GOSC: What users' need? Radar Case Study Group

Incoherent scatter radar data fusion and computation case study

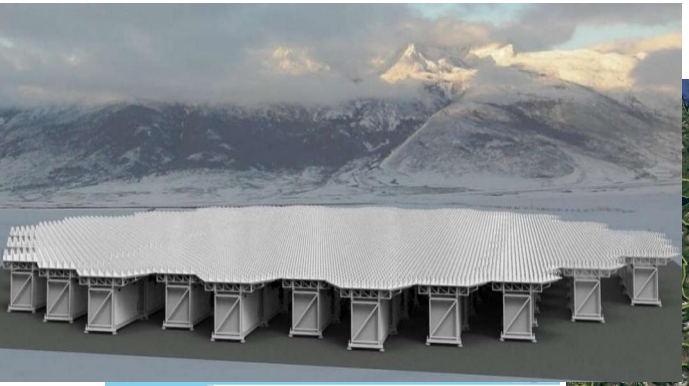
Co-chairs: Ingemar Häggström, EISCAT, Sweden
Xinan Yue, IGGCAS, China,
Secretary: Yin Chen, EGI Foundation, Netherlands
22 September 2022

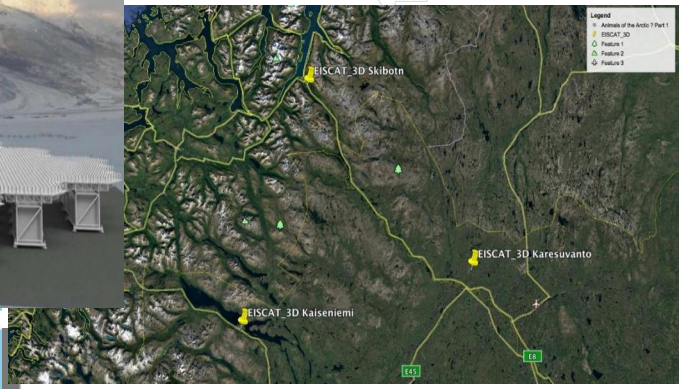
The work *is partly funded by the European Commission under H2020 Framework Programme*

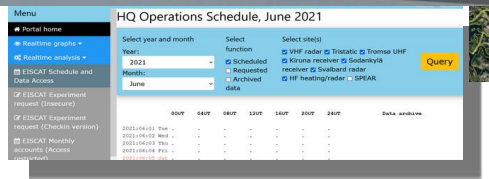
GOSC Case 1 Space Physics

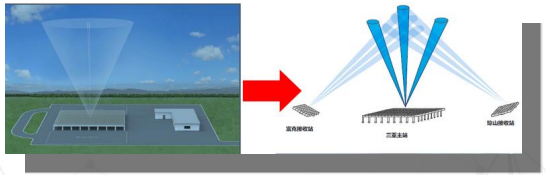
Incoherent scatter radar data fusion and computation

EISCAT_3D radar, EISCAT Scientific Association, Sweden
Sanya Incoherent Scatter Radar (SYISR), Institute of Geology and Geophysics, China



