EOL Data Stewardship

Brooks Snyder

NCAR/EOL/DMS
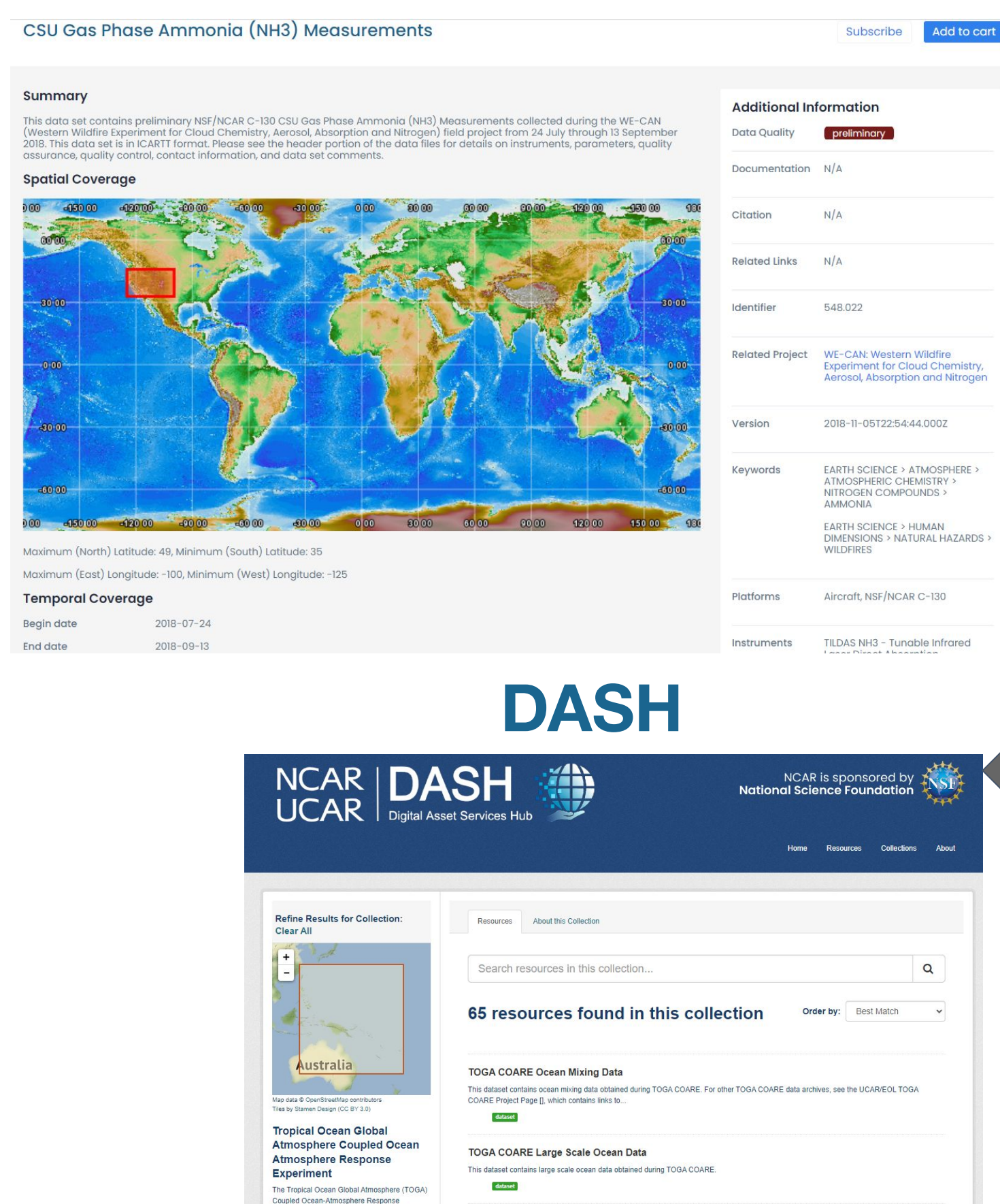
