



Archiving for Different Source Astronomical Data: Experiences and Lessons Learned

Youfen Wang , Yihan Tao
NADC Data Resource Curation Group
2024 Spring IVOA
May 23 2024, Australia, Sydney

Out Line



Experiences & Lessons

Five Type of Data

01 General Dataset

Data produced by Chinese astronomical telescopes and instruments

02 Paper Data

paper data accepted by astronomical journals

03 Scientific Project Data

Scientific project data supported by government funding

04 Mirror Data

Data from astronomical telescopes and instruments worldwide

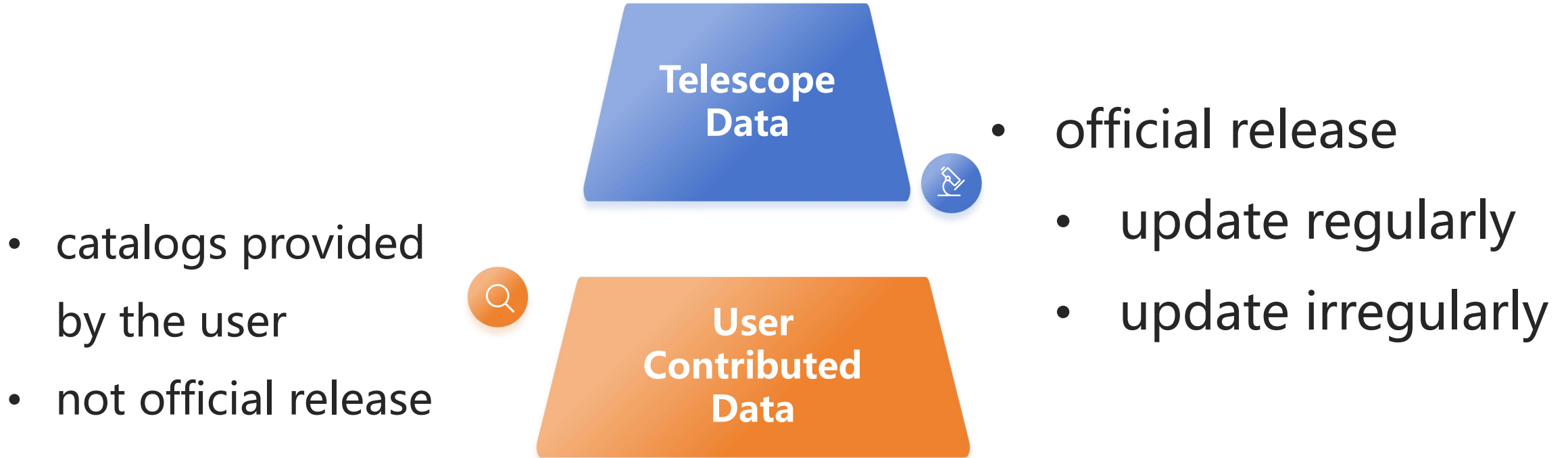
05 Outreach Data

Astronomical science and education data

06 Statistics

Statistics of the data

General Dataset



General Dataset-Telescope Data

01 LAMOST

Data save in NADC server each release with 10-20 metadata

Now have LAMOST DR1-DR5(Low resolution) , DR6-DR12(both low resolution and medium resolution)

There are Q1 Q2 Q3 V0 V1 V2 each release, V2 is for international access

Data saved in FAST server, programs and observed sources available

Update every three months

Now have DR1 to DR17 each release with one metadata

02 FAST

update regularly upto now

General Dataset-Telescope Data



Antarctic
AST3
Telescope



Beijing-
Arizona
Sky
Survey



The South
Galactic
Cap u-
band Sky
Survey



Lijiang
2.4m
Telescope



The First
Release
of the
CSTAR



Digitized
Data of
China
Astronom-
ical
Plates

More.....