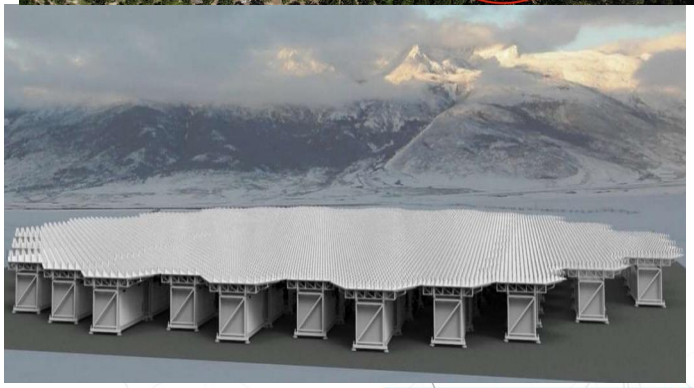






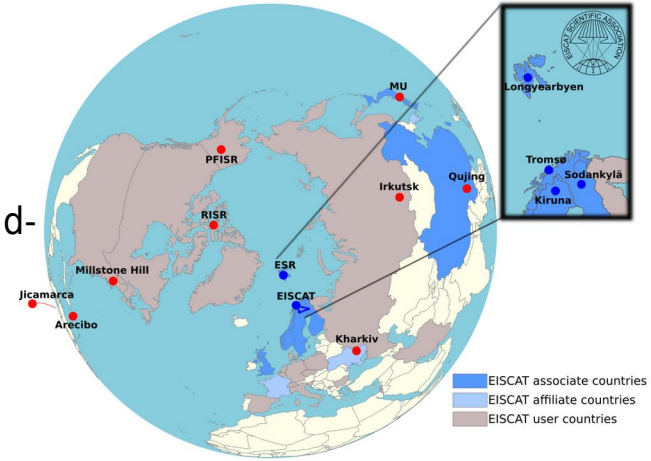
EISCAT/CAS use case

- Next generation of incoherent scatter radar systems
  - HPLA HighPowerLargeAperture
- Similar hardware
  - Systems of distributed phased array radars
    - Enable comprehensive three-dimensional observations of the atmosphere and ionosphere.
  - Increased temporal and spatial resolution
  - Continuous measurement capabilities
    - Inclusion of detailed incoherent scatter radar data into climate and Earth system modeling.
  - Large VVVV
    - Velocity, Volume, Versatility, Variability



Use Case Mission

- Objectives
  - To explore **data and metadata models** for collaborative production and sharing.
  - To develop **enhanced algorithms**, methods for cloud-based radar data processing and federated data learning.
  - Other issues to address for **data interoperability**.



- Key deliverables
  - Exploration of technical solutions for EISCAT and SYISR **(Meta)data federation**.
  - Deployment of GOSC technical solutions supporting the cloud **federated data processing** and **on-demand data movement**.
  - Recommendations for analogous research infrastructures in OSCs.

