

Justine Woo<sup>1</sup>, Nathan Pollack<sup>1</sup>, Taylor Yates<sup>1</sup>, Benjamin Patrick Michael<sup>2</sup>

<sup>1</sup>Science Systems and Applications INC./ NASA Goddard Space Flight Center, Code 61A, Greenbelt, MD, USA ([Justine.y.woo@nasa.gov](mailto:Justine.y.woo@nasa.gov))

<sup>2</sup>NASA Goddard Space Flight Center, Code 61A, Greenbelt, MD, USA

## ABSTRACT

The Crustal Dynamics Data Information System (CDDIS) is one of twelve NASA Earth Observing System Data and Information System (EOSDIS) Distributed Active Archive Centers (DAAC) and supports the space geodesy and geodynamics community through the International Association of Geodesy (IAG) services. As an EOSDIS DAAC, the CDDIS is required to meet best archival practices including the Findability, Accessibility, Interoperability, and Reuse (FAIR) Guiding Principles for scientific data management and stewardship which ultimately serves data and product providers and users. This poster will introduce some common tools to ensure clarity for users on how these best practices support their efforts and how to utilize and find the tools built to support them. These include ensuring data is findable via the use of Digital Object Identifiers (DOIs) and facilitating discovery using the Earthdata Search API and NASA's Earthdata Search. As the CDDIS continues to grow during NASA's year of open science, the CDDIS will start migrating to the cloud further ensuring clarity and accessibility of data.

## THE CDDIS AND GEODESY DATA/PRODUCTS

### Techniques:

- Global Navigation Satellite Systems (GNSS)
- Laser ranging
  - Artificial satellites (SLR)
  - Lunar (LLR)
- Very Long Baseline Interferometry (VLBI)
- Doppler Orbitography and Radio-positioning Integrated by Satellite (DORIS)

### Data

- Observation data
- Broadcast orbit information
- Meteorological parameters

### Products

- Precise orbits and orbit predictions
- Clock products
- Reference frames
- Troposphere and ionosphere products
- Displacement time series and Plate boundary aseismic transients
- Water storage time series