update irregularly

General Dataset-User Contributed Data

M
subdwarf
catalog

1

LAMOST
Spectra
Classification
Dataset

2

Stellar
Parameters
for over 20
Million
Stars from
SAGES
DR1

3

SAGES
Data
Release 1
& 1s

4

FASHI
Data
Release 1

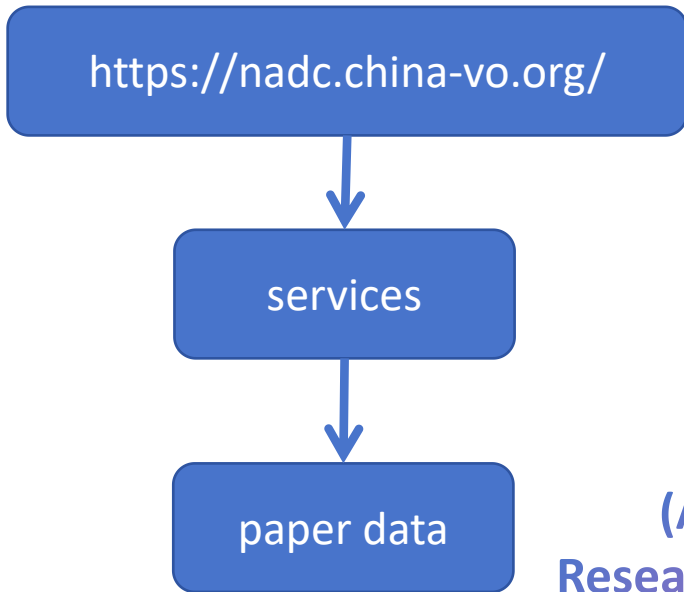
5

The first
comprehensive
Milky Way
stellar mock
catalogue for
the Chinese
Space Station
Telescope
Survey Camera

6



Paper Data



how access?

AAS Journals :
(AJ/ApJ/ApJL/ApJS/RNAAS) and
Research in Astronomy and Astrophysics :
RAA

Chart

Movie

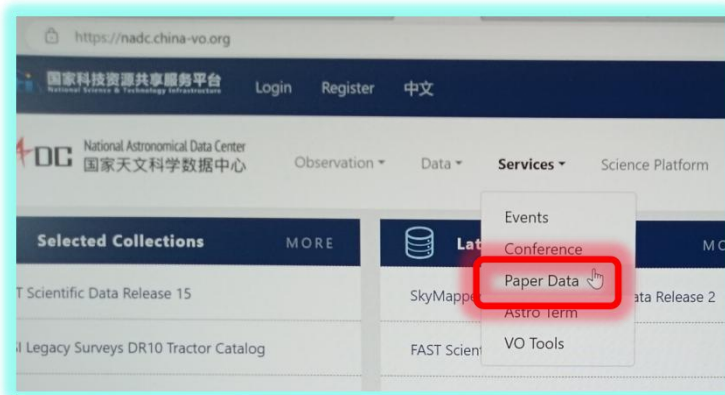
Data

accepted by
astronomical
journals

Model

Animation

Code



provides long-term storage and open access service

Paper Data

fill in metadata

title of data set and paper, data description, journal name

upload evidence

acceptance letter, full paper

calibrate data

standard format of the data(format asked by the journal)

upload data

Vol. < 5GB on the web, Vol. > 5 GB rsync

registration of DOI CSTR

complete within two working days

update data version

if need to change data, update new version

Paper Data

click 'datasources' get to the paper data

Abstract

Ionized-gas Metallicity of the Strong [O III]λ5007 Emission-line Compact Galaxies in the LAMOST Survey

Show affiliations

Liu, Siqui ; Luo, A. -Li ; Zhang, Wei ; Kong, Xiao ; Zhang, Yan-Xia ; Shen, Shi-Yin ; Zhao, Yong-Heng

