



The implementation of DOI in NADC

Yihan Tao

On behalf of National Astronomical Data Center of China

IVOA InterOp Meeting, Bologna, Italy
10 May 2023

Outline

- NADC data resources
- DOI implementation
- DOI vs. CSTR



NADC Data Resources

- Observation dataset of domestic telescopes
 - LAMOST, FAST, BASS, etc.
- Mirror Dataset
 - Gaia, SDSS, DESI, etc.
- Research Project data
- Paper-related data/software
- EPO resources

The screenshot displays the National Astronomical Data Center (NADC) website. The header includes the NADC logo and navigation links for 'Observation', 'Data', 'Services', 'Science Platform', and 'Public'. A search bar is located in the top right corner. Below the header, there is a section titled 'All categories (Total number of datasets: 88)' with a 'Hide categories' button. This section lists various telescope and project categories, such as 'Radio (13)', 'Millimeter (1)', 'Optical (62)', and 'telescope or project' categories like 'Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) (43)'. Below this, there is a 'data type' section with a 'published time 1' filter. The main content area features several data release announcements, each with a title, publish time, DOI, and CSTR. The releases include: 'LAMOST Data Release 10 V1.0' (2023-09-30), 'LAMOST Data Release 11 V0 Q1' (2023-02-23), 'AST3-2 SN Survey Data Release' (2023-02-10), 'LAMOST Data Release 9 V1.1' (2023-02-02), and 'FAST Scientific Data Release 12' (2023-02-01). Each release includes a brief description of the data and its scope.

Data types and PIDs

- NADC data resource types and PIDs

	IVOID	DOI	CSTR
Observation Dataset	✓	✓	✓
Mirror Dataset	✓		
Research Project data	✓		✓
Paper-related data	✓	✓	✓
EPO resources	✓	✓	✓

LAMOST Data Release 8 V1.0
Publish Time:2021-03-31

LAMOST Data Release 8 V1.0 includes spectra and catalogs obtained by LAMOST low/medium resolution survey during October 24th 2011 and May 25th 2020. For the low resolution survey, there are 5,207 plates observed, 10,388,423 stellar spectra, 219,776 galaxy spectra, 71,786 quasar spectra and 534,091 unknown object spectra. For the medium resolution survey, there are 6,038,218 spectra, among them 1,479,127 non time-domain spectra, 4,599,091 time-domain spectra. In addition, there are dozens of catalogs list spectral parameters for the low/medium resolution survey respectively.

Data access

Data Website

◆ INTRODUCTION ◆

Data format	application/fits , text/csv , image/png
Data volume	11 tables, 69936029 rows, 1881250.07 MB
Sharing method	online
Sharing scope	Share with conditions
DOI	10.12149/100361
CSTR	11379.11.100361
VO Identifier	ivo://China-VO/data/LAMOST/DR8/V1.0

◆ CONTACT AUTHOR ◆

Author	LAMOST Operation and Development Center
Email	support(at)lamost.org
Institution	National Astronomical Observatories, Chinese Academy of Sciences

◆ CONTACT PUBLISHER ◆

Email	support(at)china-vo.org
Phone	+86-10-64807973
Institution	National Astronomical Data Center

◆ TAGS ◆

waveband	Optical
telescope or project	Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)
subject	Cosmology and Galaxies Star and Interstellar Matter Astronomical Technology and Methods
data type	Spectrum Data Catalog Data

DOI Implementation

- Observation/EPO dataset
 - Assigns DOI when the dataset is released
 - Automatically generate xml from metadata for DOI registration
 - Submit to Chinadoi



