

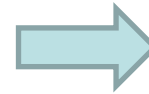
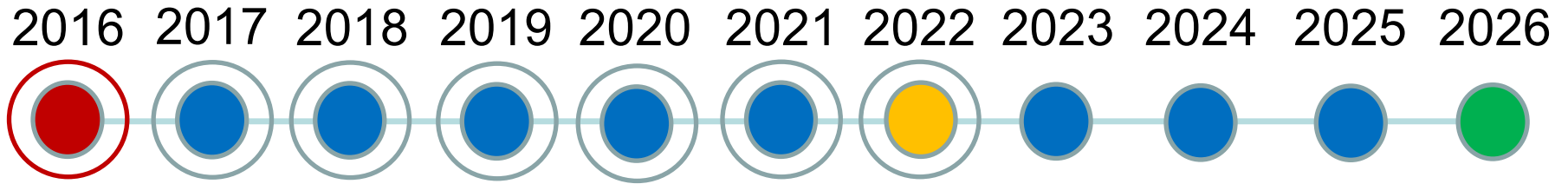
## 2022 DBAR Science Team Meeting

# DBAR Data Development

DBAR Secretariat

Beijing • China

# DBAR Timeline | Implementation Phase II



# DBAR Progress | DBAR Data Development



**SDGSAT-1**  
可持续发展科学卫星

## SDG Satellite Launched on 5 November 2021

- Thermal infrared + nighttime-light + multi-spectral
- Wide scale (300 km)
- High-resolution (10 m)



- Glimmer image of Paris /40m
- Multispectral image of entry of Yellow River /10m
- Thermal infrared image of Agsu in Xinjiang Uygur Autonomous Region /30m

*Explore new methods to sense Earth's environment*

**Data will be accessible at the end of 2022**

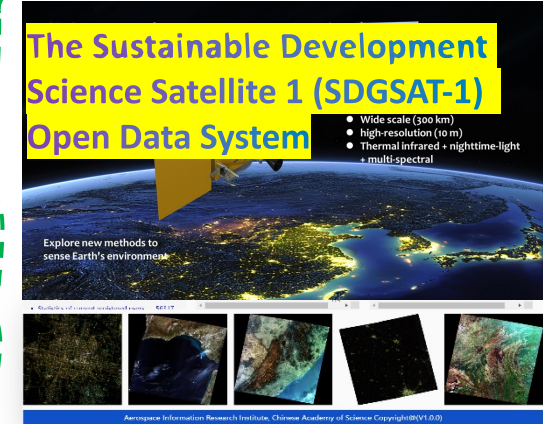
# DBAR Progress | DBAR Data Development



- **Data mining models** (classification, keyword search, tag cloud filtering, and association rules).
- **Multiple data acquisition patterns** (online downloading and API access).
- **Online services** for customized multiple data formats
- **Advanced functions** Data evaluation, validation, download, and cloud analysis for **customized needs**



- **SDG Data Hub for DBAR**
- **TerraView** (Big Earth Data applications)
- **DBAR and SDG News Library**(Multi-mode media data processing and display)
- **DBAR and SDG Knowledge Base**(Science popularization)



**SDGSAT-1 : Monitoring, evaluating and researching on indicators of SDGs. Depicts traces of anthropic activities by synergetic observing.**