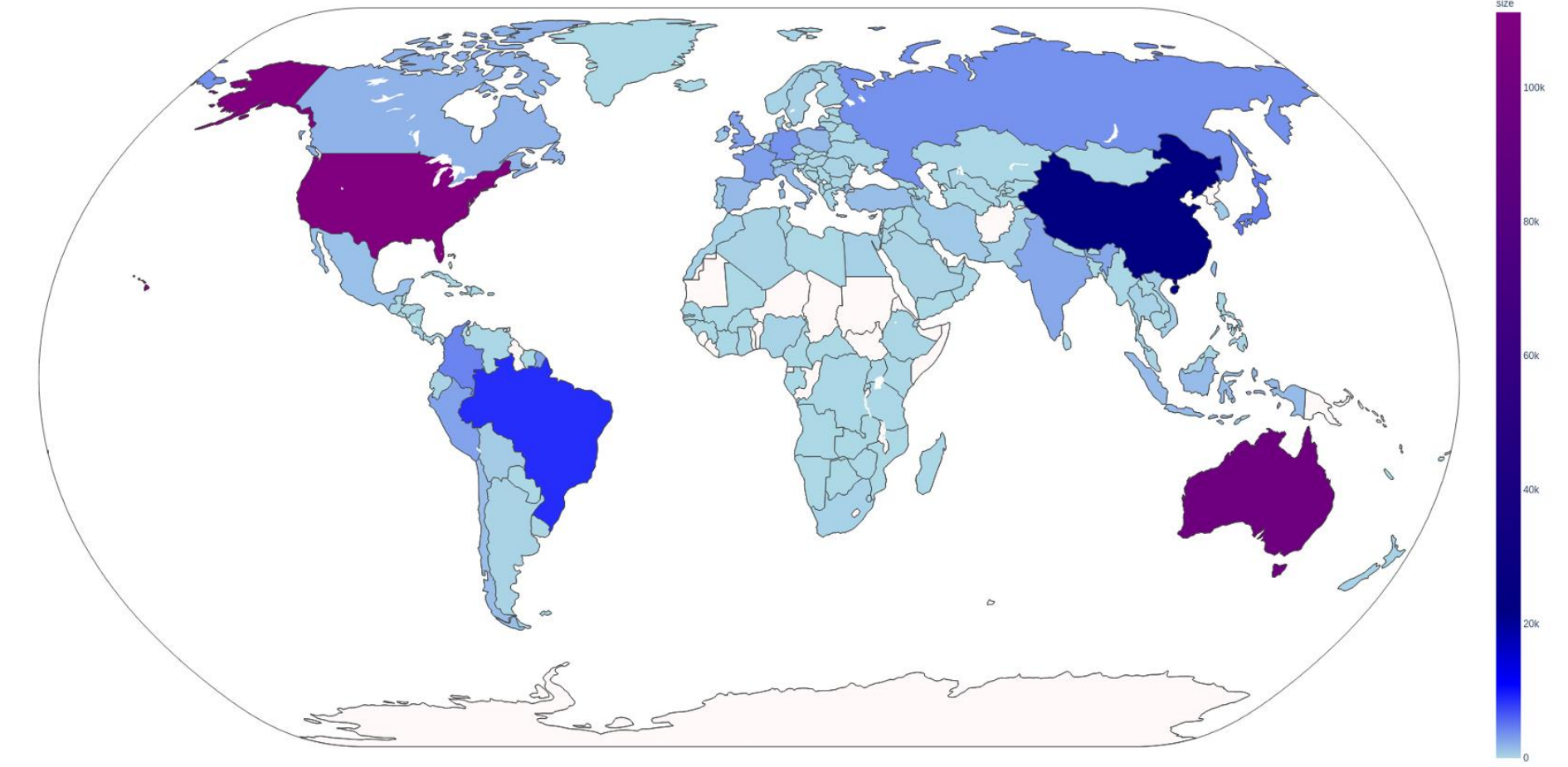


Figure 1: Number of Distinct Users of the CDDIS Archive By Country

## FINDABLE

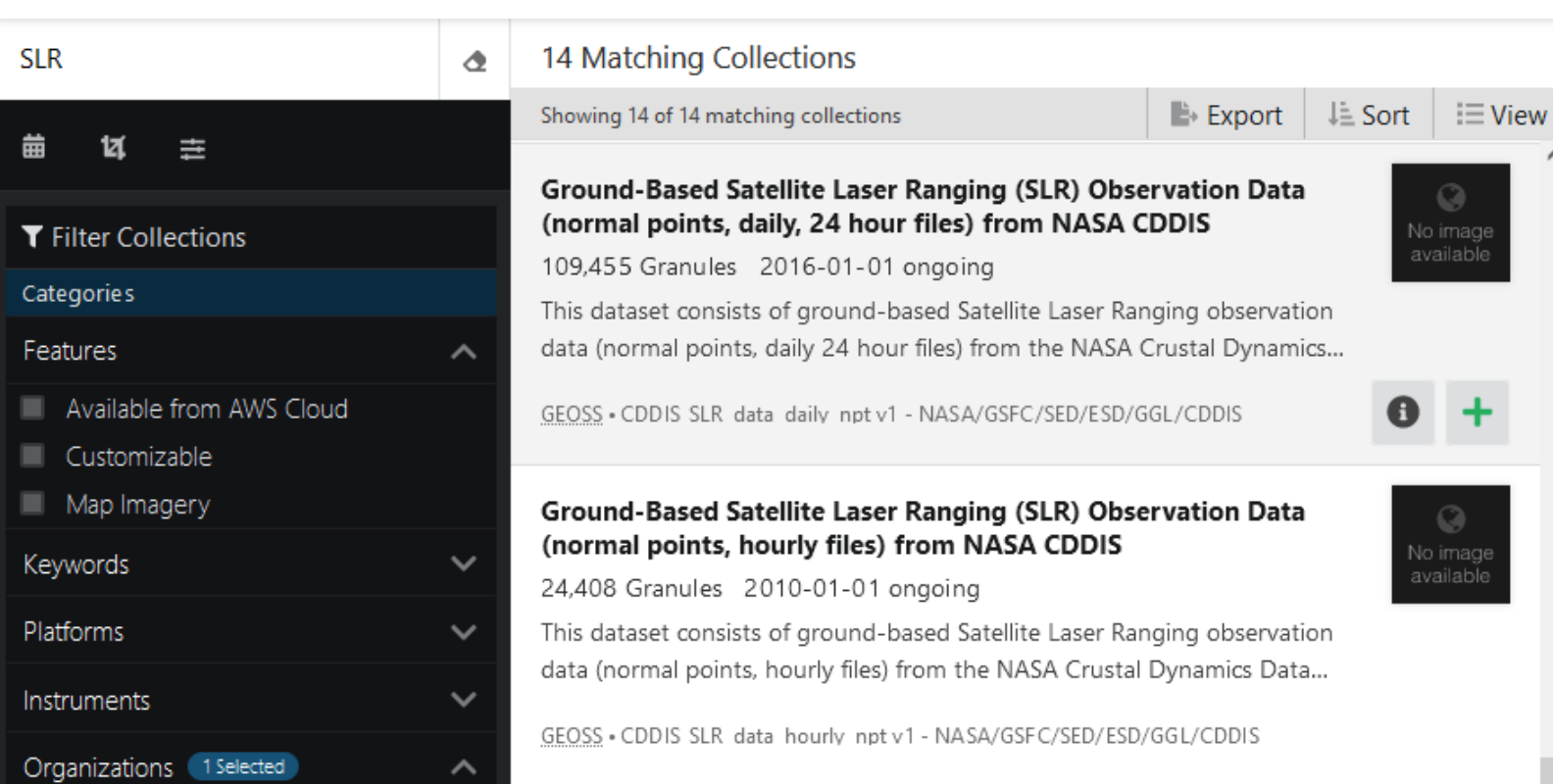


Figure 2: Earthdata Search – CDDIS SLR Collections

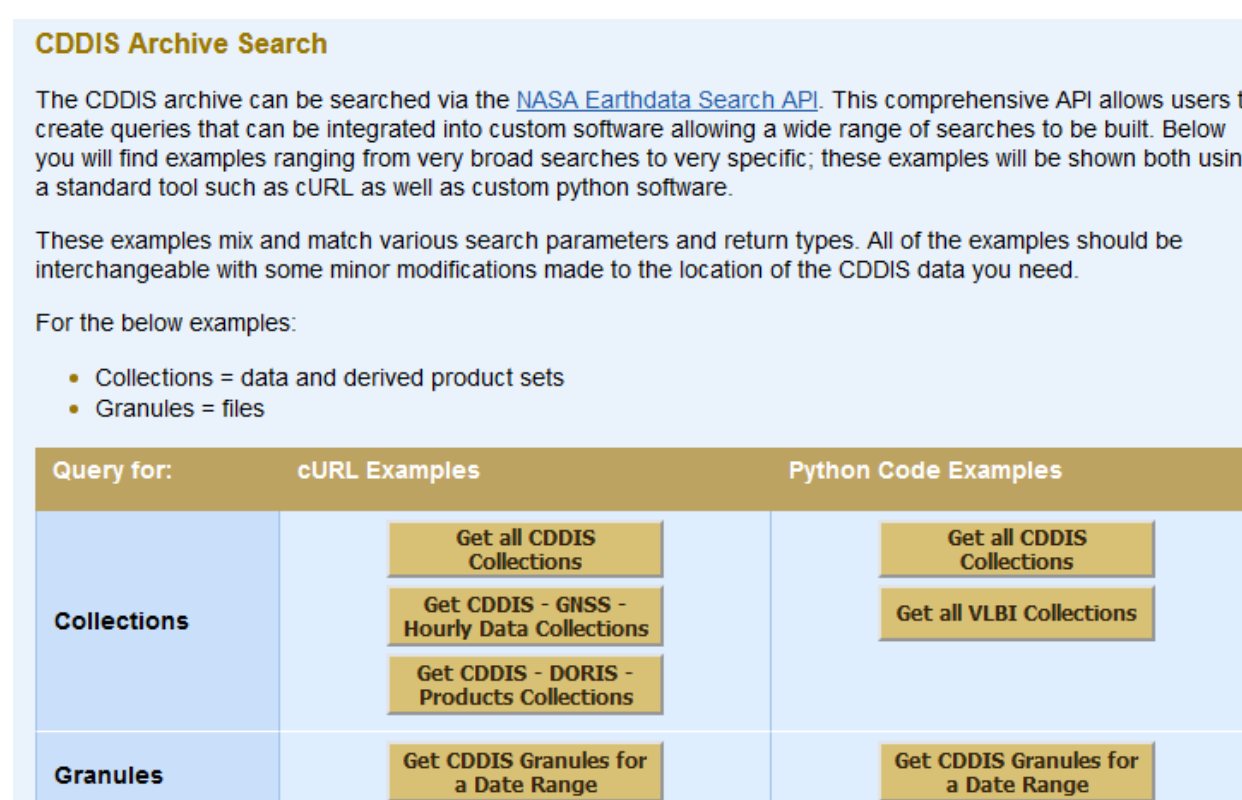


Figure 3: CDDIS Website Archive Search

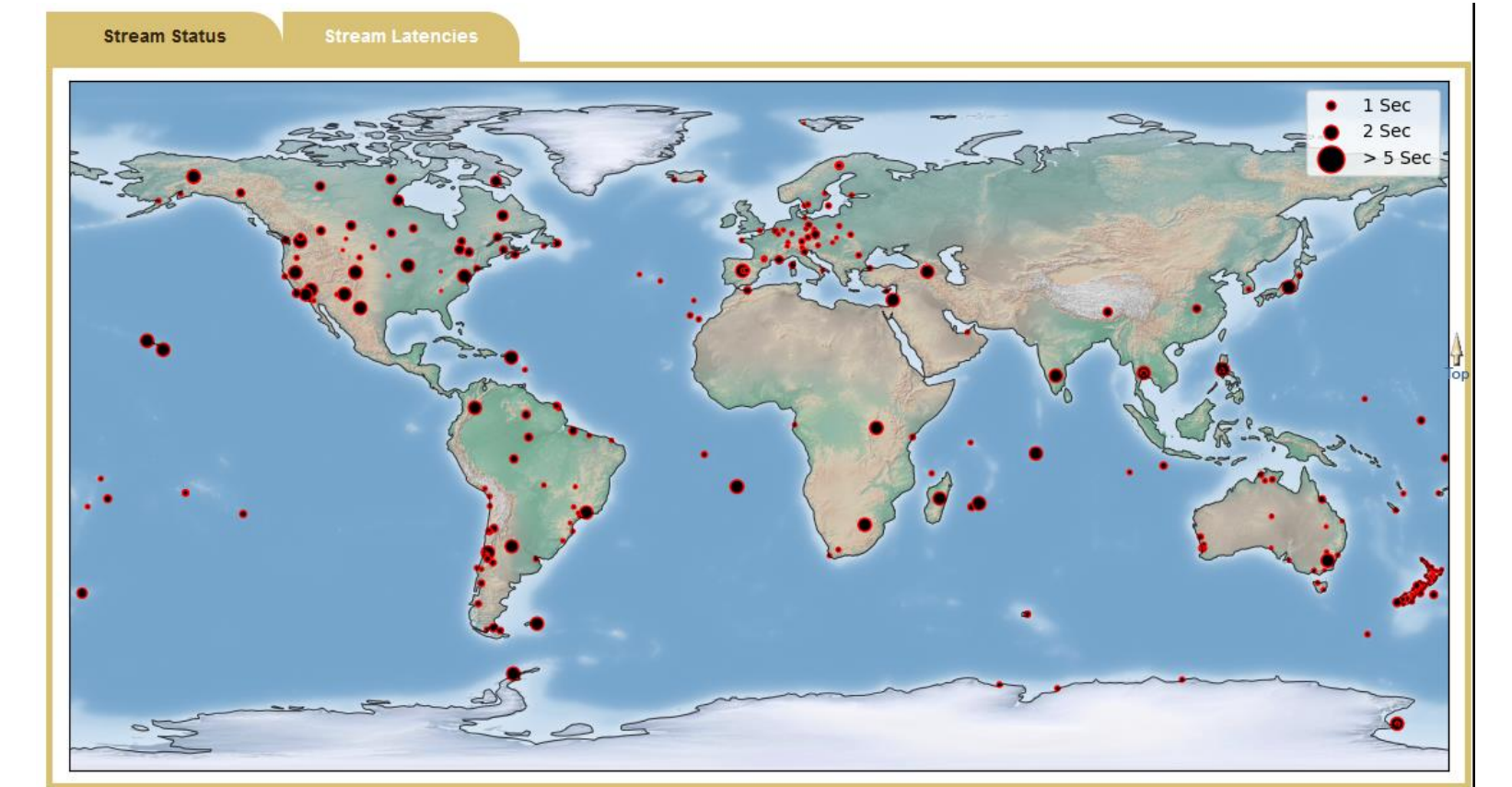


Figure 4: CDDIS Website GNSS Real-time Data Streaming Latency Status

### New Users: Earthdata Search

Earthdata Search provides a visual tool for scientists and engineers to search for data available from all the DAACs. The