This article reports a sample of 1830 strong [O III]λ5007 emission-line compact galaxies discovered with the LAMOST spectroscopic survey and the photometric catalog of the Sloan Digital Sky Survey. We newly identify 402 spectra of 346 strong [O III]λ5007 emission-line compact galaxies by finding compact isolated point sources. Combined with the samples in our previous work, this returns a sample of 1830 unique strong [O III]λ5007 emission-line compact galaxies with 2033 spectra of $z \leq 0.53$. For the sources with 2:[O III]λ4363 detections, we calculate the gas-phase metallicity with the direct- T_e method, and verify that the strong-line metallicity diagnostics calibrated with the direct- T_e method also applies to this sample. The strong [O III]λ5007 emission-line compact galaxies fall below several T_e -calibrated mass-metallicity relations. The N/O measurements of the strong [O III]λ5007 emission-line compact galaxies mainly locate at a plateau at low metallicity, indicating the product of primary nucleosynthesis. The Ne3O2 and O32 relation follows a tight linear relation with no redshift evolution. The Ne3O2 anticorrelates with the stellar mass, and at fixed stellar mass the Ne3O2 increases with the redshift. Eight sources with asymmetric [O III]λ5007 emission-line profiles have been identified, however with no [O III]λ4363 detection, which proves the rich metal content and complex ionized-gas kinematics within the galaxies. Higher-resolution spectroscopy will be necessary to identify the ionized-gas components in detail.

FULL TEXT SOURCES

Publisher

Proprint

DATA PRODUCTS

SIMBAD (8)

CDS (1)

RELATED MATERIALS (1)

Catalog: 2023yCat...22670016L

GRAPHICS

Click to view more

main version direct to the newest,
and the subversion direct to itself

Near-Infrared Ca II Triplet-based stellar activity database

Huang Xin

A new stellar database of the Ca II Triplet, which serves as an indicator for characterizing the chromospheric activity of stars. It includes R, R+ index for Ca II IRT, stellar parameters and other valuable information, the uncertainties of the index also be given. More detail about it please read our paper. In version 1.0, there are a very small number of entries (66) that lack Gaia magnitude data. Upon investigation, this is due to Gaia having two different sourceid for the same target. This has been corrected. Additionally, the mrt format file for this database has been added. The provided manuscript is the version at the time of acceptance. Please refer to the formal published article for accuracy.

Files

File Name	Size	Download
Call_IRT_index_LAMOST_DR9_LRS.csv	130.35 MB	
Call_IRT_index_LAMOST_DR9_LRS.mrt	168.74 MB	
manuscript_accepted.pdf	4.47 MB	

download here

Paper Information

Paper Title: Near-Infrared Ca II Triplet As A Stellar Activity Indicator: Library and Comparative Study

Publication: ApJS

Identifier

ctr: 11379.11.101395

DOI: 10.12149/101395

VO Identifier: ivo://China-VO/paperdata/101395

Publication date: 2024-03-15

Citation Guidelines

Huang Xin et al. 2024. Near-Infrared Ca II Triplet-based stellar activity database. Version 2.0. <https://doi.org/10.12149/101395>

@misc{10.12149/101395, doi = {10.12149/101395}, url = {https://doi.org/10.12149/101395}, author = {Huang Xin}, title = {Near-Infrared Ca II Triplet-based stellar activity database}, version = {2.0}, publisher = {National Astronomical Data Center of China}, year = {2024}}

Version 2.0 (current)