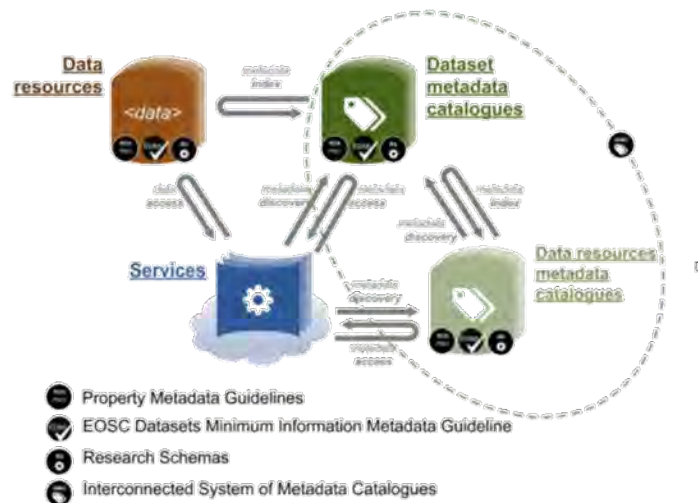
# DBAR Progress | DBAR Data Development

- Data interoperability is a problem affecting the interaction of entities at very different levels, and thus not only the technical operations.
- When merging two (or more) datasets into one common target dataset, we need to ensure that data from heterogeneous sources can be used in combination to view, query and analyse, or in short: we need to harmonize the datasets.

Therefore, set up a data harmonization process.

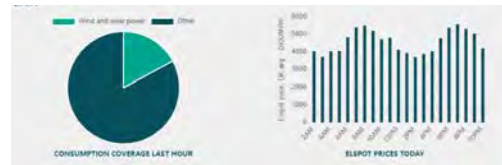


Coordination with CBAS Data Infrastructure



# DBAR Progress | DBAR Data Development

- Develop a catalogue of products and services that can be linked together or provide data that is interoperable for different science applications.
- This exercise will also be critical for compiling a major report on achievements of DBAR program in 2026.



**12 datasets found**

**Commercial gas amounts:**  
Daily commercial flow in the Danish gas market. The commercial flow can vary slightly from the physical flow.

**Gas**  
The quality of the gas in the Danish gas system has been determined by the Security Agency in the Gas Regulations section 11. The gas can come from Denmark, from Europe, from the North Sea, or originate from biomass from longer pipelines originating from the gas storage in LNG-Terms and Skovde. Consumers can get gas that is a mixture from several supply areas.

The gas must at all times comply with the Danish Gas Regulations and the quality specifications in the Gas Regulation. Information about variations in gas quality is published here by hour and month by month.

**Gas composition and quality for transmission**  
The dataset contains the hourly-min-calculated gas composition and other gas quality parameters for the Danish Transmission gas measured at Transtank - a gas quality measuring station in Esbjerg.

**Gas composition and quality for transmission, monthly**  
Gas quality data for the Danish Transmission gas that is representative for gas delivered from the transmission grid. The data are available but the data cannot be used for billing purposes. The basis of the dataset is hourly values and the temporal and minimal value is therefore dependent on the monthly and minimal hourly value.