CDDIS currently has more than 270 collections which users can find by narrowing down the "Organization" to the CDDIS. For more information on how to narrow down your search, please see the Interoperable section.



### General Use: The CDDIS Website and Earthdata Search API

The CDDIS website contains information on the space geodesy data it has available, broken down by technique. Collection information is under the "Data and Products" tab. This website is useful for those who know that they need geodesy data but need more information.



The Earthdata Search API also allows users to get a listing of the CDDIS collections and granules. To simplify the process for users, the CDDIS website has sample queries and guidance.

### Specific Use: Real-time (NTRIP) System Status

The CDDIS has been supporting GNSS real-time data streaming since 2018 via the CDDIS-Caster NTRIP broadcaster. This year, to provide additional usability and clarity to users, the CDDIS created a real-time map showing which streams are active and their latencies.



## ACCESSIBLE

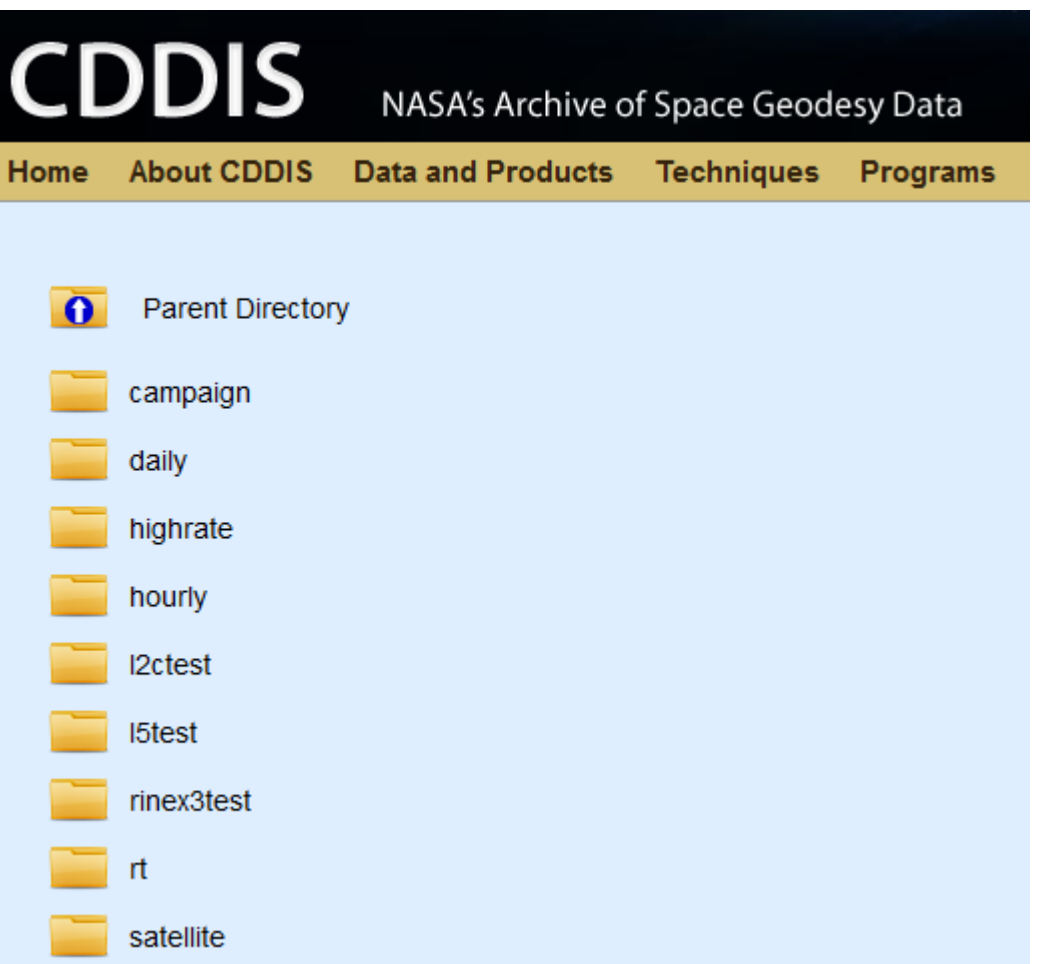
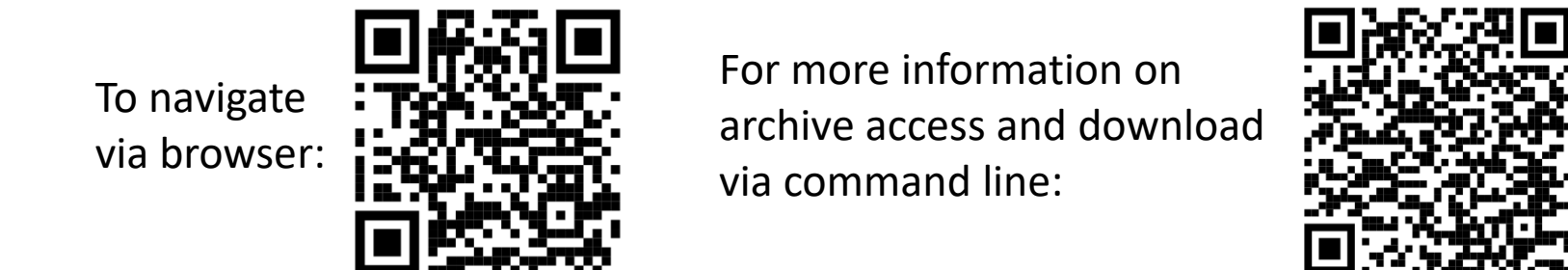