https://data.eol.ucar.edu
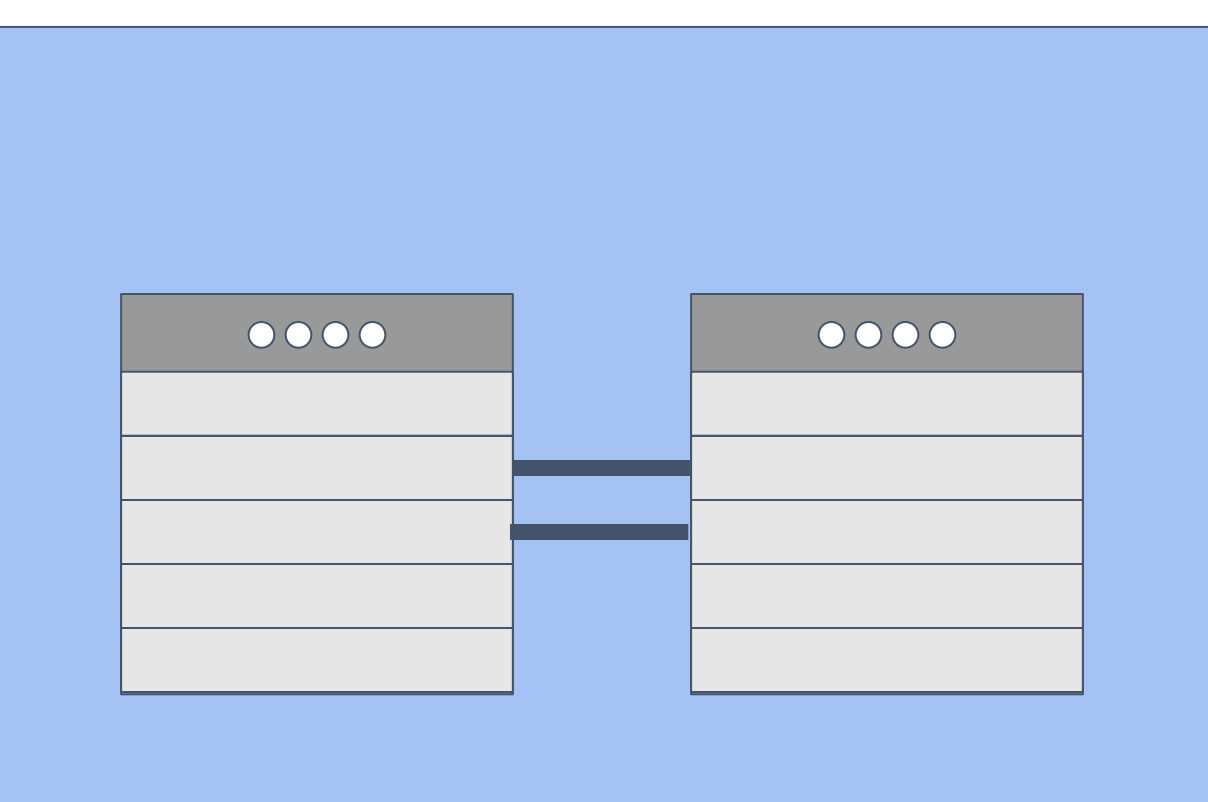
Data Archive Systems