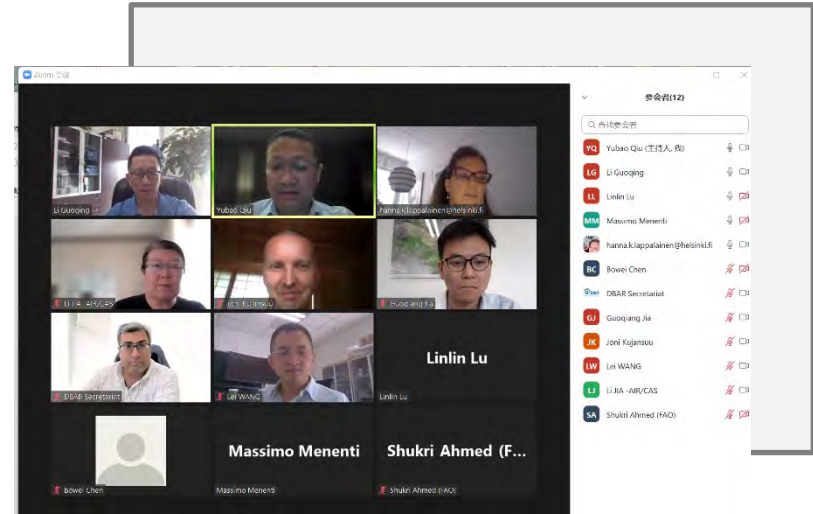


Good Example for Catalogue

## Featured Session

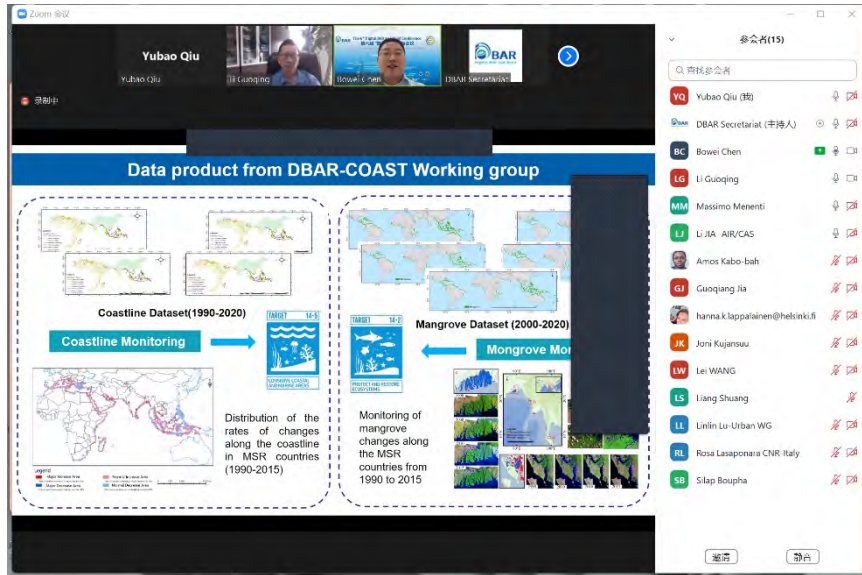
### Data Work Session (DBAR Secretariat, DBAR Data and DBAR EEC)

- Special session during DBAR 2022 on **Data Sharing and Applications by DBAR WGs and ICoEs.**
- Identify solutions to existing data challenges within DBAR community inside the DBAR.



9 DBAR WGs and 3 ICoEs online for the data exercise

## Key Points: Data Legacy, Complimentary, Data Connections

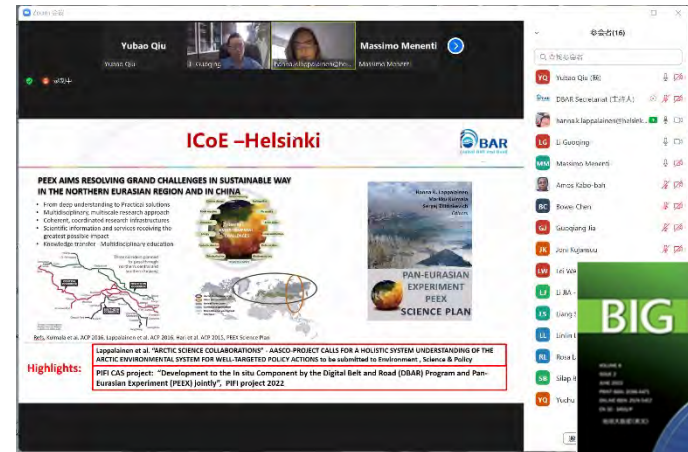


- It was acknowledged that DBAR has **rich data legacy** due continuous development of datasets over several years.
- It was observed that there are **complimentary datasets** from different working groups that can be linked with due acknowledgements
- It was suggested that **DBAR** can also collect data from around the world and archive for new studies.



## Key Points: External users, Algorithm, System, CBAS Connections

- An important issue highlighted was the lack of visible high-level DBAR thematic products for **external users**.
- It was also suggested that **algorithms** are important resource to share for flexibility of user to apply on desirable data.
- It was emphasized that DBAR **data processing system** should have clear users and operators and good resource support.
- It was advised that for any physical infrastructure DBAR can partner with CBAS for support.