10.12149/101395 2024-03-15

Version 1.0

10.12149/101246 2024-02-19

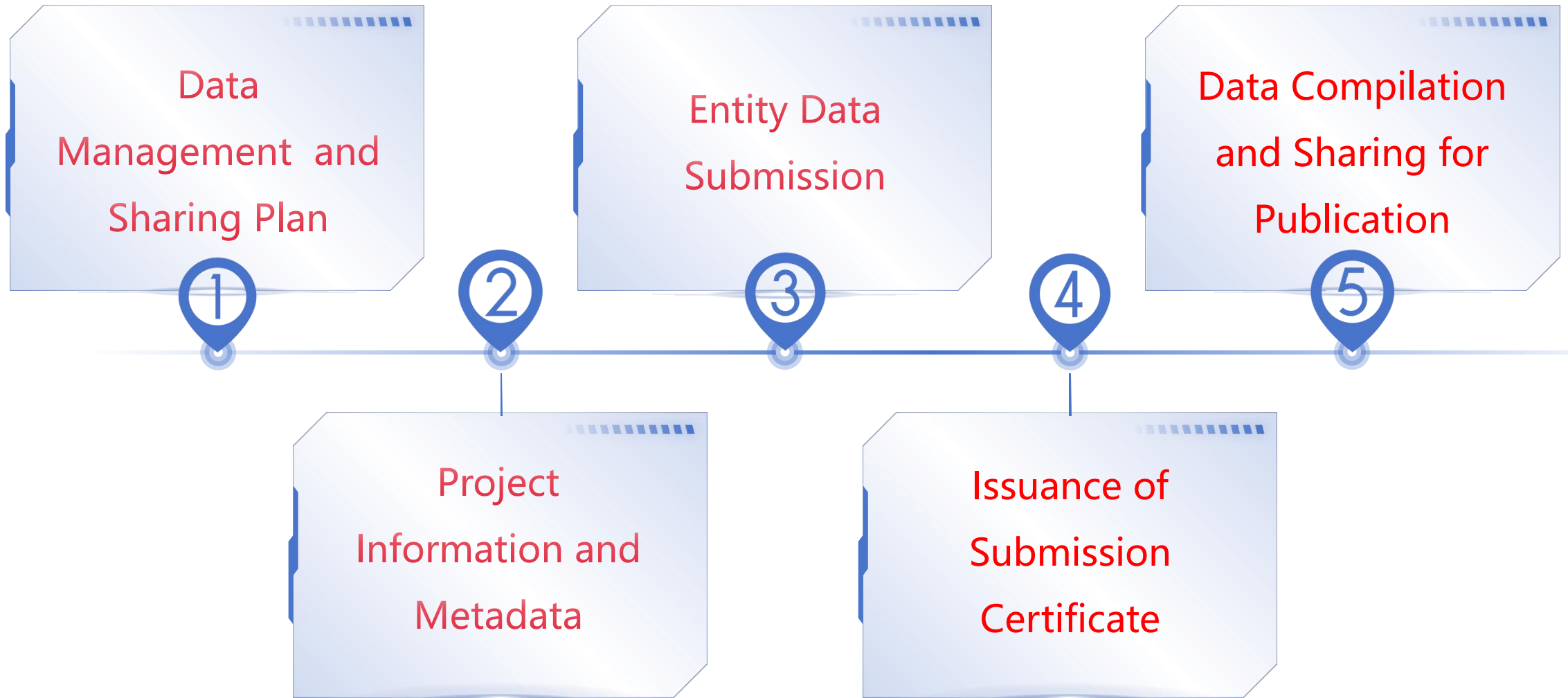
Main

This DOI represents all versions, and will always resolve to the latest one.