Field Data Archive Web Interface
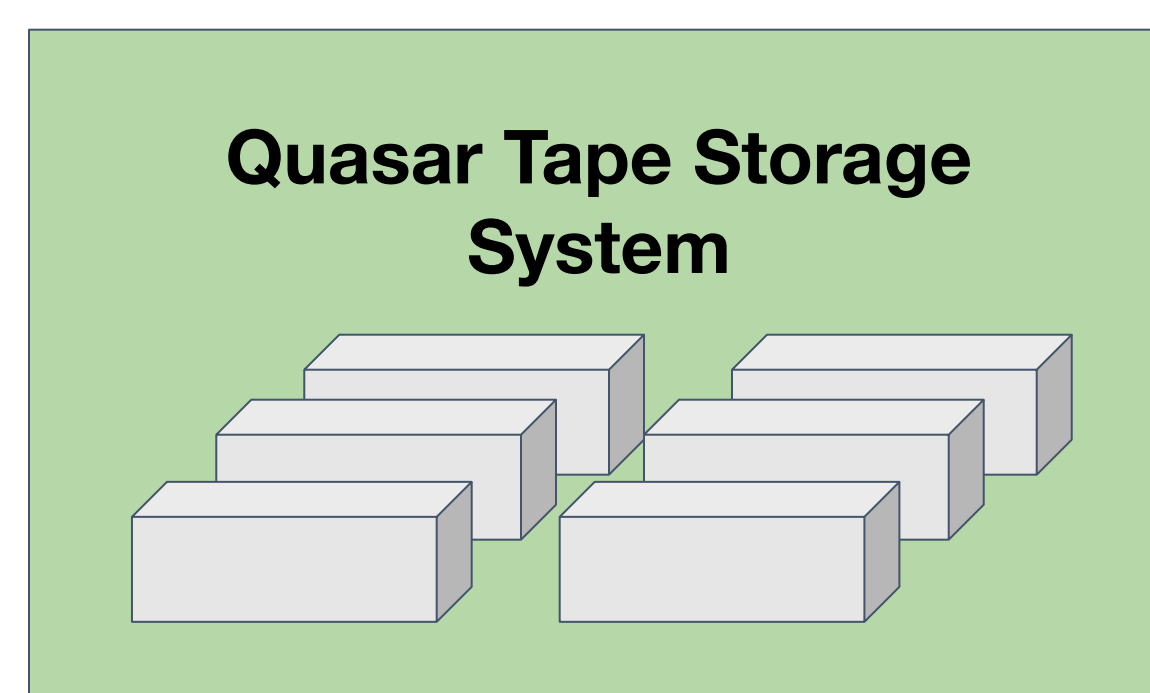
CSU Gas Phase Ammonia (NH3) Measurements