Data and Algorithm



## Actions needed!

- Further development on the high quality data
- Further application to support the SDGs

# Example Talk by HiMAC Data Interoperability Portal

by Yubao Qiu

# Content

---

- **Interoperability**
- **Examples by CKAN**
- **HiMAC Data Portal**

# Interoperability

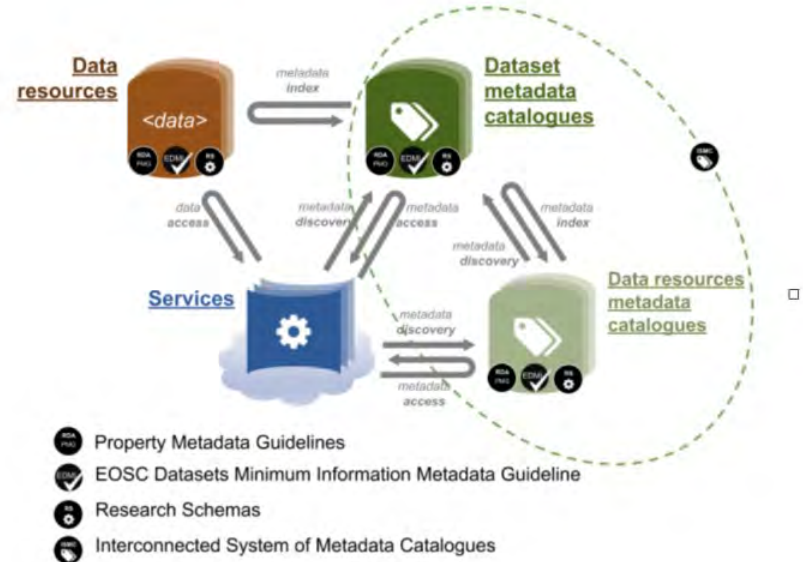
- Interoperability is the ability of **different systems, devices, applications or products** to connect and communicate in a **coordinated way, without effort from the end user**.
- Functions of interoperable components include **data access, data transmission** and **cross-organizational collaboration regardless of its developer or origin**.
- Similar to compatibility, interoperability helps organizations achieve higher efficiency and a more holistic view of information.



# Data interoperability

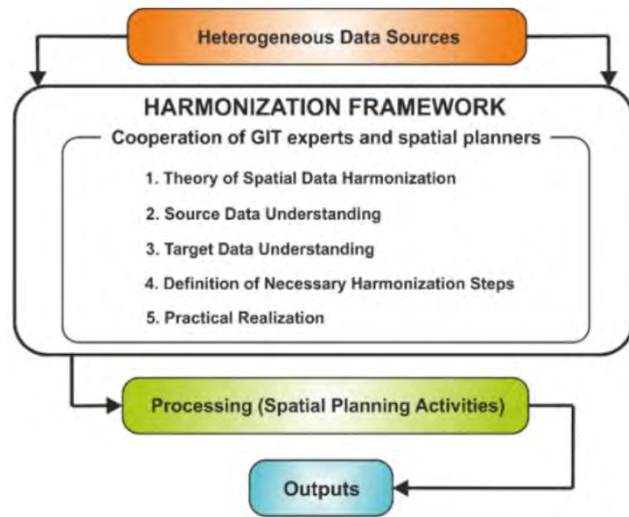
- Data interoperability is a problem affecting the **interaction of entities at very different levels**, and thus **not only the technical operations**.
- When merging two (or more) datasets into one common target dataset, we need to ensure that data from heterogeneous sources can be used in combination to view, query and analyse, or **in short: we need to harmonize the datasets**.

Therefore, we need to **set up a data harmonization process**.



# Data harmonization process

The description of **data harmonization process** itself can be described in many ways , this the following 5-step harmonization approach was selected in some papers.



- ① Understanding the **theory of spatial data harmonization** - understanding techniques which can be used for converting data between different data structures, while losing as little information as possible.
- ② **Source data understanding** - deep understanding of source data scheme up to the level of attributes.
- ③ **Target data understanding** - deep understanding of target data scheme up to the level of attributes.
- ④ Definition of harmonization steps - analysis of source and target data differences. Development of geometry and attribute matching scheme which describes the conversion of source data into target data scheme, layer by layer, table by table, attribute by attribute.
- ⑤ **Practical realization**

Implementation of the above-defined harmonization steps in a selected software

# Some software for increase interoperability

## Metadata Harmonization

**Socrata** is based on the RDF metadata (Dublin Core and DCAT) with enrichment from custom metadata fields.

**CKAN** stores the datasets as a folder that hosts datasets or resources. The metadata is served as RDF and the platform supports DCAT, Dublin Core and INSPIRE format.