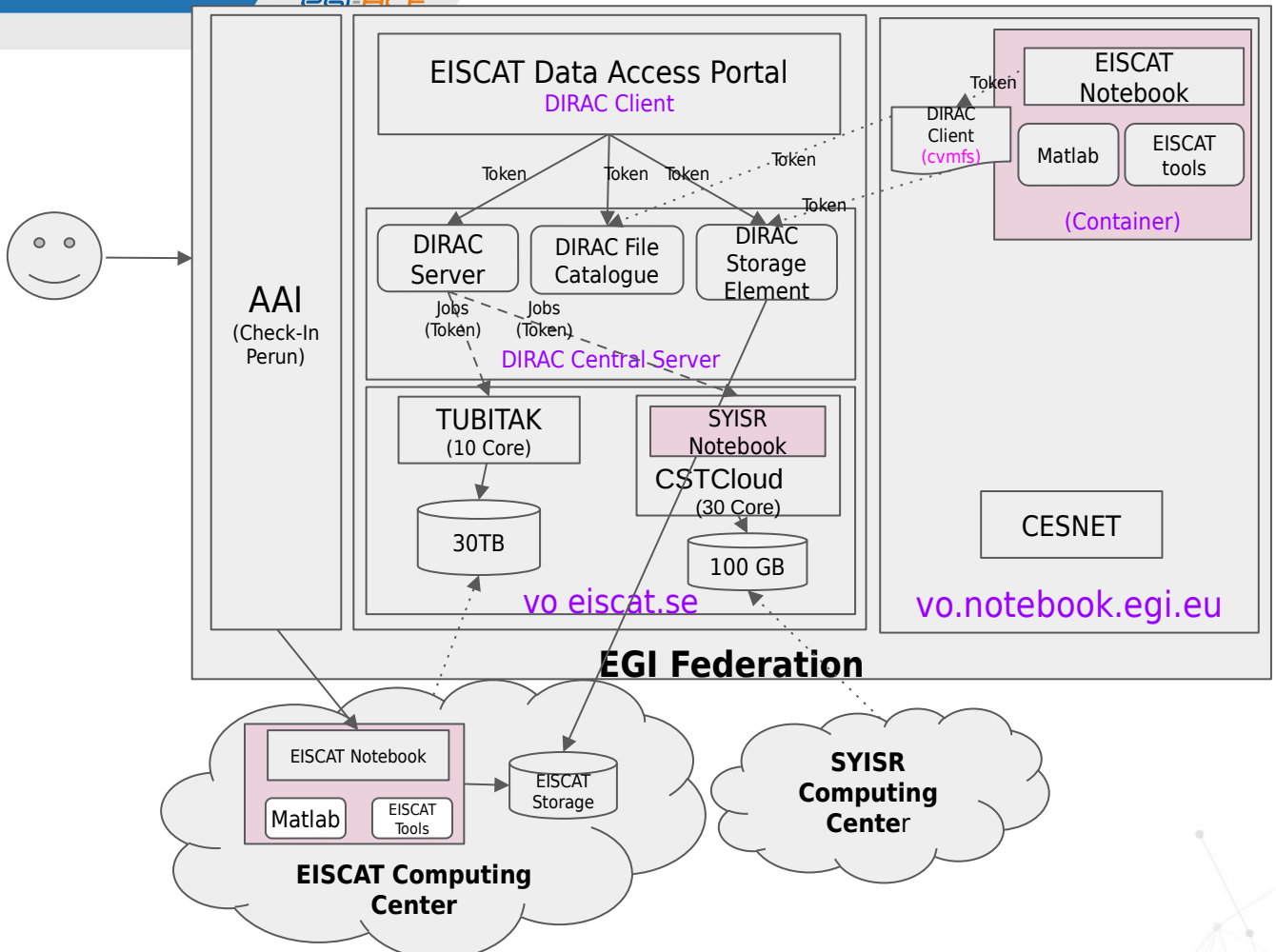
Challenges

- **Data interoperability**
  - Convert the different data to a radar independent format
    - Wrapper routines
    - Processing in each others analysis software
    - Harmonize on-site DSP routines
- **E-Infrastructure interoperability**
  - Across continental data sharing and transmission on Cloud
- **Policies issues**
  - Resource provision need to align with national policies, funding objectives, computing centers' own interests
  - Embargo period is applied to data access
- **Governance & sustainability issues**
  - Funding to sustain the development
  - Long-term resource and technical support

VO eiscat.se @ EGI

- **Cloud resources**:
  - Turkish TUBITAK(10 core CPU VM, 30TB)
  - CNIC/CSTCloud (30core CPU VM, 100GB, later can be up to 2TB)
- **AAI**
  - User login via check-in and Perun
  - User registration is approved by EISCAT admin in Perum
- **DIRAC-based EISCAT Data Access Portal**
  - Metadata-based data search and download, job submission for data processing
  - AAI (Check-In, Perun)
  - DIRAC Server – access to Cloud (TUBITAK, CSTCloud) for job submission
  - DIRAC File Catalogue – EISCAT Metadata, EISCAT User access control based on EISCAT data access policy
  - DIRAC Storage Element – enable search for EISCAT data in EISCAT's own storage
  - Notebook users should be able to access storage for data processing (DIRAC SE, FTS, …)