100632	LAMOST光谱巡天第八次数据发布第2.0版 LAMOST Data Release 8 V2.0	普通数据集	2022-09-19 00:00:00	2022-09-19 00:00:00	Published/Active	Edit Preview View Metadata DOI xml	Upload
100633	LAMOST光谱巡天第八次数据发布第2.0版 低分辨率数据 LAMOST Data Release 8 V2.0 Low Resolution	普通数据集	2022-09-19 00:00:00	2022-09-19 00:00:00	Published/Active	Edit Preview View Metadata DOI xml	Upload
100634	LAMOST光谱巡天第八次数据发布第2.0版 中分辨率数据 LAMOST Data Release 8 V2.0 Medium Resolution	普通数据集	2022-09-19 00:00:00	2022-09-19 00:00:00	Published/Active	Edit Preview View Metadata DOI xml	Upload

```
1 <doi_batch version="1.0.0">
2 <head>
3 <doi_batch_id>2023042811231</doi_batch_id>
4 <timestamp>2023042811231</timestamp>
5 <depositor>
6 <name>Chinese Astronomical Data Center</name>
7 <email_address>support@china-vo.org</email_address>
8 </depositor>
9 <registrant>Chinese Astronomical Data Center</registrant>
10 </head>
11 <body>
12 <science_data>
13 <database>
14 <contributors>
15 <person_name contributor_role="author" sequence="first">LAMOST Operation and Development Center</person_name>
16 <organization contributor_role="author" sequence="first">National Astronomical Observatories, Chinese Academy of Sciences</organization>
17 </contributors>
18 <title>
19 <title>LAMOST Data Release 8 V2.0 Low Resolution</title>
20 </title>
21 <description>LAMOST Data Release 8 LRS data V2.0 includes spectra and catalogs obtained by LAMOST low resolution survey during October 24th, 2011 and May 27th 2020. For the low resolution survey, there are 10,633,915 spectra obtained totally, including 10,336,752 stellar spectra, 224,752 galaxy spectra, 72,061 quasar spectra, and related catalogs published.</description>
22 <publisher>
23 <publisher_name>Chinese Astronomical Data Center</publisher_name>
24 <publisher_place>20A Datun Road, Chaoyang District, Beijing, China, 100101</publisher_place>
25 </publisher>
26 <doi_data>
27 <doi>10.12149/100633</doi>
28 <timestamp>2023042811231</timestamp>
29 <resource><|<url>[https://nadc.china-vo.org/res/r100633]</resource>
30 </doi_data>
31 </database>
32 </science_data>
33 </body>
34 <dataset dataset_type="record">
35 <contributors>
36 <person_name contributor_role="author" sequence="first">LAMOST Operation and Development Center</person_name>
37 <organization contributor_role="author" sequence="first">National Astronomical Observatories, Chinese Academy of Sciences</organization>
38 </contributors>
39 <title>
40 <title>LAMOST DR8 V2.0 LRS General Catalog</title>
41 </title>
42 <dataset_date>
43 <creation_date>
44 <year>2022</year>
45 <month>09</month>
46 <day>18</day>
47 </creation_date>
48 <publication_date media_type="online">
49 <year>2022</year>
50 <month>09</month>
51 <day>20</day>
52 </publication_date>
53 <update_date>
54 <year>2022</year>
55 <month>09</month>
56 <day>19</day>
57 </update_date>
58 </dataset_date>
59 <item_number>100635</item_number>
60 <description>LAMOST DR8 V2.0 LRS General Catalog lists information of targets obtained by LAMOST low resolution survey during October the 24th 2011 and May the 27th 2020. There are parameters for all of the published spectra. Two formats of the LAMOST LRS general catalog, a FITS table and a CSV table, are available on the website http://drs.lamost.org/v2.0/catalogue.
61 </description>
62 <format MIME_type="application/csv">application/csv</format>
63 <doi_data>
64 <doi>10.12149/100635</doi>
65 <timestamp>2023042811231</timestamp>
66 <resource><|<url>[https://nadc.china-vo.org/res/r100635]</resource>
67 </doi_data>
68 </dataset></dataset dataset_type="record">
69 <contributors>
70 <person_name contributor_role="author" sequence="first">LAMOST Operation and Development Center</person_name>
71 <organization contributor_role="author" sequence="first">National Astronomical Observatories, Chinese Academy of Sciences</organization>
72 </contributors>
73 <title>
74 <title>LAMOST DR8 V2.0 LRS A, F, G and K Type Star Catalog</title>
75 </title>
76 <dataset_date>
77 <creation_date>
78 <year>2022</year>
79 <month>09</month>
80 <day>19</day>
81 </creation_date>
82 <publication_date media_type="online">
83 <year>2022</year>
84 <month>09</month>
85 <day>30</day>
86 </publication_date>
87 <update_date>
88 <year>2022</year>
89 <month>09</month>
90 <day>18</day>
91 </update_date>
92 </dataset_date>
93 <item_number>100636</item_number>
94 <description>LAMOST DR8 V2.0 LRS Parameter Catalog lists atmospheric parameters of targets observed by LAMOST medium resolution survey during October the 24 th 2011 and May the 27th 2020. There are atmospheric parameters for FGK stars. Two formats of the LAMOST LRS Parameter catalog, a FITS table and a CSV table, which are available from the website http://drs.lamost.org/v2.0/catalogue.</description>
95 </description>
96 <format MIME_type="application/csv">application/csv</format>
97 <doi_data>
98 <doi>10.12149/100636</doi>
```