	Fields	Description
CKAN	<b>Title</b>	Field used to label datasets. This attribute is intended to allow search, sharing and linking of datasets
	<b>Unique identifier</b>	This attribute assigns a unique URL to a dataset. This is one of the Dublin Core recommendations
	<b>Groups</b>	A customisable group that the dataset belongs to
	<b>Description</b>	Human readable description of the dataset
	<b>Data preview</b>	Quick preview in the comma separated value (CSV) format of the dataset
	<b>Revision history</b>	Provides revision history
	<b>Licence</b>	Allows user to check what licence a given dataset is
	<b>Tags</b>	Allocating tags to datasets makes them more discoverable through tag search and faceting by tags
	<b>Formats</b>	Provides information on the format datasets is available for download in
	<b>API key</b>	Allows for a developer access to the metadata fields
	<b>Customizable extra fields</b>	Such as location data or extra information relevant to the publisher or the dataset
		<b>Name</b>
<b>ID</b>		Unique identifier for the dataset
<b>Description</b>		The human-readable description of the asset
<b>Attribution</b>		The attribution of the dataset
<b>Type</b>		What sort of asset is described
<b>Updated at</b>		Timestamp
<b>Page views</b>		Set to provide statistics on page view of a dataset per day/week/month or all time
		array of column names in the dataset
		array of the descriptions matching the column name
		s serves as an identifier for columns and describes the field nes of columns
		egories are assigned using statistically derived models
		js are also assigned based on statistically derived models
		en by the owning domain
		ay of tags assigned to the dataset by the owning domain
		y' and 'value' of any custom metadata given to this asset by the ring domain

## Datatank:

The DataTank is an open source RESTful data management system managed by Open Knowledge Belgium. It is a web application where the administrator can register different datasets, which are published in various formats. The user can browse the available datasets, and view them in a preferred format.



**Open Knowledge Belgium**

Open Knowledge Belgium is a not-for-profit organization founded in 2012 dedicated to building tools, projects and communities that promote open knowledge in all its forms.





# Some software for increase interoperability

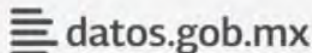
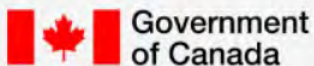


CKAN makes it easy to publish, share and work with data.  
It's a data management system that provides a powerful platform for **cataloging, storing and accessing datasets with a rich front-end, full API** (for both data and catalog), visualization tools and more.

## Key advantages of CKAN

- Is open-source and widely supported by the developer community
- Features extensive and comprehensive documentation
- Allows deep customization of its features
- Can be fully under institutions control
- Supports unrestricted (non standards-compliant) metadata
- Has faceted search with fuzzy-matching
- Records datasets change logs and versioning information

Powering open data at:



# Examples by CKAN

## Science Data Business Demand

data sources, multiple data

- (2) Use advanced geospatial and visualization feature of CKAN for research needs.
- (3) Research teams usually develop several plugins and customizations to meet the needs of researchers.

## Government Business Demand

Government of Canada. Everyone information they want on the

- (2) uses CKAN to make open data portal which can be anything from municipal infrastructure data to socio-economic composition.
- (3) Using Filestore features of CKAN to create news module to blogs, events, etc.

## Enterprise Business Demand

CKAN to manage data or files as enterprise operation data ,

plans, reports and insights.

