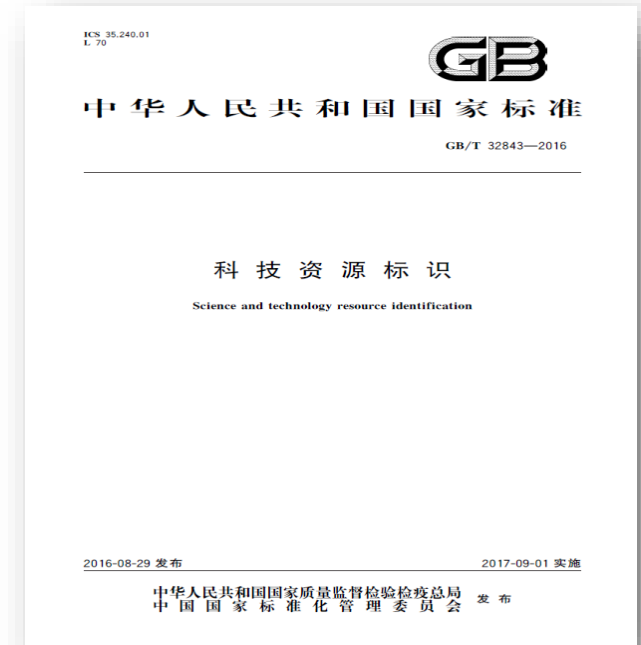
[DataCite](#)



A persistent digital identifier that distinguishes one researcher from other researchers and a record that supports automatic links among all his professional activities.

China Science & Technology Resource identifier (CSTR)

覆盖20余类科技资源类型

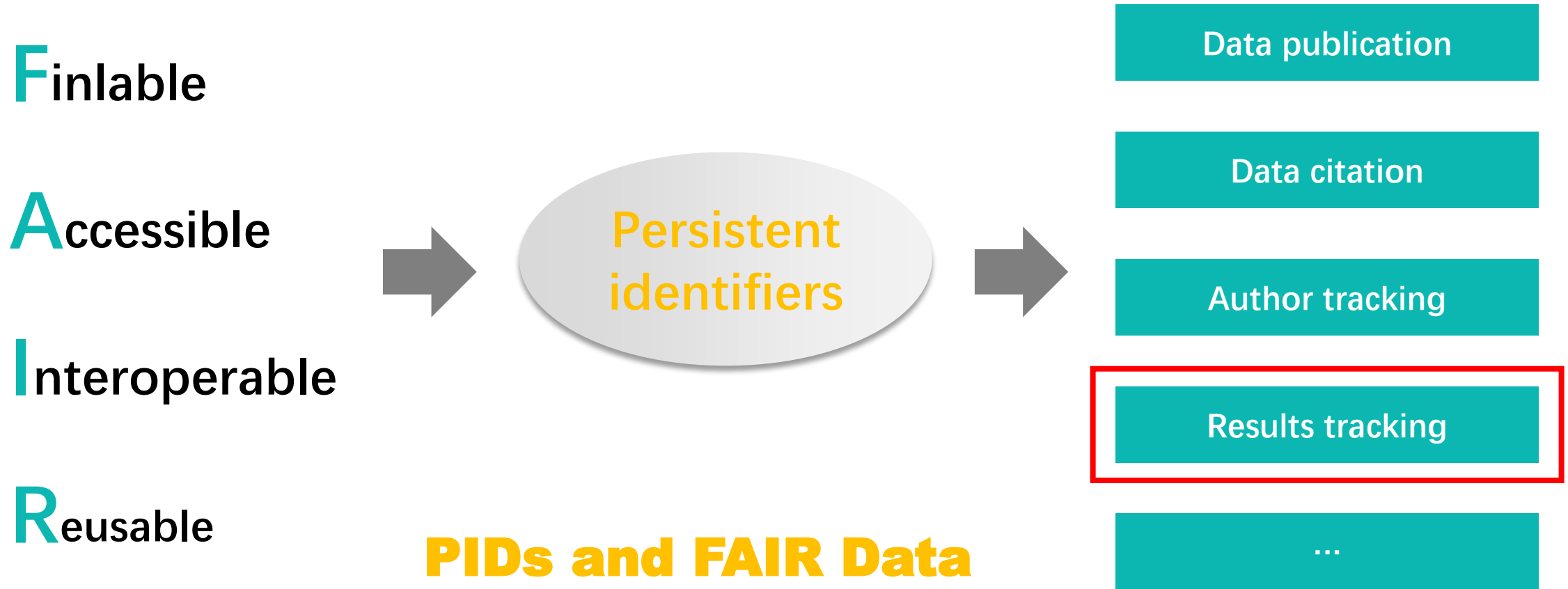


INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE

International standard serial number, an eight-digit number assigned to many serial publications such as newspapers, magazines, annuals, and series of books.