DOI Implementation

- Granularity

- Datasets will have subsets/catalogues, defined case by case

- The main dataset is declared as <database>, the subsets/catalogues are declared as <dataset>, each given a DOI

- Can be download and submit for DOI registration in one xml file

```
11 <body>
12 <science_data>
13 <database>
14 <contributors>
15 <person_name contributor_role="author" sequence="first">LAMOST Operation and Development Center</person_name>
16 <organization contributor_role="author" sequence="first">National Astronomical Observatories, Chinese Academy of Sciences</organization>
17 </contributors>
18 <titles>
19 <title>LAMOST Data Release 8 V2.0 Low Resolution</title>
20 </titles>
21 <description>LAMOST Data Release 8 LRS data V2.0 includes spectra and catalogs obtained by LAMOST low resolution survey during October 24th,
22 survey, there are 10,633,515 spectra obtained totally, including 10,336,752 stellar spectra, 224,752 galaxy spectra, 72,061 quasar spectra, a
23 <publisher>
24 <publisher_name>Chinese Astronomical Data Center</publisher_name>
25 <publisher_place>20A Datun Road, Chaoyang District, Beijing, China, 100101</publisher_place>
26 </publisher>
27 <doi_data>
28 <doi>10.12149/100633</doi>
29 <timestamp>20220428113231</timestamp>
30 <resource><![CDATA[https://nadc.china-vo.org/res/r100633]]></resource>
31 </doi_data>
32 </database>
33 <dataset dataset_type="record">
34 <contributors>
35 <person_name contributor_role="author" sequence="first">LAMOST Operation and Development Center</person_name>
36 <organization contributor_role="author" sequence="first">National Astronomical Observatories, Chinese Academy of Sciences</organization>
37 </contributors>
38 <titles>
39 <title>LAMOST DR8 V2.0 LRS General Catalog</title>
40 </titles>
41 <dataset_date>
42 <creation_date>
43 <year>2022</year>
44 <month>09</month>
45 <day>19</day>
46 </creation_date>
47 <publication_date media_type="online">
48 <year>2022</year>
49 <month>09</month>
50 <day>30</day>
51 </publication_date>
52 <update_date>
53 <year>2022</year>
54 <month>09</month>
55 <day>19</day>
56 </update_date>
57 </dataset_date>
58 <item_number>100635</item_number>
59 <description>LAMOST DR8 V2.0 LRS General Catalog lists information of targets obtained by LAMOST low resolution survey during October the 24th
60 parameters for all of the published spectra. Two formats of the LAMOST LRS General Catalog, a FITS table and a CSV table, are available on th
61 </description>
62 <format MIME_type="application/csv">application/csv</format>
63 <doi_data>
64 <doi>10.12149/100635</doi>
65 <timestamp>20220428113231</timestamp>
66 <resource><![CDATA[https://nadc.china-vo.org/res/r100635]]></resource>
67 </doi_data>
```

DOI Implementation



批处理操作

用户可以选择DOI注册、DOI查询、元数据查询、元数据修改4种不同的操作类型，并选择文件上传，完成相应的批量处理工作。

在提交了批处理操作后，可通过操作记录对批处理操作结果进行浏览或查询，从而确保您的工作有效进行。

请选择操作类型:

DOI注册

请选择上传的文件:

[选择文件](#) 未选择任何文件

[上传](#)