## International Organizations Business Demand

International



### • Geospatial



### • Visualization



### • Access Datasets by Topic



12 datasets found

Commercial gas amounts

Entry/exit gas quality

Gas composition and quality for transmission

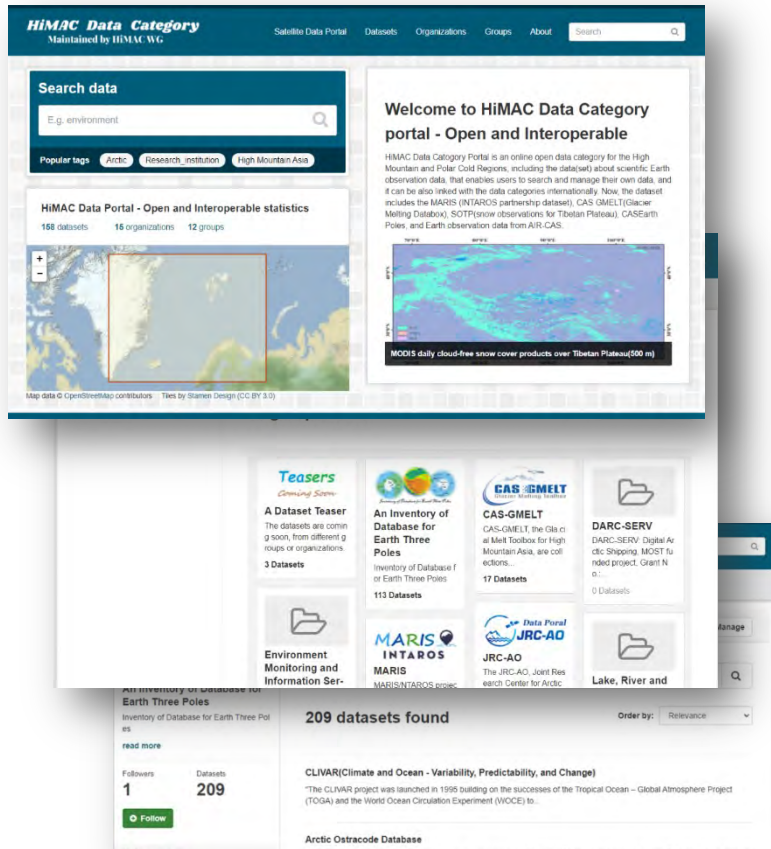
Gas composition and quality for transmission, monthly

HiMAC – a Work Group for the Digital Belt and Road (DBAR) Program



- To build **research and observation network** of High Mountain and Cold Regions (HiMAC);
- To build the HiMAC **Big Earth Data information service** platform;
- To promote **capacity building**, and advance the **applications and services** of HiMAC-Big Earth Data;
- To develop an **indicator system** for the **Sustainable Development for the HiMAC region**, and provide consultant reports for the "B&R" areas.

# HiMAC Data Category



The screenshot shows the HiMAC Data Category website. At the top, it says "HiMAC Data Category Maintained by HIMAC WG". Below this is a navigation bar with "Satellite Data Portal", "Datasets", "Organizations", "Groups", and "About". A search bar is also present. The main content area features a "Search data" section with a search box containing "E.g. environment" and "Popular tags" for "Arctic", "Research\_institution", and "High Mountain Asia". There are two maps: one showing the "HIMAC Data Portal - Open and Interoperable statistics" with 158 datasets, 16 organizations, and 12 groups, and another showing "MODIS daily cloud-free snow cover products over Tibetan Plateau(500 m)". Below the maps are several "Teasers" for datasets: "A Dataset Teaser" (3 Datasets), "An Inventory of Database for Earth Three Poles" (113 Datasets), "CAS-GMELT" (17 Datasets), "DARC-SERV" (0 Datasets), "Environment Monitoring and Information Ser...", "MARIS INTAROS" (113 Datasets), "JRC-AO" (0 Datasets), and "Lake, River and...". At the bottom, a search result for "Inventory of Database for Earth Three Poles" is shown, indicating "209 datasets found" and "209 Followers".

