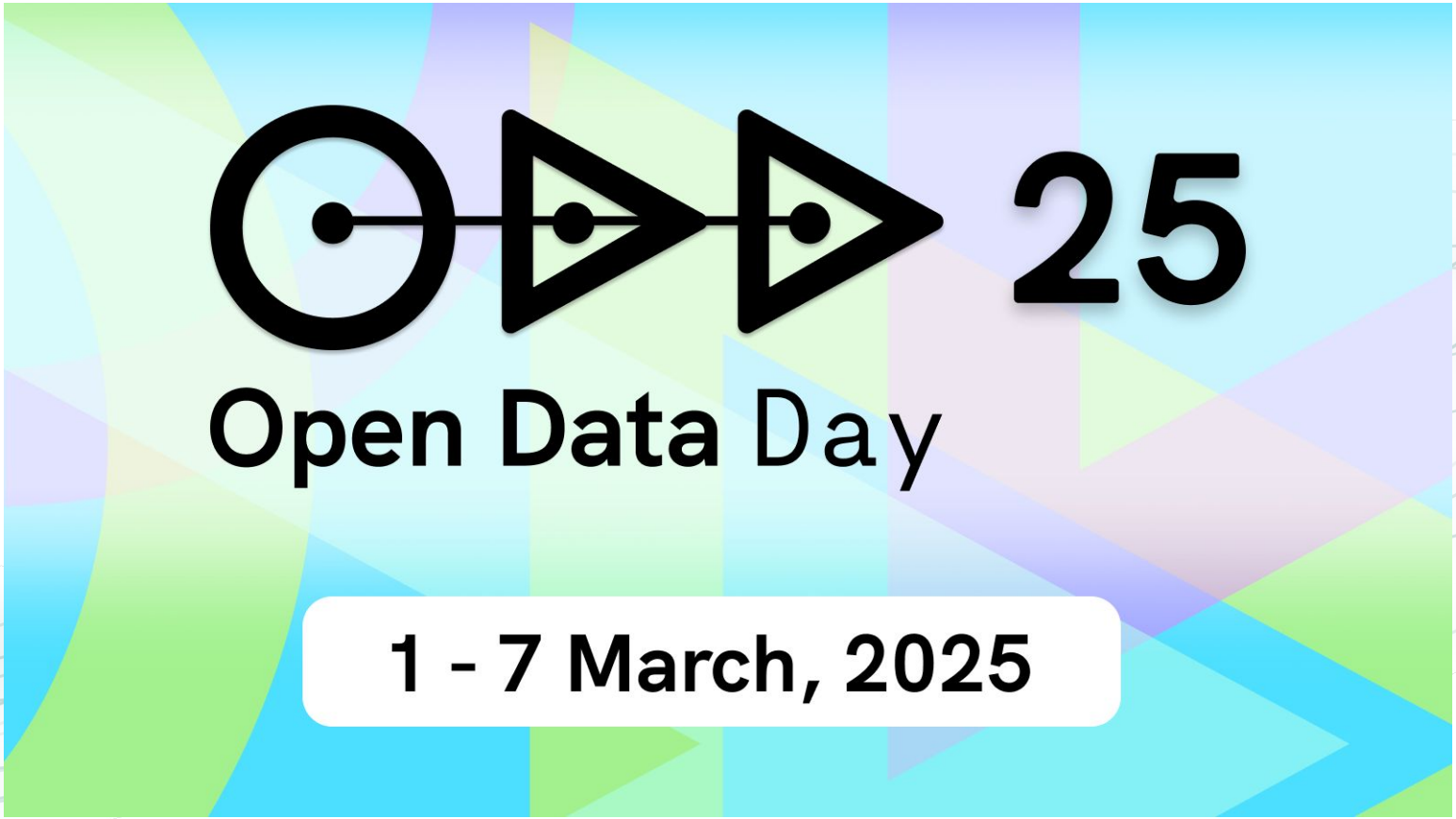
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

Open Data Day: environmental data sources for exploring the polycrisis

Sadie Bartholomew
Computational Scientist, NCAS-CMS & University of Reading Met. Dept.
NCAS Harmony in the Skies: Bridging Horizons in Atmospheric Science
2025-03-07