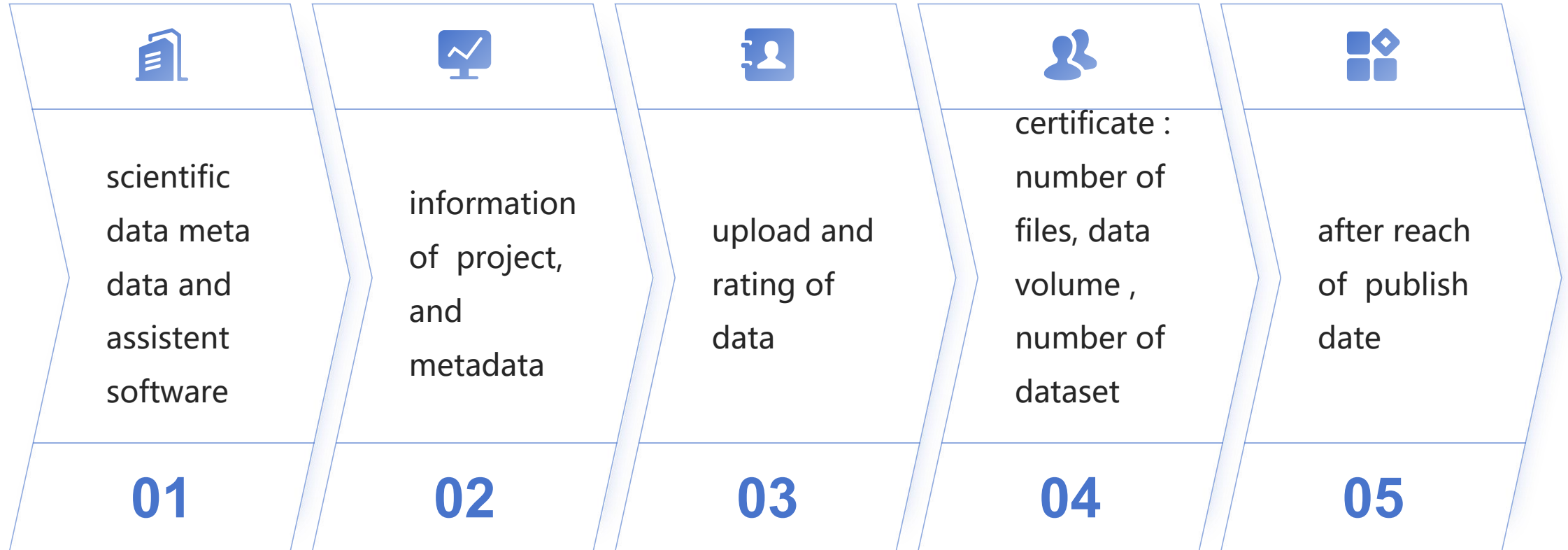
10.12149/101245 2024-02-19

citation and versions

Scientific Project Data



Scientific Project Data



10 scientific projects, involving 99 datasets, with a data volume of 329TB!

Outreach Data

nine datasets

- Precursor—Guo Shoujing's Astronomical Achievements
- Eclipse Calculator
- Recommended Chinese Names of Mars Features
- Recommended Chinese Names of Mars Features (2nd release)
- Atlas of remote sensing images of Jilin-1 of world-famous observatories
- Map of the surface of the Moon (the Second Edition)
- Four seasons star chart
- Multiband Astronomical Observations and Data Processing Graduate Course Optical Infrared Data Processing Data
- Map of the surface of the Moon (the Third Edition)