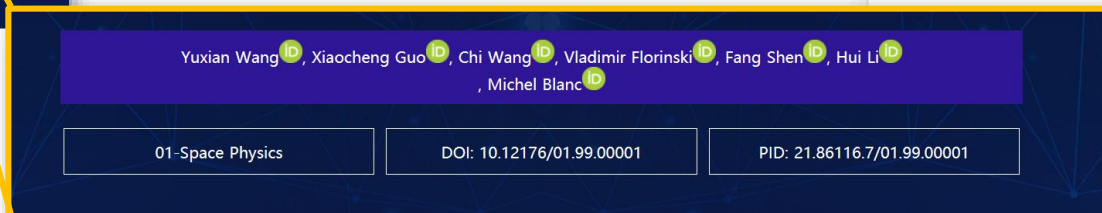
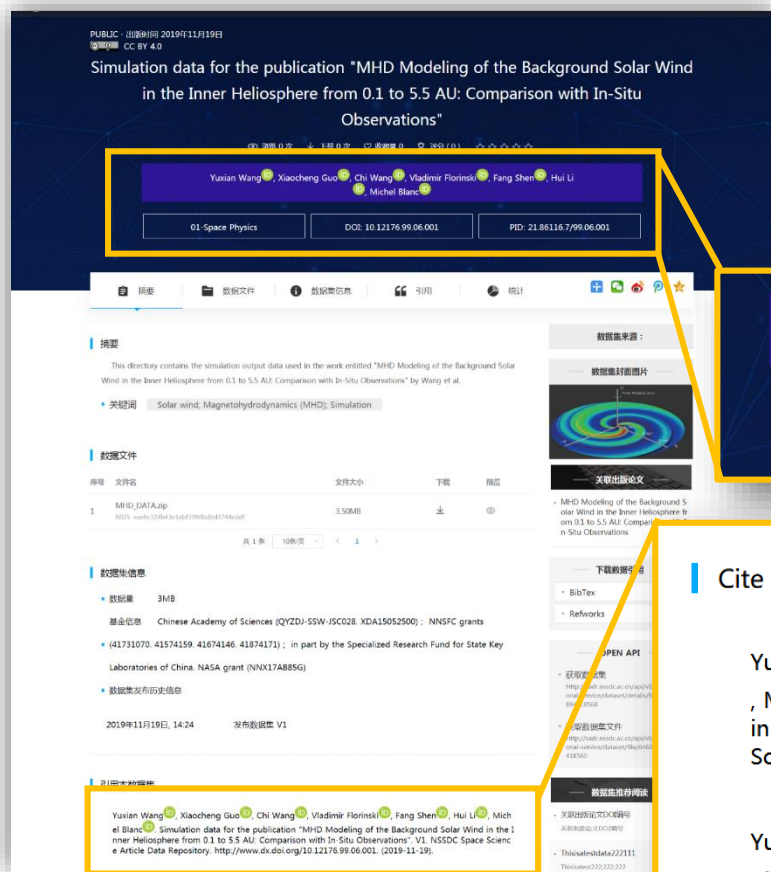
三、联合观测的数据发布与服务探讨

3.数据溯源和应用效益追踪。



三、联合观测的数据发布与服务探讨

3.数据溯源和应用效益追踪。



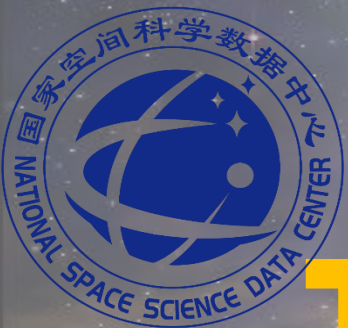
Cite this dataset

Yuxian Wang^{ID}, Xiaocheng Guo^{ID}, Chi Wang^{ID}, Vladimir Florinski^{ID}, Fang Shen^{ID}, Hui Li^{ID}, Michel Blanc^{ID}. Simulation data for the publication "MHD Modeling of the Background Solar Wind in the Inner Heliosphere from 0.1 to 5.5 AU: Comparison with In-Situ Observations". V1. NSSDC Space Science Article Data Repository. <http://hdl.pid21.cn/21.86116.7/01.99.00001>. (2020-02-21).

Yuxian Wang^{ID}, Xiaocheng Guo^{ID}, Chi Wang^{ID}, Vladimir Florinski^{ID}, Fang Shen^{ID}, Hui Li^{ID}, Michel Blanc^{ID}. Simulation data for the publication "MHD Modeling of the Background Solar Wind in the Inner Heliosphere from 0.1 to 5.5 AU: Comparison with In-Situ Observations". V1. NSSDC Space Science Article Data Repository. <https://dx.doi.org/10.12176/01.99.00001>. (2020-02-21).



国家空间科学数据中心的论文关联数据出版系统（SADR）截图



Thank you for listening !

xuqi@nssc.ac.cn

huxiaoyan@nssc.ac.cn

mzou@nssc.ac.cn

tongjz@nssc.ac.cn