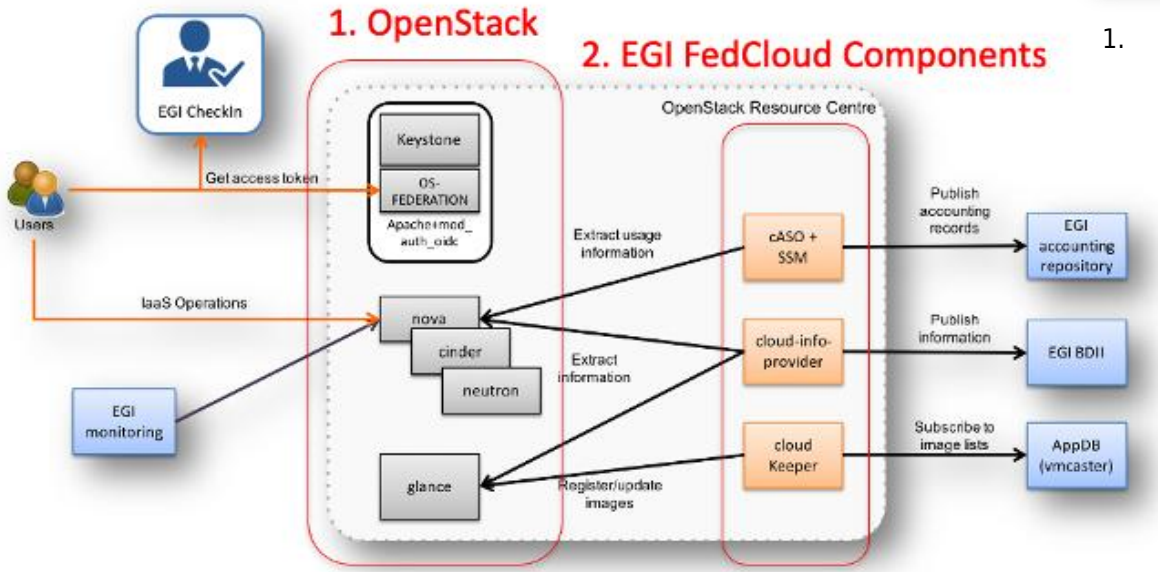
System Setup

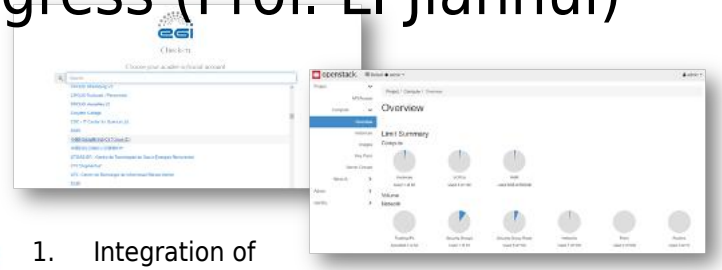


- **EISCAT Data Access Portal (**DIRAC)
  - User access via Check-In/Perun
  - Access token is passed to DIRAC File Catalogue (by DIRAC Client)
  - DIRAC File Catalogue return token with user information
  - (DIRAC Client) access DIRAC Storage Element with token
  - DIRAC Storage element enables search of EISCAT data (in the EISCAT storage)
  - DIRAC Server submits Jobs to Cloud
    - TUBITAK (10core, 30TB) + CSTCloud (30Core, 100GB)
    - Token access to the Cloud is ongoing
- **EISCAT Compute Center**
  - Jupyter notebook
    - User access via Perun
    - Matlab enabled
    - Not yet access to DIRAC Storage Element – cannot perform search, and have lower performance
  - EISCAT Storage: EISCAT data
- **EISCAT Notebook Container**
  - Enable EISCAT user access
  - Notebook+DIRAC image
    - dirac.egi.ei cvmfs repository is configured alongside eiscat cvmfs repository (eiscat.egi.eu)
    - DIRAC client is available via cvmfs
    - pass user access token to the DIRAC File Catalogue to get user information
    - Access the DIRAC Storage Element to search and retrieve data from the EISCAT storage
  - Will include Matlab
    - Token to use EISCAT license?
  - Will include EISCAT Tools

Achievements and progress (Prof. LI Jianhui)