Mirror Data

21 datasets

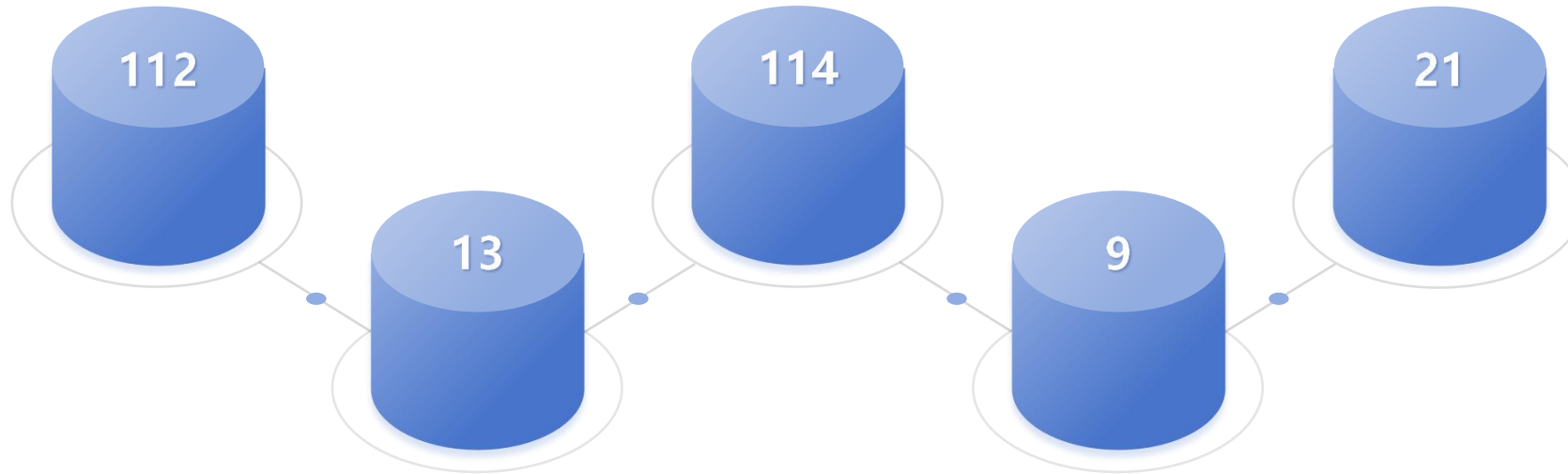
- U. S. Naval Observatory CCD Astrograph Catalog(UCAC1 - 5)
- Wide - field Infrared Survey Explorer (AllSky All WISE CatWISE)
- Sloan Digital Sky Survey(DR12 DR14)
- The Pan - STARRS1 Data Release 1
- GAIA Data Release 1 - DR2 - EDR3 - DR3
- DESI Legacy Surveys Tractor Catalog(DR9 DR10)
- 2MASS All - Sky Data Release
- AAVSO Photometric All Sky Survey (APASS) Data Release 10
- Distances to 1.47 billion stars in Gaia EDR3 (2021)
- SkyMapper Southern Sky Survey Data Release 2

Statistics

Geleran Dataset

Paper Data

Mirror Data



Scientific Project Data

Outreach Data

first level

Meta data

43 items for the core metadata: including 22 necessary and 21 not necessary

will update metadata for different data type

e.g. paperdata

表 2-1: 核心元数据

字段	定义	必选	说明
title	VARCHAR	Y	数据集名称 (中英文)
shortname	VARCHAR	N	短名称 (英文)
description	TEXT	Y	说明文档 (中英文)
create_date	DATE	Y	数据生产日期
update_date	DATE	Y	数据更新日期
publciate_date	DATE	N	(预计) 发布日期
authors	VARCHAR	Y	数据作者、生产者信息 (中英文)
affiliation	VARCHAR	Y	单位 (中英文)
address	VARCHAR	Y	作者地址 (中英文)
email	VARCHAR	Y	电子邮件
telephone	VARCHAR	Y	电话
contributors	VARCHAR	N	数据贡献者 (中英文)
cb_affiliation	VARCHAR	N	数据贡献者单位 (中英文)
cb_address	VARCHAR	N	数据贡献者地址 (中英文)
cb_email	VARCHAR	N	数据贡献者电子邮件
cb_telephone	VARCHAR	N	数据贡献者电话
contacts	VARCHAR	N	数据联系人 (中英文)
ct_affiliation	VARCHAR	N	数据联系人单位 (中英文)
ct_address	VARCHAR	N	数据联系人地址 (中英文)
ct_email	VARCHAR	N	数据联系人电子邮件