Figure 5: CDDIS GNSS Archive via Web Browser

On October 2020, the CDDIS shifted away from FTP downloads and transitioned to HTTPS which supports human and computer readability. Access requires an Earthdata account which can be created for free. The archive can be accessed via a browser and command line.



## INTEROPERABLE

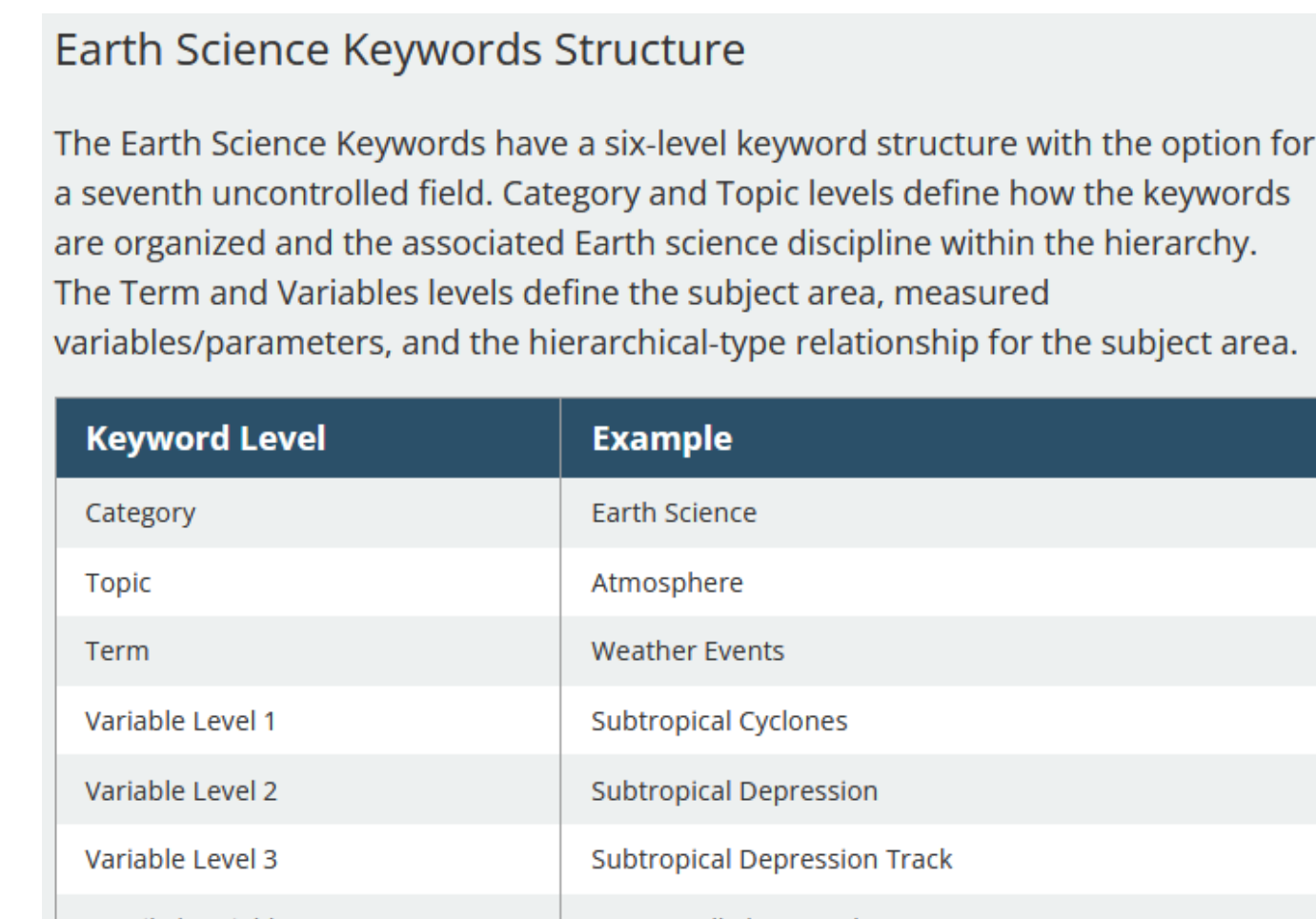


Figure 6: Screenshot from the GCMD Website

To ensure interoperability with other DAACs, the CDDIS follows the Global Change Master Directory (GCMD) structure and keywords to provide insight into its data and products. This allows for clearer connections

between different DAACs and their data/products and is primarily used for Earthdata Search.