Installation procedure





1. Integration of EGI Check-In

2. . Installation of **OpenStack** and **EGI FedCloud Virtual Appliance**

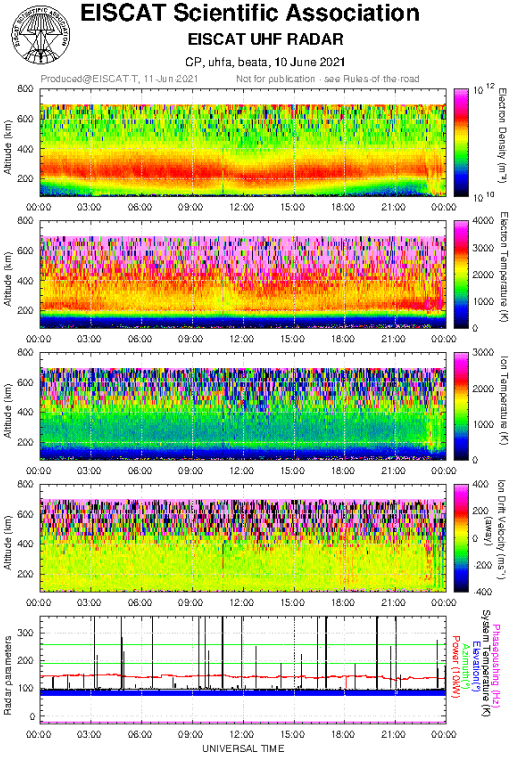
3. Capture by EGI Monitoring

4. Configuration of VO: **EISCAT.SE**

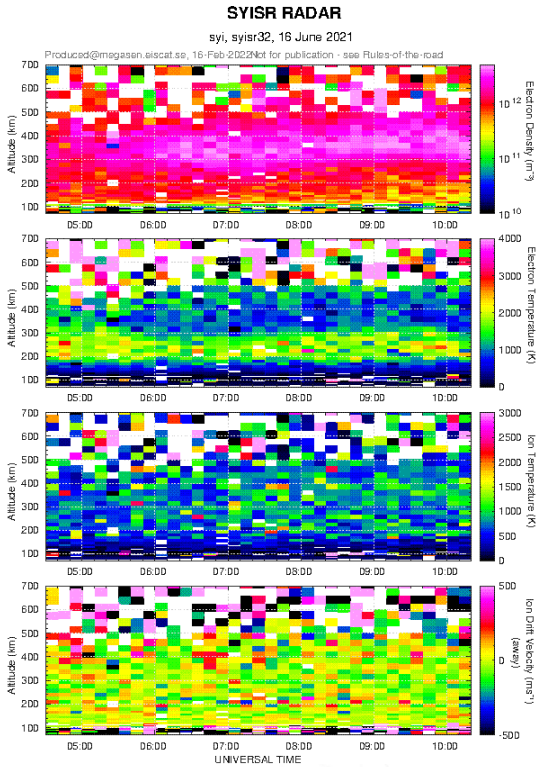
Integration of China's CSTCloud with EGI FedCloud is completed (Joint activities with GOSC TI WG)

Analysis interoperability

- **Analysis (GUISDAP)**
  - Internally radar independent
  - EISCAT wrapper
    - Single beam setup
  - SYISR wrapper
    - Multi beam setup
      - Test also for EISCAT_3D