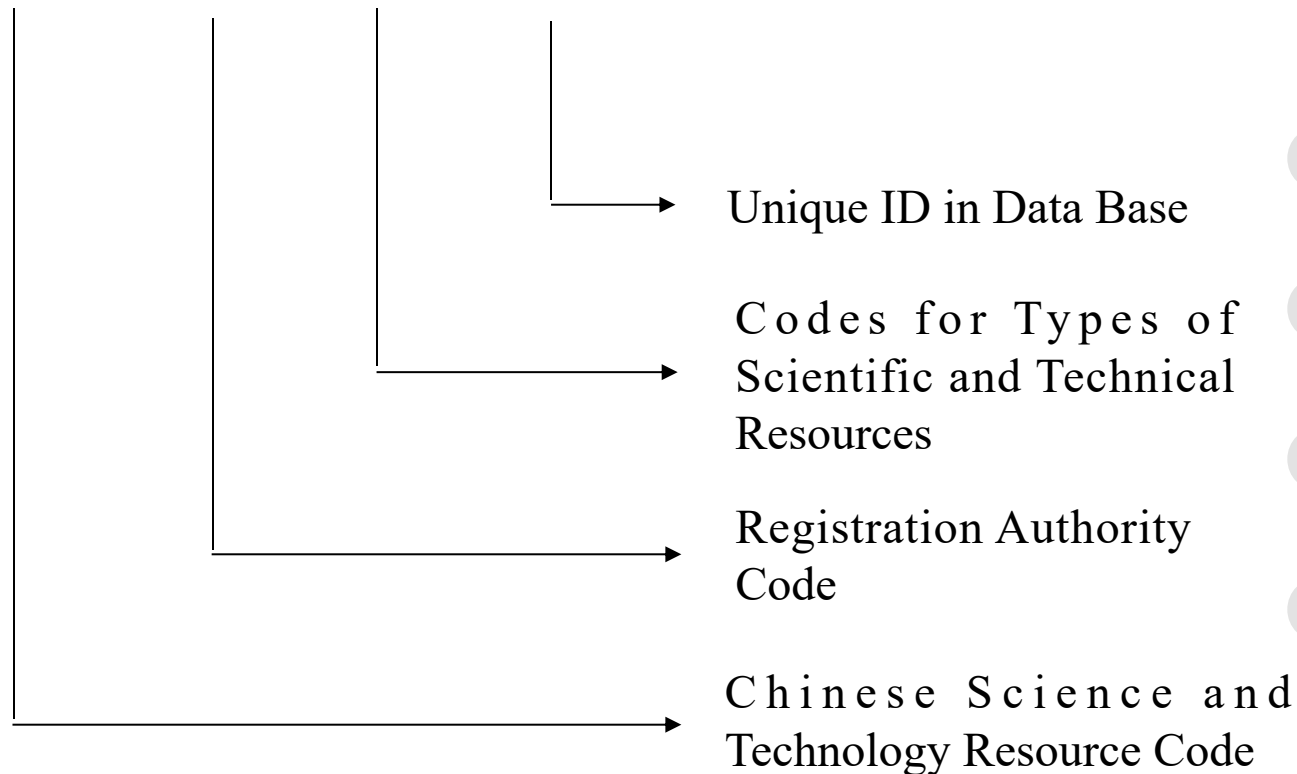
表 2-1 - 接上页

字段	定义	必选	说明
ct_telephone	VARCHAR	N	数据联系人电话
keywords	VARCHAR[]	Y	关键词 (3-5 个)
type	VARCHAR[]	Y	数据实体类型
waveband	VARCHAR[]	Y	数据波段
project	VARCHAR	N	数据来源项目
facility	VARCHAR	N	生产装置
subjects	VARCHAR[]	Y	子学科
production_type	VARCHAR	Y	数据生产方式
data_type	VARCHAR[]	Y	数据格式类型
data_level	VARCHAR	Y	数据级别
production_age	VARCHAR	Y	生产年代
content_level	VARCHAR[]	Y	内容级别
data_class	VARCHAR	Y	数据发布分类
share_method	VARCHAR[]	Y	共享途径
share_scope	VARCHAR	Y	共享范围
application_procedure	TEXT	N	申请流程
files	INT	Y	文件数
filesize	INT	Y	文件量 (单位字节)
catalogs	INT	N	星表数目
catalog_rows	INT	N	星表记录总数目
usage	TEXT	N	数据使用说明
bibcode	VARCHAR	N	相关的论文编码
URL	VARCHAR[]	N	相关网址
acknowledgement	TEXT	N	致谢文本 (中英文)