## REUSABLE

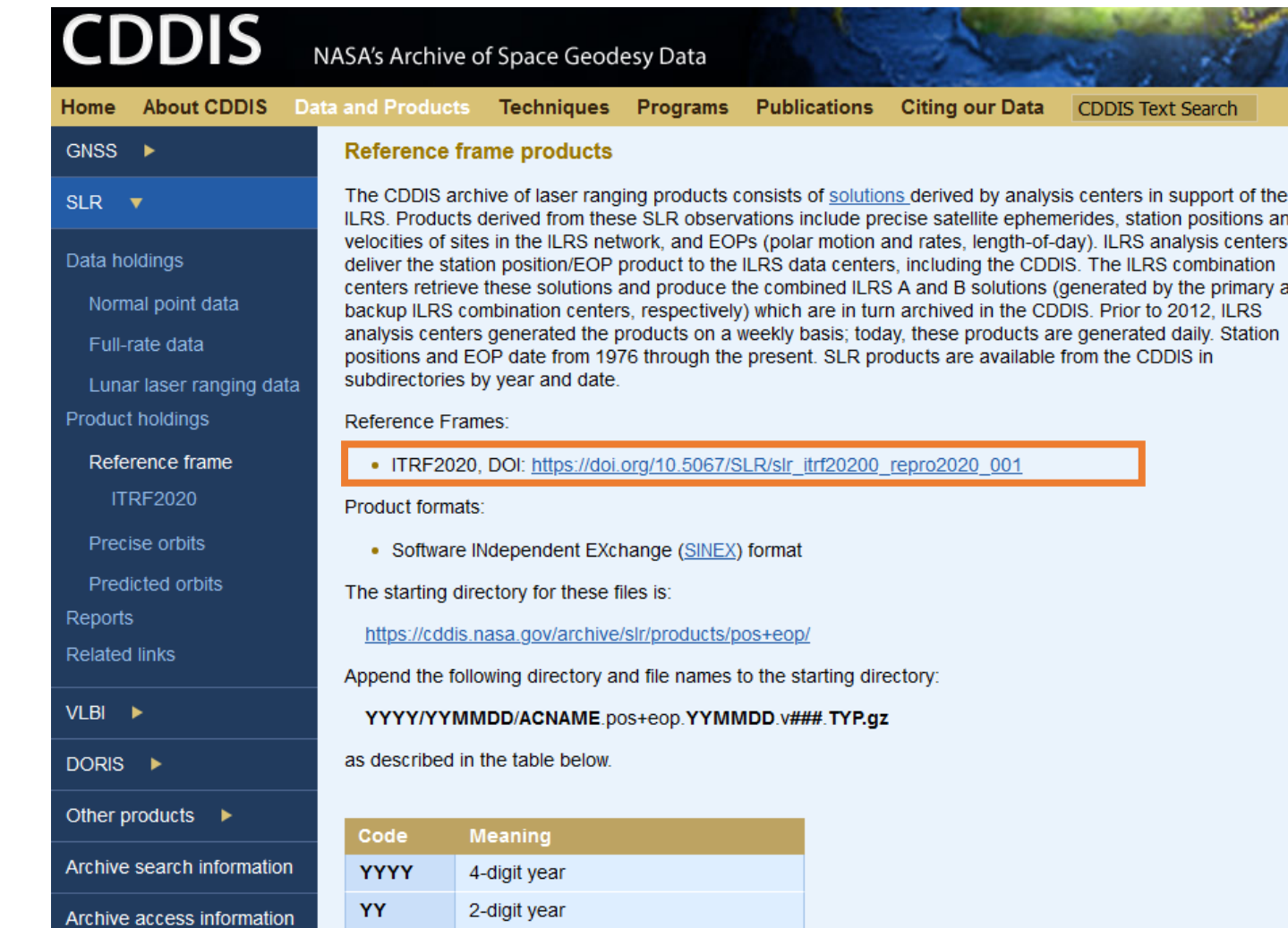


Figure 7: CDDIS Website DOI Example

The CDDIS ensures the reusability of its data and products by ensuring they can be cited via Digital Object Identifiers (DOIs). DOIs can be added to references in papers and publications to ensure traceability. The DOIs can be found on the CDDIS website by selecting the "Data and Products" tab, followed by the specific data or product you hope to cite.

The link will either lead directly to a landing page with the DOI or the DOI link will be available which leads to the landing page. If you have any issues finding an applicable DOI please contact the CDDIS.

## DATA DESCRIPTION FILES (DDF)

```
# CDDIS file definition Version 1.0 of 2019.10.22
#
# 2019.07.16 Leonid Petrov created
# 2019.08.19 Dirk Behrend updated
# 2020.08.07 Dirk Behrend changed Format_file_name location
Short_description: 24-hour master file
Long_description: Master file for 24-hour sessions
Format_file_name: /vlbi/ivsformats/master-format.txt
Reference: n/a
DOI: n/a
FileNaming_scheme: master@date@{y}.txt
File_location: /vlbi/ivscontrol
Product_ID: VLBI_AUX_MASTER_24
Data_type: VLBI
Data_content_type: Misc
Data_format: ascii
Validate_proc: master.py
Magic: ## Master file format version 1.0
Compression_type: none
2001.08.21 CCT&NRV
```

Figure 8: Sample DDF

In 2020, the CDDIS collaborated with the International VLBI Service (IVS) to create a new processing structure for VLBI data. The software architecture emphasizes a modular design maintained using DDFs with quality checking software provided by the IVS. A sample DDF for processing VLBI master files is given in Figure 8.

The file contains all the information needed to identify and process the file, including some of the metadata needed for findability.