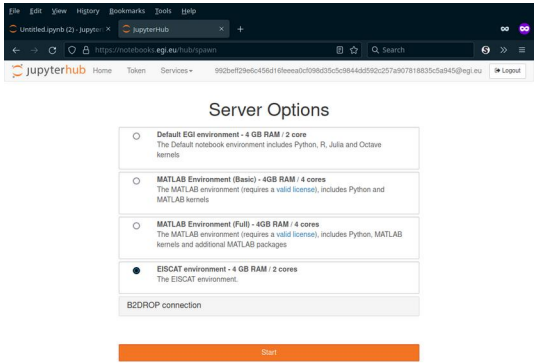
EISCAT analysis results



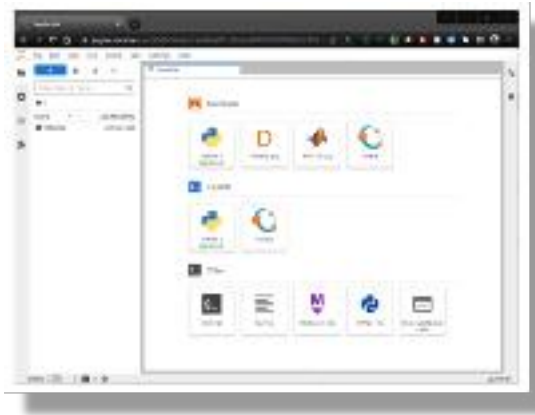
SYISR analysis results

Jupyter notebooks

1. Login by EGI Check-In

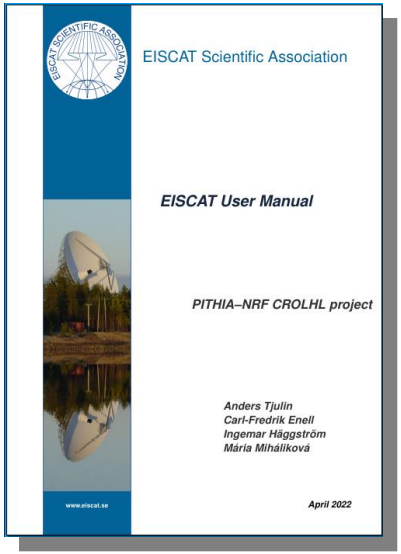


2. Select server



3. EISCAT Jupyter Notebook landing page

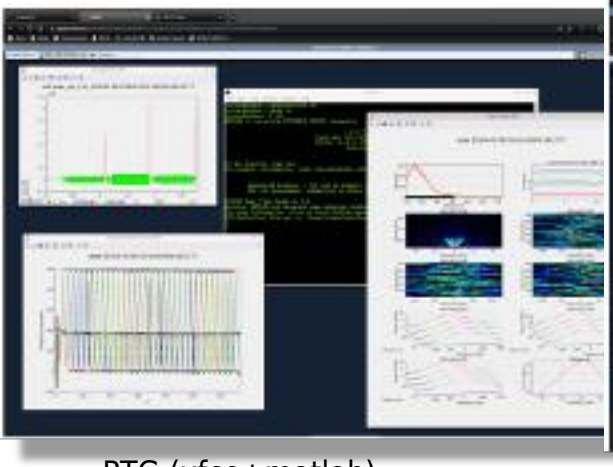
•Jupyter Notebook Desktop workspace with EISCAT GUISDAP application analysis running



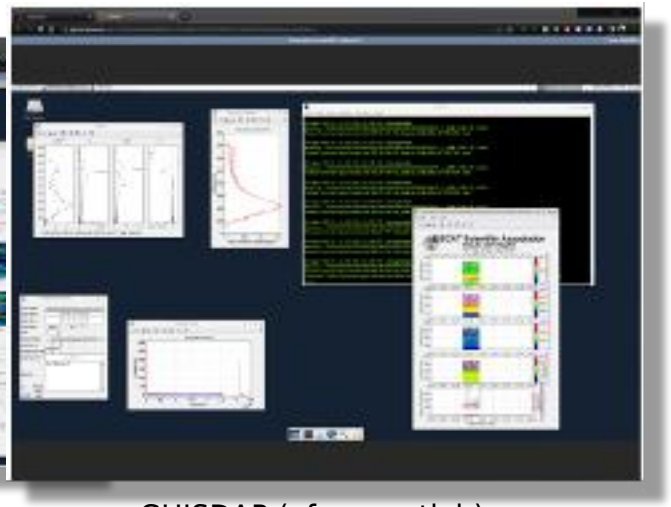
User Manual

EISCAT Notebook is in production:
https://notebooks.egi.eu/
https://jupyter.eiscat.se/