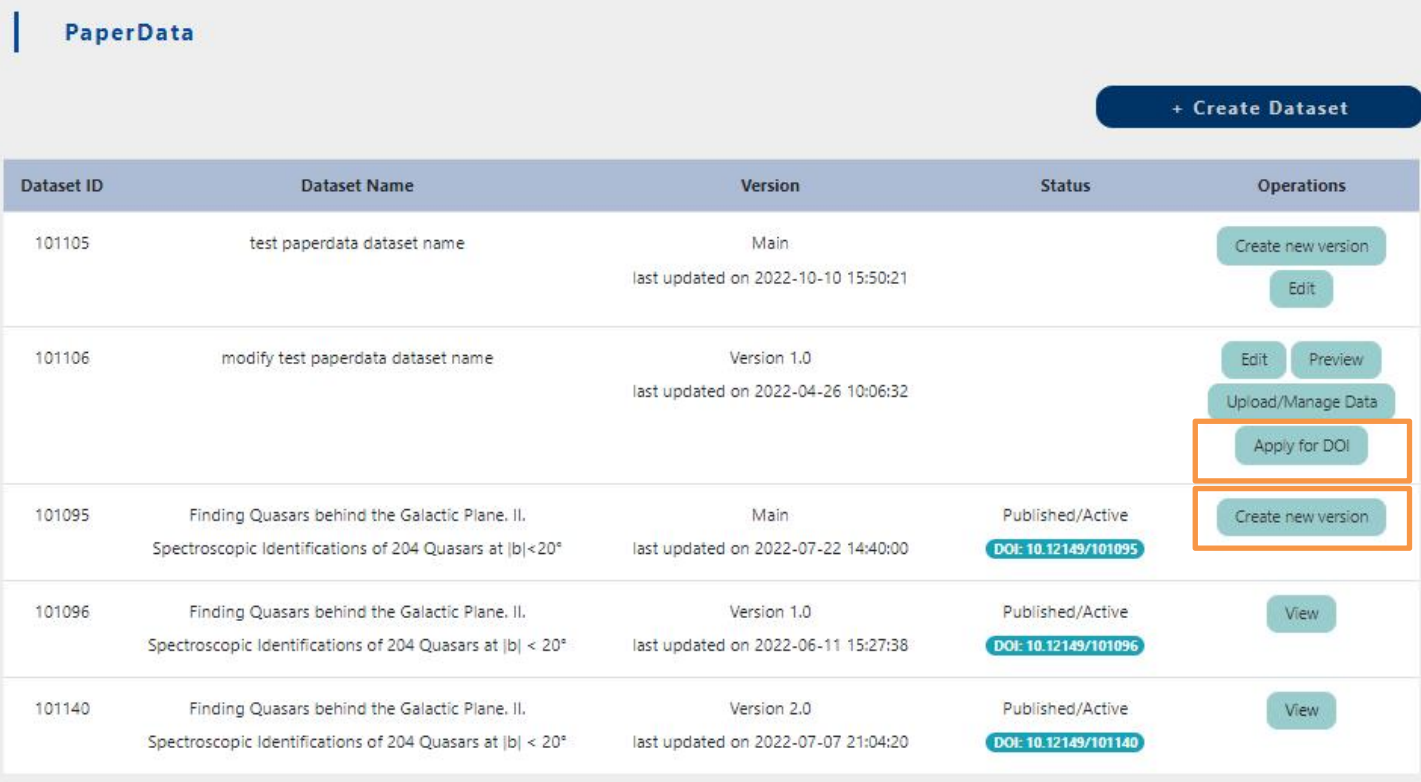
操作记录

请选择状态 请选择类型 提交时间 至 [搜索](#)

提交文件名	系统处理文件名	提交时间 ▲ ▼	操作类型	目前状态	完成时间 ▲ ▼	操作结果
101130.xml	73820230421095...	2023年04月21日...	元数据修改	未审核	2023年04月21日...	处理中
101130.xml	73820230421095...	2023年04月21日...	DOI注册	已完成	2023年04月22日...	查看
101241.xml	73820230418134...	2023年04月18日...	DOI注册	已完成	2023年04月19日...	查看
101241.xml	73820230418134...	2023年04月18日...	DOI注册	已完成	2023年04月19日...	查看
101239.xml	73820230413123...	2023年04月13日...	DOI注册	已完成	2023年04月14日...	查看
101234.xml	73820230403132...	2023年04月03日...	DOI注册	已完成	2023年04月04日...	查看
100718.xml	73820230330141...	2023年03月30日...	DOI注册	已完成	2023年03月30日...	查看
100717.xml	73820230330140...	2023年03月30日...	DOI注册	已完成	2023年03月30日...	查看
100716.xml	73820230330140...	2023年03月30日...	DOI注册	已完成	2023年03月30日...	查看
101226.xml	73820230329153...	2023年03月29日...	DOI注册	已完成	2023年03月30日...	查看
101208.xml	73820230328140...	2023年03月28日...	DOI注册	已完成	2023年03月29日...	查看
101216.xml	73820230314130...	2023年03月14日...	DOI注册	已完成	2023年03月15日...	查看
101218.xml	73820230313212...	2023年03月13日...	DOI注册	已完成	2023年03月14日...	查看
101210.xml	73820230301103...	2023年03月01日...	DOI注册	已完成	2023年03月01日...	查看
100697.xml	73820230227154...	2023年02月27日...	DOI注册	已完成	2023年02月28日...	查看

DOI Implementation

- Paper-related Data
 - On a request basis
 - One dataset per paper
 - Metadata check by administrators
 - paper acceptance letter needs to be uploaded
 - Once a DOI is issued, change of content/deletion is not allowed, but new version could be created.