## PROCESSING V4 – PRE-CLOUD

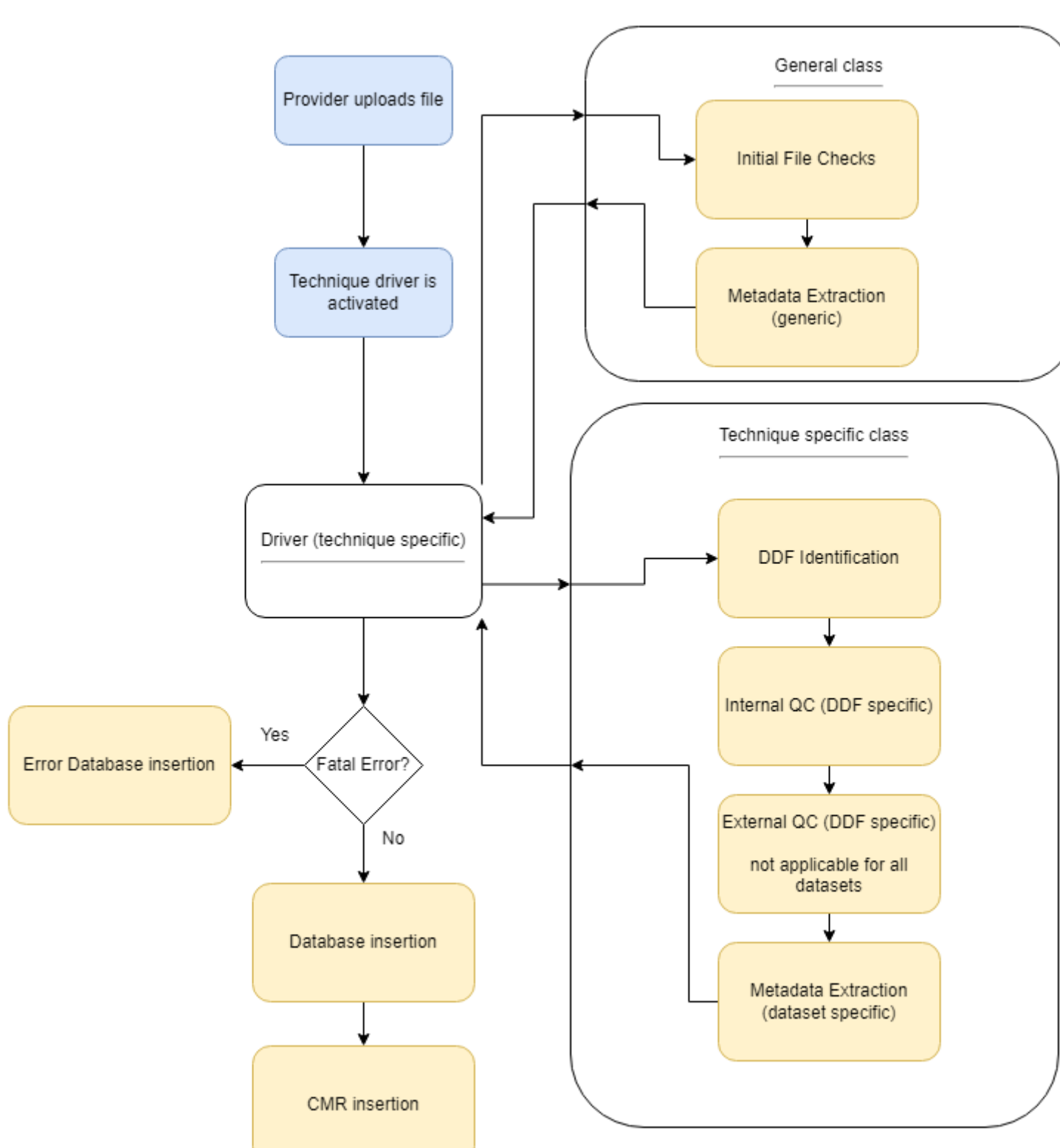


Figure 9: Ingest Processing Flow Chart

Given the simplicity and effectiveness of DDFs, the CDDIS has been working to transition its processing software to this format for all the techniques it supports.

The new architecture also considers the CDDIS's future transition to the Earthdata Cloud by structuring its software to fit AWS services.

The new architecture have the following improvements that providers and users will experience:

- Automatic processing after upload; processing will no longer be run at discrete time intervals
- Faster availability on Earthdata Search
- Simplified process for adding or removing data or products from the processing stream

## OPEN SCIENCE - EARTHDATA CLOUD TRANSITION

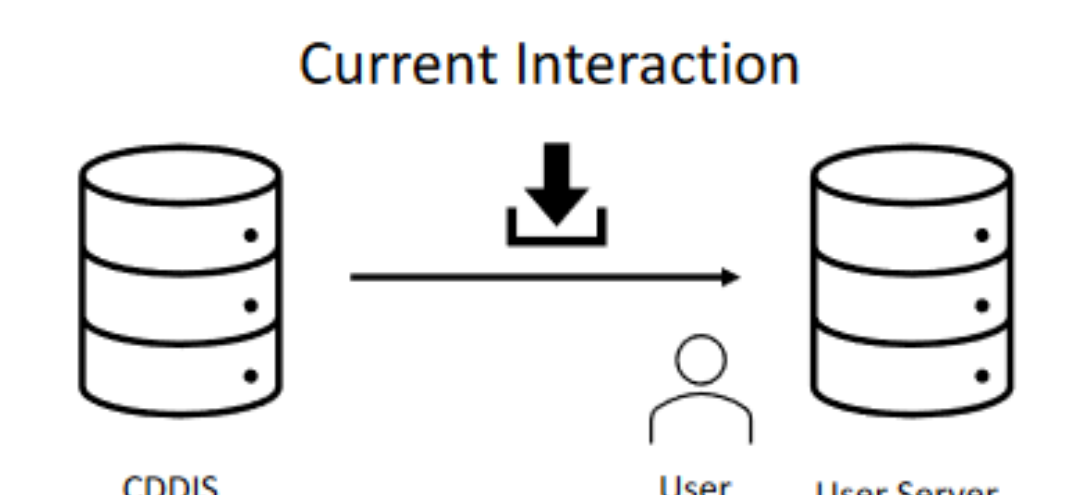


Figure 10: Current User Interaction with the CDDIS Server

The CDDIS is transitioning to the Earthdata Cloud in order to fully realize NASA's goals for open science. The primary focus of these efforts is to ensure the accessibility of NASA data by reducing the barriers to entry within the science community while increasing traceability and reproducibility of research.