DASH

Metadata Database

Data Discovery

Disk Archives

Local EOL Archive

File metadata loaded to database

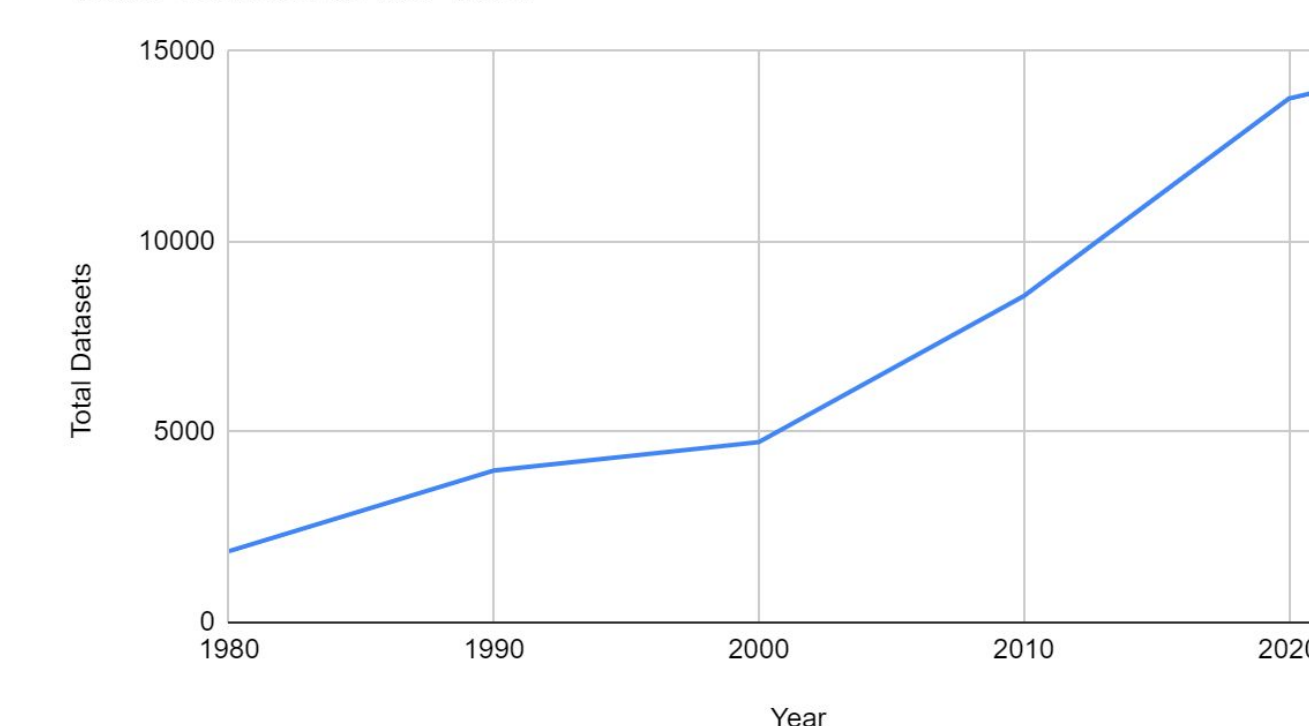**Campaign Storage Archive**

Quasar Tape Storage System

Quasar Backup Utility

Daily process that bundles and transfers new files to backup tape archive

EOL Archive Stats

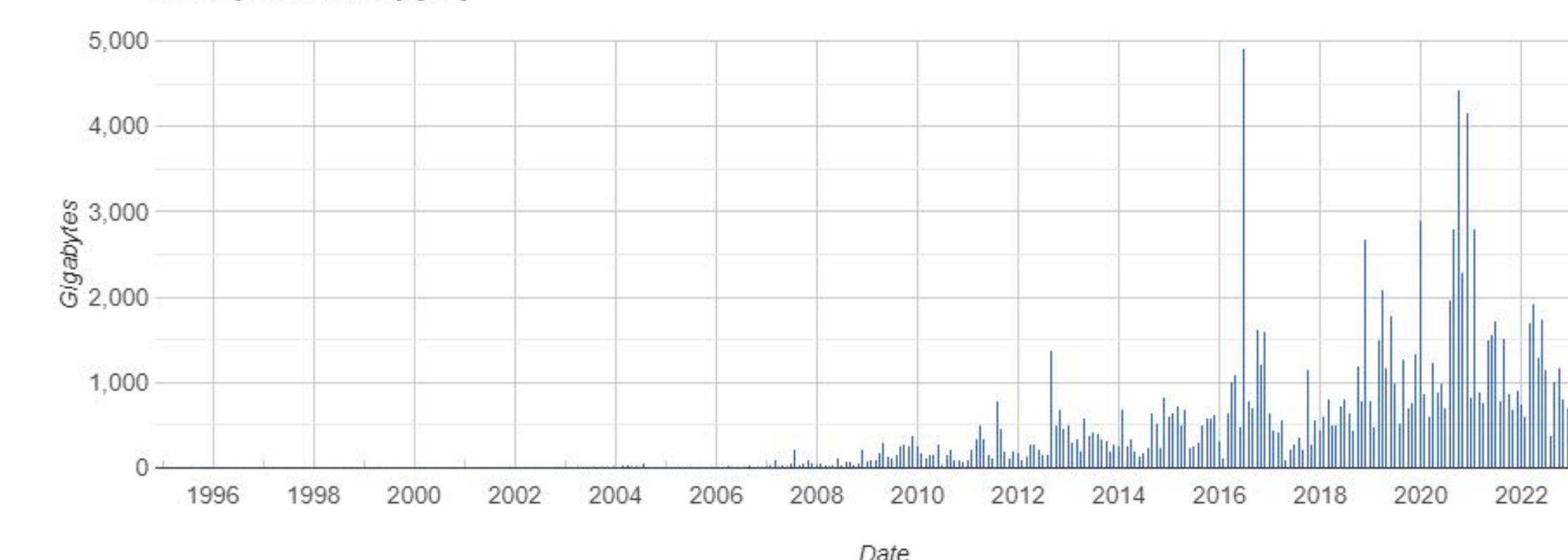### Archive Growth

Total Datasets vs. Year

Archive Size

As of August 14, 2023 the EOL Field Data Archive has:

