

LAMOST & China-VO Data System

Dongwei Fan

fandongwei@nao.cas.cn



2022/4/26

Data Websites

u NADC

- u Chinese Domestic Datasets
- u <https://nadc.china-vo.org/data/>
- u **CORE TRUST SEAL** verified in 2018

u LAMOST

- u special for LAMOST
- u serve DR1 ~ DR10 and their sub versions
- u <http://www.lamost.org/data/>

Data Release	Version	Low Resolution			Medium Resolution			Release Date	Public/Internal
		Plate	Footprint	Spectra	Plate	Footprint	Spectra		
DR10	v0	99		87,632	181		2,356,750	2022-03-03	Internal
DR9	v1.0	5,533		11,226,252	1,529		29,900,588	2022-03-31	Internal
	v0	326		538,137	440		7,778,953	2021-08-15	Internal
DR8	v1.1	5,207		10,927,525	1,089		22,141,635	2022-02-28	Internal
	v1.0	5,207		11,214,076	1,089		22,356,885	2021-03-31	
	v0	285		446,669	410		8,088,917	2020-08-11	
DR7	v2.0	4,922		10,421,197	679		14,096,967	2021-09-30	Public
	v1.3	4,922		10,640,255	679		14,210,021	2021-04-20	
	v1.2	4,922		10,599,979	679		14,375,293	2020-12-28	
	v1.1	4,922		10,602,012	680		14,225,777	2020-06-16	
	v1	4,926		10,606,416	680		14,391,013	2020-04-01	
	v0	355		558,412	351		8,524,640	2019-07-31	

LAMOST DR9 TOTAL								
Observed Plates: 5,633	Total Spectra: 11,226,252	Star: 10,907,516						
Galaxy: 242,669	QSO: 76,167	AFGK Stellar Parameters: 7,060,436						
M Giants, Dwarfs and Subdwarfs: 832,756	Line Indices of A Type Stars: 654,940	WD Star Catalog: 16,382						
CV Star Catalog: 489	Galaxy Population Synthesis Catalog: 23,770	Oso Emission line Feature Catalog: 17,825						

Low Resolution								
	Observed Plates	Total Spectra	Star	Galaxy	QSO	AFGK Stellar Parameters	M Giants, Dwarfs and Subdwarfs	Line Indices of A Type Stars
Pilot survey (20111024-20120817)	404	882,478	871,461	9,470	1,547	417,753	87,107	65,397
First year survey (20120928-20130603)	612	1,607,261	1,587,141	13,856	6,264	934,247	67,845	79,965

Datasets

- ∪ Antarctic Survey Telescopes
- ∪ Beijing-Arizona Sky Survey
- ∪ Chinese Small Telescope ARray Catalog
- ∪ Large Sky Area Multi-Object Fiber Spectroscopic Telescope — LAMOST
 - ∪ **biggest spectrum dataset in the world**, >11m low-resolution, >8m med-resolution
- ∪ Single-exposure data for LAMOST DR5
- ∪ Legacy Plates - Thousands of historical films
- ∪ The South Galactic Cap u-band Sky Survey

- ∪ Low Frequency Radio spectrum Explorer
- ∪ Mingantu Ultrawide Spectral Radioheliograph
- ∪ Solar Broad Band Spectrometer

- ∪

Data Types

- u Catalogue
- u Image
- u Spectrum
- u Light Curve
- u some other solar binary format
- u other

Services

- flexible mechanism to support instruments' data publishing
- User System and Access Control
 - Single Sign in/out
- Query Forms
- Online SQL - not TAP
- Interactive spectrum visualization
- Cross-match
- CMS for documents online editing
- IVOA Services
 - Web SAMP
 - Cone Search
 - SSAP
 - HiPS - Images & Catalogues
 - SIA - support image search, without image cutout

VORegistry

- u China-VO VORegistry upgrading...
 - u <https://registry.china-vo.org/>
- u based on China-VO metadata system
 - u collecting data and assign VO-ID, DOI, CSTR ...
- u not all services registered yet
- u btw, thanks for the registry scan services

adapt more VO specifications?

- u Of Course!
- u full SIA
- u full VORegistry
- u SSO
- u TAP
- u VOSpace
- u MOC
- u Datalink
- u ...

Difficulties

- arbitrary field query on billion rows catalog
 - 10 billion catalog is coming
 - testing postgres data sharding extension
- Huge Image Cutout
 - ~30TB images
- TAP implementation
 - need java based library
 - considering switch to Python
- how to download millions small files
 - no good solution
 - try don't download --> **work on online science platform**

if start again ...

- u distributed structure
- u high Availability/Reliability
- u to support more dataset by configuration, not just add some new code
- u more api & better documents

THANKS
Q & A

fandongwei@nao.cas.cn