The screenshot shows the PaperData interface with a table of datasets. The table has columns for Dataset ID, Dataset Name, Version, Status, and Operations. The 'Apply for DOI' button for dataset 101106 and the 'Create new version' button for dataset 101095 are highlighted with orange boxes.

Dataset ID	Dataset Name	Version	Status	Operations
101105	test paperdata dataset name	Main last updated on 2022-10-10 15:50:21		Create new version Edit
101106	modify test paperdata dataset name	Version 1.0 last updated on 2022-04-26 10:06:32		Edit Preview Upload/Manage Data Apply for DOI
101095	Finding Quasars behind the Galactic Plane. II. Spectroscopic Identifications of 204 Quasars at $ b < 20^\circ$	Main last updated on 2022-07-22 14:40:00	Published/Active DOI: 10.12149/101095	Create new version
101096	Finding Quasars behind the Galactic Plane. II. Spectroscopic Identifications of 204 Quasars at $ b < 20^\circ$	Version 1.0 last updated on 2022-06-11 15:27:38	Published/Active DOI: 10.12149/101096	View
101140	Finding Quasars behind the Galactic Plane. II. Spectroscopic Identifications of 204 Quasars at $ b < 20^\circ$	Version 2.0 last updated on 2022-07-07 21:04:20	Published/Active DOI: 10.12149/101140	View

DOI Implementation

- Versioning
 - Different DOI for each version of a dataset, also a DOI for the main version

Finding Quasars behind the Galactic Plane. II. Spectroscopic Identifications of 204 Quasars at $|b| < 20^\circ$

Yuming Fu

Quasars behind the Galactic plane (GPQs) are important astrometric references and valuable probes of Galactic gas, yet the search for GPQs is difficult due to severe extinction and source crowding in the Galactic plane. In this paper, we present a sample of 204 spectroscopically confirmed GPQs at $|b| < 20^\circ$, 191 of which are new discoveries. This GPQ sample covers a wide redshift range from 0.069 to 4.487. For the subset of 230 observed GPQ candidates, the lower limit of the purity of quasars is 85.2%, and the lower limit of the fraction of stellar contaminants is 6.1%. Using a multicomponent spectral fitting, we measure the emission line and continuum flux of the GPQs, and estimate their single-epoch virial black hole masses. These GPQs have higher black hole masses and continuum luminosities in comparison to the SDSS DR7 quasar sample. This is due to a selection effect raised from Galactic extinction and target magnitude. The spectral-fitting results and black hole mass estimates are compiled into a main spectral catalog, and an extended spectral catalog of GPQs. The successful identifications prove the reliability of both our GPQ selection methods and the GPQ candidate catalog, shedding light on the astrometric and astrophysical programs that make use of a large sample of GPQs in the future.

Description of `gpq2_204_GPQ_spec.tar.gz`:

- Name: Spectra of the 204 identified Quasars behind the Galactic Plane
- Short name: GPQ Spectra
- This tar package includes reduced one-dimensional spectra of the 204 identified GPQs in FITS format. The FITS files are named in the format of "`spec_3hmmss.ss+ddmss.s_Telescope.fits`". The FITS files can be viewed with the `plot` task of IRAF or read with other standard FITS I/O routines (e.g. `astropy.io.fits`, `specutils`).

Description of `gpq2_contam_spec.tar.gz`:

Identifier

DOI: 10.12149/101140
Publication date: 2022-07-22

Versions

Version 2.0 (current)
10.12149/101140 2022-07-22
Version 1.0
10.12149/101096 2022-06-13
Main
This DOI represents all versions, and will always resolve to the latest one.
10.12149/101095 2022-03-18

[Back to PaperData Catalogue](#)

DOI vs. CSTR

- **China Science & Technology Resource Identifier(CSTR)** is proposed by the Ministry of Science and Technology (MOST) of China for the identification, cataloguing, registration, publication, maintenance and management of scientific and technological resources in China.
- NADC is a Identification Registration Agency of CSTR .