- **108,726,790 total files**
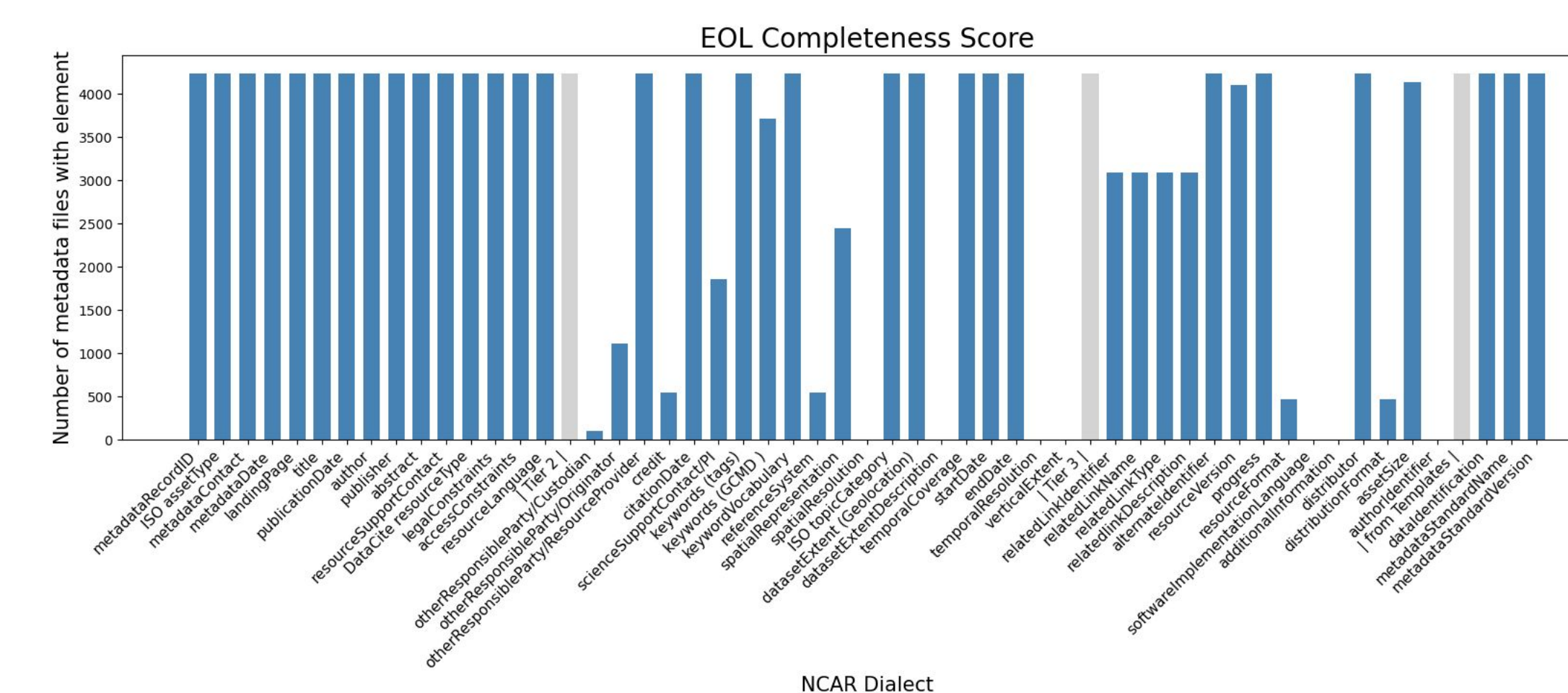- Totaling **549 TB**
- 3,058 physical media (tapes, etc.)