Once the CDDIS has shifted to the Earthdata Cloud, users will have access to the following benefits:

- Access to full features on Earthdata Search including searchability and download directly from Earthdata
- Ability to work with the CDDIS data and products in the cloud – downloads are no longer required, reducing the need to invest in on-premises resources
- Easier identification of collocated data from across the NASA DAACs
- Access to new tools for subscribing and analyzing data

\* Please note that, even as the CDDIS transitions to the cloud, users will still be able to download data to their servers and the original archive structure will still be maintained. EOSDIS is also investing in helping scientists transition to the cloud through training initiatives

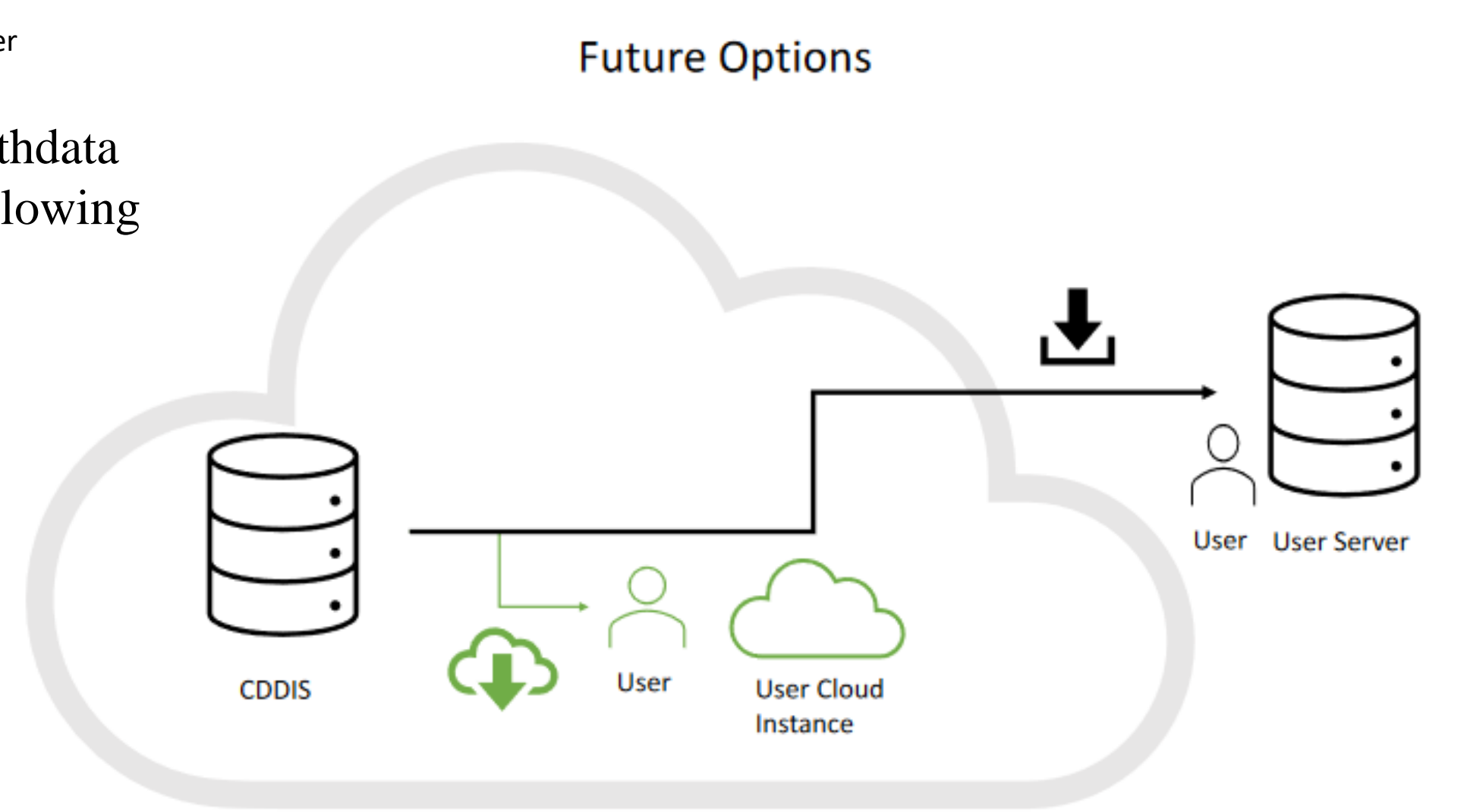


Figure 10: Future Options for User Interaction with the CDDIS Server

- The CDDIS is also working to ensuring the following as data and products are migrated to the cloud:
- Ensuring the data and products have a DOI and landing page, especially for older data sets
  - Reorganize the data/products into clearer collections with the input from representatives of the techniques supported
  - More thorough QC is created for each of the file types
  - Data can be downloaded based on temporal specifications