

## **CEDA Sentinel ARD data access**

Ed Williamson
Earth Observation Data Scientist (CEDA)













# **Today talk**

1. CEDA Overview

2. API access

3. JASMIN and JASMIN notebooks





#### **CEDA** overview

#### Who are CEDA?

- Part of UKRI
- Part of STFC RAL Space, but provide the Earth Observation and Atmospheric Science data centres for NERC.
- The CEDA mission is To provide data and information services for environmental science.
- The CEDA team is roughly 30 staff. Split between data scientists and developers.
- CEDA store over 15PB of archived data which is in ~292 million data files.
- CEDA have over 66,000 registered data users.





# Searching for data at CEDA

#### Methods of finding data at CEDA

- CEDA have many different methods of finding data:
  - CEDA Catalogue (<a href="https://catalogue.ceda.ac.uk/">https://catalogue.ceda.ac.uk/</a>)
  - Data browser (<a href="http://data.ceda.ac.uk/">http://data.ceda.ac.uk/</a>) Earth observation datasets can be found under the NEODC directory.
  - CEDA Satellite data finder (<a href="http://geo-search.ceda.ac.uk/">http://geo-search.ceda.ac.uk/</a>)
    - Elasticsearch interface ceda-eo index

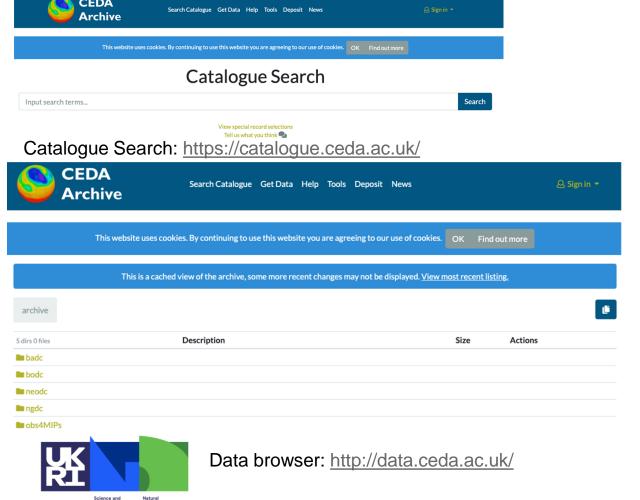
(https://help.ceda.ac.uk/article/4694-ceda-elasticsearch)

- CEDA provide different ways to download/access the data:
  - OpenDAP behind the Data browser
  - FTP
  - Direct access using JASMIN.

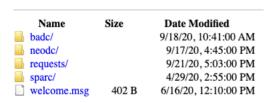


# Searching for data at CEDA

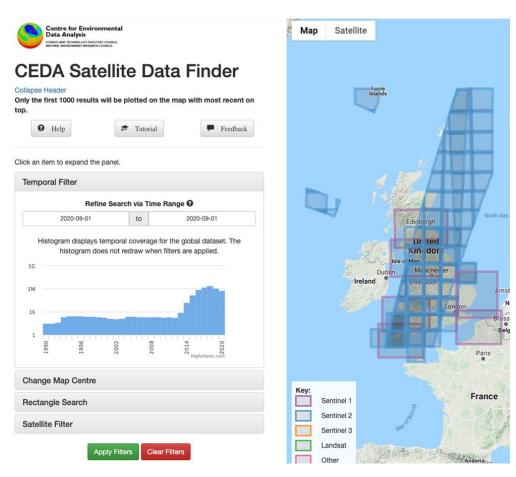
Methods of finding data at CEDA



Index of /



FTP: ftp://ftp.ceda.ac.uk/



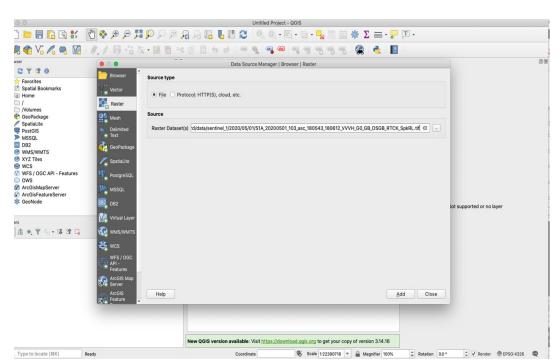
Satellite data finder: http://geo-search.ceda.ac.uk/

### **API** data access

#### CEDA ARD data can be accessed directly via OpenDAP

- Data can be accessed directly using the OpenDAP servers.
- Can be pulled directly into GIS software using 'vsicurl'.
- Similar to access via cloud computing.
- /vsicurl/http://dap.ceda.ac.uk/neodc/sentinel \_ard/data/...
- Tested using QGIS, gdal and tuiview.





### **JASMIN** access

### CEDA data can be accessed directly using JASMIN

- JASMIN is a 'super-data-cluster' which delivers infrastructure for large scale data analysis and is administered by CEDA.
- JASMIN in total has 44PB storage capacity.
- JASMIN provides direct access to the CEDA archive for data processing.
- Scientific analysis data servers to run and test data processing.
- LOTUS: batch processing cluster running almost 8000 cores.



JASMIN webpage: http://jasmin.ac.uk/

JASMIN help pages: <a href="https://help.jasmin.ac.uk/">https://help.jasmin.ac.uk/</a> JASMIN helpdesk: <a href="mailto:support@jasmin.ac.uk">support@jasmin.ac.uk/</a>



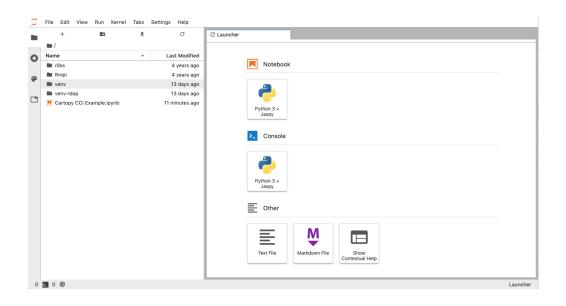
Credit: STFC

### **JASMIN Notebook service**

#### Accessing ARD data via JASMIN notebook service.

- New in 2020
- Anyone with JASMIN access can use the JASMIN notebook service
- Has direct read access the CEDA archive
- Allows you to run Python notebooks.
- You can bring previously written notebooks or write new ones from scratch.
- Some examples of notebooks can be found here: <a href="https://github.com/cedadev/ceda-notebooks">https://github.com/cedadev/ceda-notebooks</a>
- ARD examples will be added soon!







# Helpful links

CEDA help pages: <a href="https://help.ceda.ac.uk/">https://help.ceda.ac.uk/</a>

**JASMIN** help pages:

https://help.jasmin.ac.uk/

JASMIN: support@jasmin.ac.uk

CEDA: support@ceda.ac.uk

Twitter - @cedanews

Website - www.ceda.ac.uk