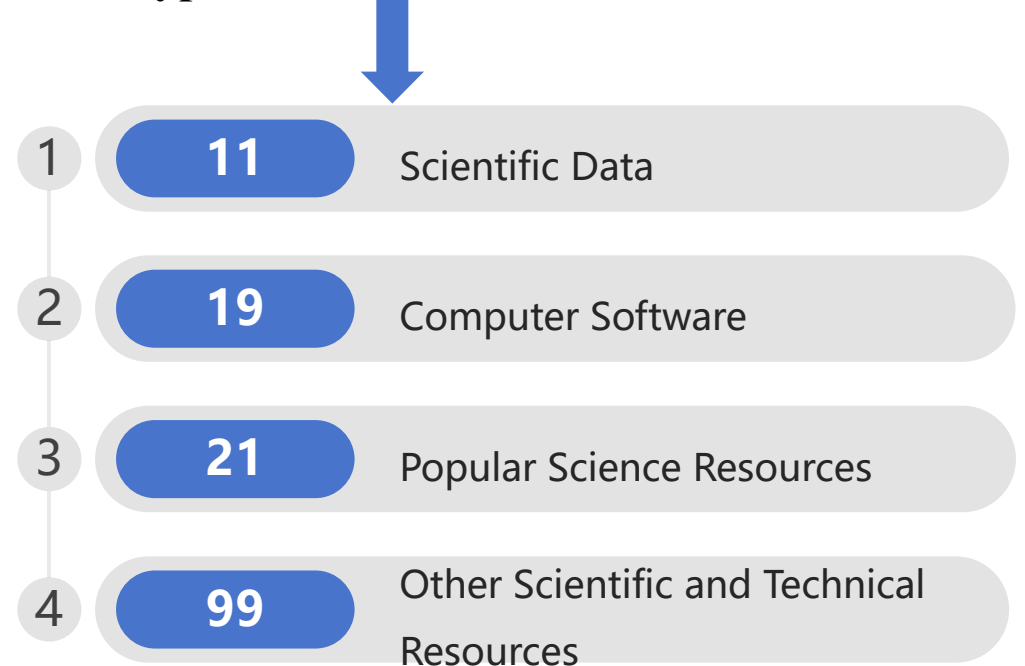
CSTR and DOI Registration

China Science and Technology Resources

CSTR.11379.11.100436



Codes for Types of Scientific and Technical Resources

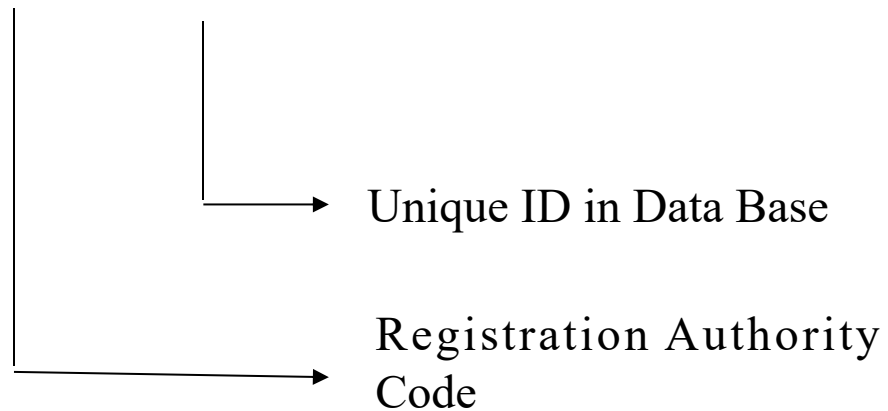


NO CSTR for mirror data!

CSTR and DOI Registration

Digital Object Identifier

10.12149/101179



- registration complete usually within two working days
- no DOI registration for mirror data

Summary

Five data types:Geleran
Dataset, Scientific
Project Data, Paper
Data, Outreach Data,
Mirror Data

1

There are 269
datasets in tole:
112 13 114 9 21

2

There are 43 items
for the core
metadata

3

CSTR and DOI are
provided, except for
mirror data

4

Experiences & Lessons