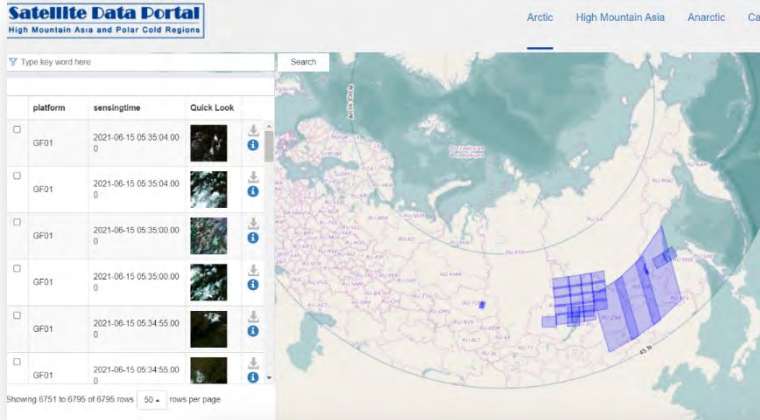
<http://115.29.142.79/group/inventory-of-database-for-earth-three-poles>

HiMAC is an online open data category system for the High Mountain and Polar Cold Regions, especially the data(set) about scientific data of Earth Three Poles, and other cold regions, and mountain areas, that enables users to search and manage their own data and also link with the data categories internationally.

The dataset includes:

- ✓ CAS GMELT(Glacier Melting Databox)
- ✓ SOTP(snow observations for Tibetan Plateau)
- ✓ CASEarth Poles
- ✓ MARIS (INTAROS partnership dataset)
- ✓ Earth observation data from the AIR-CAS
- ✓ International datasets of collaboration.







# HiMAC Satellite Data



Satellite Data Portal  
High Mountain Asia and Polar Cold Regions

Arctic High Mountain Asia Antarctic Ca

Type key word here Search

platform	sensingtime	Quick Look
GF01	2021-06-15 05:35:04.00	
GF01	2021-06-15 05:35:04.00	
GF01	2021-06-15 05:35:00.00	
GF01	2021-06-15 05:35:00.00	
GF01	2021-06-15 05:34:55.00	
GF01	2021-06-15 05:34:55.00	

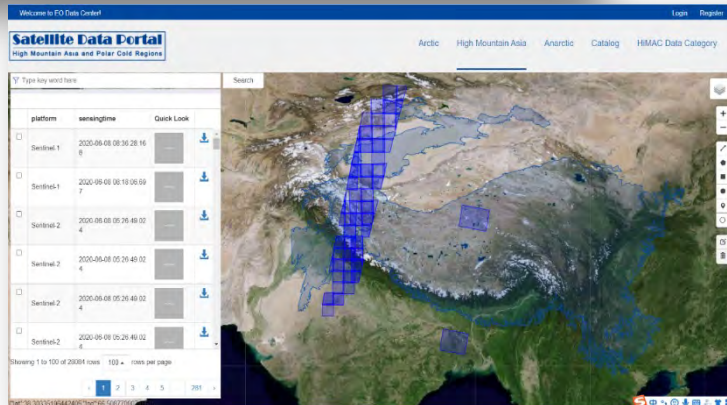
Showing 6751 to 6755 of 6795 rows 50 rows per page

## Chinese Satellite Data

Gaofen Series for the Sub-Arctic Region and HMA

## European Satellite Data

ESA Sentinel Data for the Arctic Ocean









Welcome to EO Data Center! Login Register

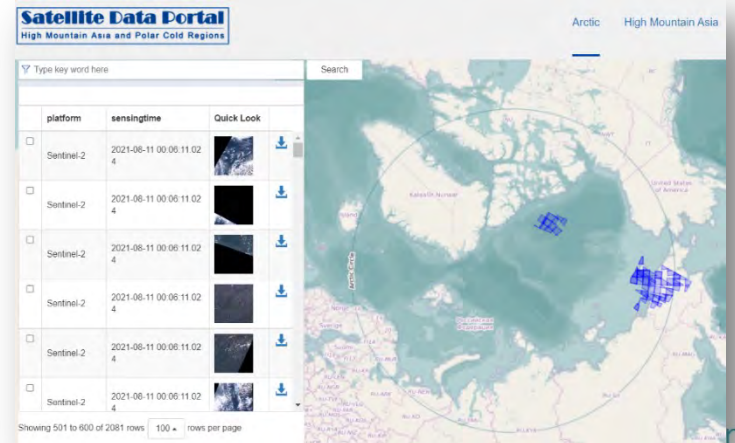
Satellite Data Portal  
High Mountain Asia and Polar Cold Regions