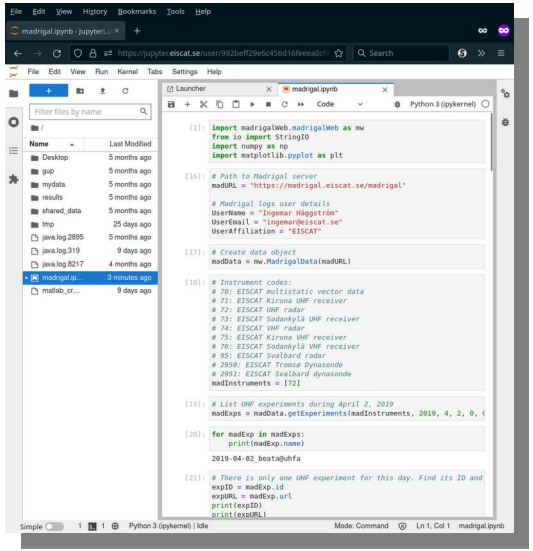
Notebook screenshots



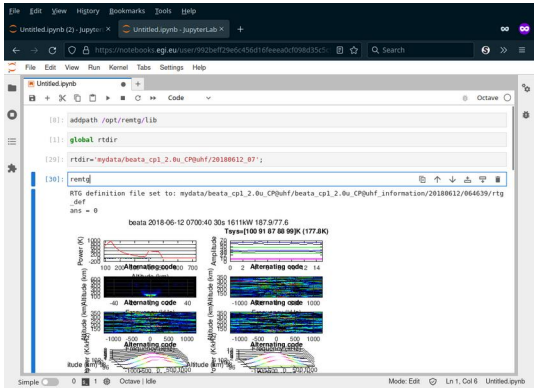
RTG (xfce+matlab)



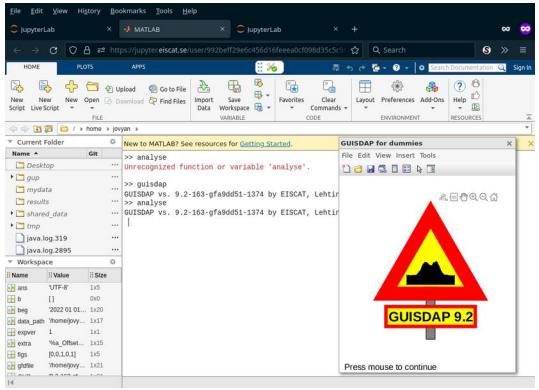
GUISDAP (xfce+matlab)



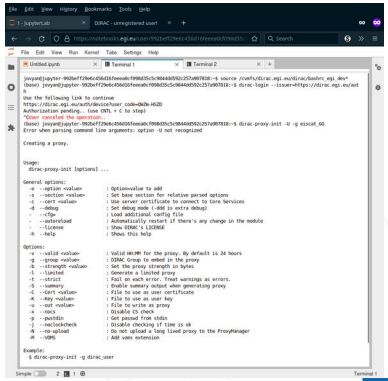
Madrigal (python)



RTG (octave)



GUISDAP (Matlab)



Dirac-cvmfs (Terminal)

Summary

- **Why important**
  - CSTCloud and EGI integration is one of the **first** Cloud federation across continent
  - EISCAT and SYISR science communities are the **first** test
    - e.g., user access from Europe/China, data interoperability, data sharing & movement, federated data process, etc.
- **Next steps**
  - **Notebook developments**
    - EISCAT+Matlab: Matlab with Jupyter environment
    - EISCAT: radar independent mode setup w. lagprofilers
    - SYISR: calibration routines
    - Some legalities (RoR, ToU...)
  - **Technical implementation**
    - EGI: DIRAC integration with Jupyter notebook
    - EGI+EISCAT: Job submission from DIRAC to HPC
    - CNIC: notebook to be accessible by all
    - EISCAT+EGI: integration EISCAT notebook with DIRAC
    - SYISR+CNIC: testing EISCAT notebook, data interoperability