Types of S&T resources

01 Large-scale scientific equipments	13 Reports
02 Major S&T infrastructures	14 Papers
03 Research and experimental bases	15 Books
04 Plant genetic resources	16 Patents
05 Amlinal genetic resources	17 Standards
06 Microbial genetic resources	18 Measurement criteria
07 Human genetic resources	19 Software
08 Reference material	20 New products, new processes and new materials
09 Experiment materials	21 Education resources
10 Specimen	22 S&T cases
11 Scientific data	99 Other S&T resources
12 Atlases	

doi://10.12149/<resource_id>
cstr:11379.11.<resource id>

Chinese National standard
GB/T 32843-2016



Access and Citation Tracking

Common Science and Technology Resource Identification

OUR SERVICE COOPERATION ABOUT US Tao Yihan EN 中文

Together for Open Identifier, Together for Open Science

Publications Name, Keywords, Authors

244,680,723 PUBLICATIONS 208,642 DISSERTATIONS 7,781,196 SCIENCE DATA 22,875 PREPRINTS 254,771 PATENTS 138,284 SPECIES LISTS 13,284 INSTRUMENTS 606,981 PROJECTS 176,059 GERMPLASMS 109,740 INSTITUTIONS 12,336,757 RESEARCHERS

New Data Science Data ↑ 15 Priprint ↑ 2 Publications ↑ 1060 Update time: 2023-05-10

Newest Cited

Data source: CSCD | Google Scholar

Is the Fama and French three factor model robust to the pricing... 2023-04-09
Oghenovo Obrimah
CSTR:31253.11.sciencedb.01458

Cited by
Obrimah, O.A. 2022c. Idiosyncratic Risk, Return Volatility, and Return Skewness[DB/OL]. Science DataBank[2023-03-08].
https://cstr.cn/31253.11.sciencedb.01458

Newest Registrations

on CSTR

Scientific and technological achievements cross-media interacti... 2023-05-09
国家基础学科公共科学数据中心
科技冬奥VR,跨媒体互动
CSTR:16666.11.NBSDC.MW9XO2ZI

Rock mechanics data of subsalt superdeep carbonate rocks in T... 2023-05-09
国家基础学科公共科学数据中心
三轴力学测试,盖层破裂数值模拟,岩石力学性质,平均突破压力
CSTR:16666.11.NBSDC.YUJUFKSA

Newest Resolutions

on CSTR

WuDaoCorpora Text 2023-05-09
科学数据银行
CSTR:31253.11.SCIENCEDB.000126.00004
Source: United States | Resolutions: 99

光电微弱信号检测系统 2023-05-09
CASIP
CSTR:32117.16.20120817.CN201220410473.5
Source: Singapore | Resolutions: 18

Most Cited

Data source: CSCD | Google Scholar

Protective Zika vaccines engineered to eliminate enhancement ... 2023-05-07
Lianpan Dai,Kun Xu,Jinhe Li,Qingrui Huang,Jian Song,Yuxuan Han,Tianyi Zhe...
CSTR:57295.14.s41590-021-00966-6
SEMANTIC SCHOLAR 16 Crossref 15 Google Scholar 0

Dual ligand engagement for noncanonical inflammasome activa... 2023-05-07
Zhang-Hua Yang,Jiahuai Han
CSTR:57295.14.s41590-022-01188-0
SEMANTIC SCHOLAR 2 Crossref 2 Google Scholar 0

中文DOI

“casdc” 欢迎进入会员系统, 现在是: 2023年5月10日 05:30:23 星期三 安全退出

欢迎页 批处理操作 操作记录 实时查询 统计查询 用户信息

解析统计查询 注册统计查询

期刊于 2023-04 的排名前十的DOI量为:

DOI	解析量
10.12149/101234	62
10.12149/101242	53
10.12149/101216	46
10.12149/101211	12
10.12149/101240	11
10.12149/101204	8
10.12149/101208	8
10.12149/101235	7
10.12149/101070	7
10.12149/101188	5

Summary

- DOI is one of the PID that NADC adopted
- DOI is useful for sharing, citing and locating data
- Using DOI to track data access and citations for every datasets could be explored further
- Learn recommendations and best practices from IVOA