Arctic High Mountain Asia Antarctic Catalog HiMAC Data Category

Type key word here Search

platform	sensingtime	Quick Look
Sentinel-1	2020-05-08 08:36:28.16	
Sentinel-1	2020-05-08 08:10:36.09	
Sentinel-2	2020-05-08 05:26:49.02	
Sentinel-2	2020-05-08 05:10:49.02	
Sentinel-2	2020-05-08 05:26:49.02	
Sentinel-2	2020-05-08 05:29:49.02	







Showing 1 to 100 of 20881 rows 100 rows per page



Satellite Data Portal  
High Mountain Asia and Polar Cold Regions

Arctic High Mountain Asia

Type key word here Search

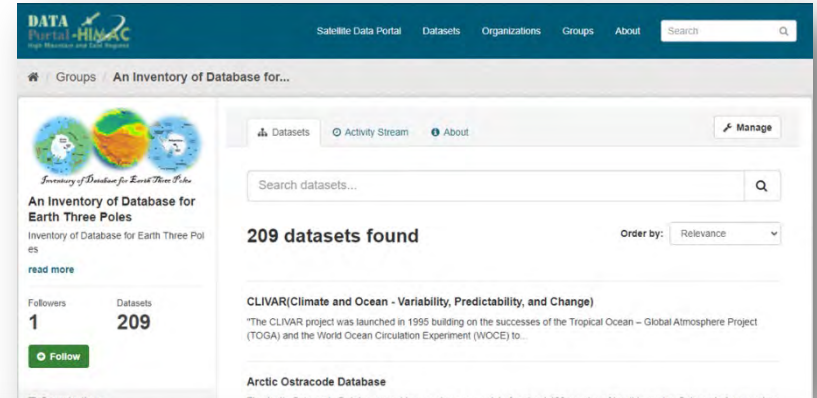
platform	sensingtime	Quick Look
Sentinel-2	2021-06-11 00:06:11.02	
Sentinel-2	2021-06-11 00:06:11.02	
Sentinel-2	2021-06-11 00:06:11.02	
Sentinel-2	2021-06-11 00:06:11.02	
Sentinel-2	2021-06-11 00:06:11.02	
Sentinel-2	2021-06-11 00:06:11.02	

Showing 501 to 600 of 2081 rows 100 rows per page

# Science and Development to HiMAC

## 1. An Inventory of Database for the Earth Three Poles

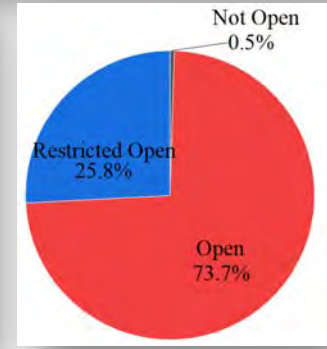
- 209 datasets found in 22 countries
- Survey on the data policy, sharing principle and data management
- Mainly in USA (42), Norway (12), China (10), Canada (10) and Denmark (10), etc.
- Far way to an opening data world



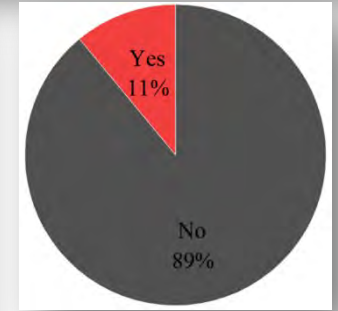
Basic Info		Organizer	Data	Standard
Name	Host Name	Data Policy	DOI	
URL	URL	Sharing Principle	Data Standard	
Description	Country	Data Type	...	
Keywords	...	Openness		
Region		FAIR		
Theme		...		
...				



Statistics for databases affiliated countries and organization



Data openness



FAIR (findable, accessible, interoperable, reusable)



---

# Thanks

**DBAR Secretariat**

No.9 Dengzhuang South Road  
Haidian District, Beijing 100094, China

Tel: +86 10 82178980

Fax: +86 10 82178959

[www.dbeltroad.org](http://www.dbeltroad.org)