Data Orders

- To date, the EOL Field Data Archive has had **166,864 data orders**, totaling **135.5 TB**.
- In 2022 alone **6,736 orders** were delivered, totaling **13.5 TB**.

Monthly Data Delivery [GB]

Completeness of XML Metadata Exported to DASH

EOL Completeness Score

NCAR Dialect

Data Backup

Local EOL Archive

- EOL's local archive is hosted on an NFS server gemini7 and synced twice daily to gemini6.
- Backups will be made to CISL-managed hardware in the future.

Campaign Storage Archive

- Campaign Storage and Quasar are CISL-managed hardware.
- Campaign Storage archive uses clustered RAID technology.
- EOL's Campaign Storage Archive is backed up to the Quasar tape storage system.
- EOL developed a Quasar Backup Utility to package and transfer newly archived files to Quasar each day for backup. Backup metadata are tracked in EOL's database.

Disaster Recovery 🔥🌪️

Files backed up to the Quasar tape system are stored on tapes which, when full, are removed and placed in a fire-proof safe.
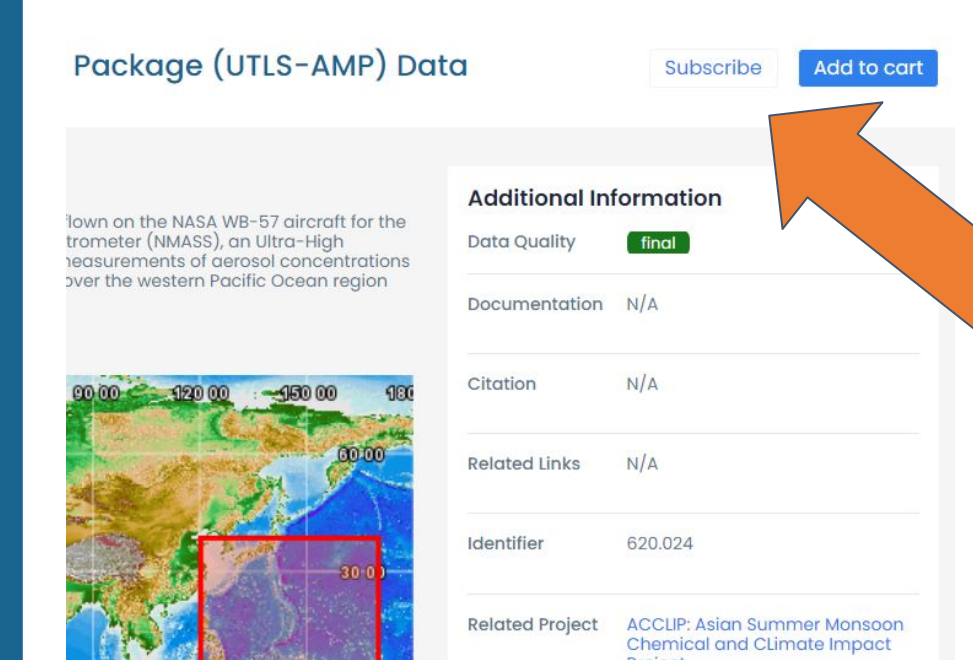
EOL's Data

EOL's Field Data Archive (FDA) is unique in that the focus is on **observational data** derived from field projects that take place all over the world.

Datasets are collected using a mix of **research aircraft**, **airborne instrumentation**, and **ground-based systems** and are currently associated with over **600 field projects.** Ongoing real-time collections of **operational datasets** compliment the field-collected data.

Data Integrity

EOL is committed to long-term stewardship and does not plan to ever delete any data that has been archived.

In order to ensure data fixity EOL has begun an initiative to maintain checksums for all data files in the archive, and uses a custom-built Checksum Utility to facilitate this objective.

Package (UTLS-AMP) Data

The data archive interface offers the option to **subscribe** to datasets so that users can receive email notifications when important changes are made or new versions are released.