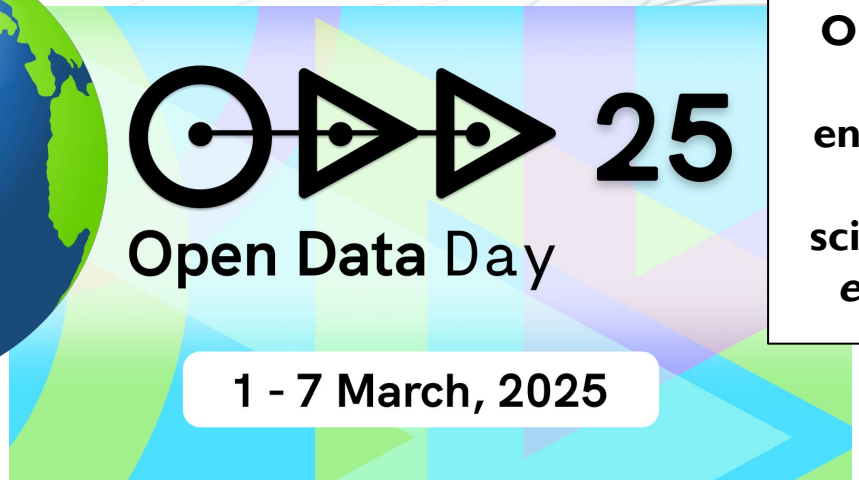


University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Workshop Overview

- **It's Open Data Day (Week) across the world! Brief summary of ODD**
- I summarise various sources of open environmental data & why open data is so useful, in general and in the domain
- I summarise the polycrisis, which underlies this year's ODD theme
- We explore sources of open environmental data together: **a (relaxed) competition**!



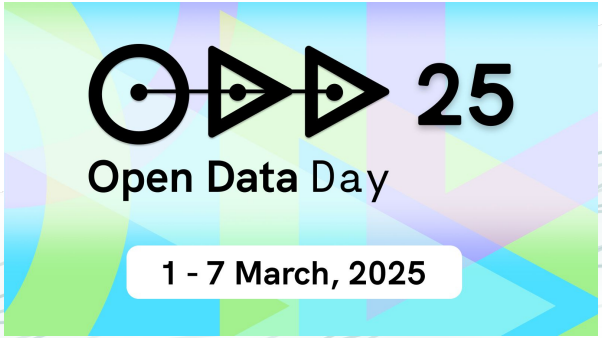
Our scope: open data with relevance to environmental (earth, space, etc.) science/research: open *environmental* data

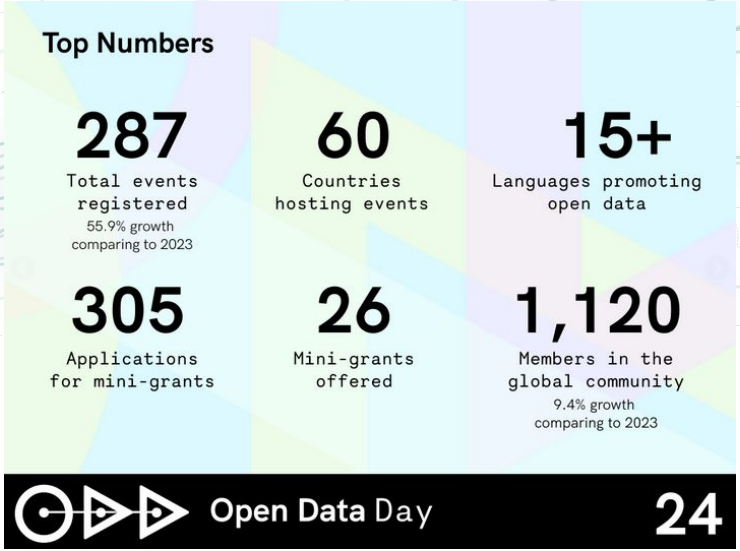
University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Overview: what is Open Data Day (/Week)?

- ODD is an annual international celebration of open data
- 15-year history: was first proposed in 2010 by David Eaves, a public policy entrepreneur
- Groups from likely every country create local events on the day where they will use open data in their communities - like our workshop today :)
- ODD is maintained and supported by the Open Knowledge Foundation (OKFN) through the Open Knowledge Network (note: this session was my idea and has no sponsor!)





University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Overview: (but wait,) what is *Open Data*?

Open data is data that *anyone* can *access*, *use* and *share* because it:

- has been made available somewhere easily accessible
- is stored in a standard, machine-readable format
- is licensed - conditions of use are made clear
- is free (to download and use)!

University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Overview: what is *the polycrisis*?

- The theme of ODD'25 is 'Open Data to Tackle the Polycrisis'. Let's stay on theme!
- Polycrisis = "*describes a complex situation where multiple, interconnected crises converge and amplify each other, resulting in a predicament that is difficult to manage or resolve… unlike single crises that may have clearer causes and solutions, a polycrisis involves overlapping and interdependent issues*" [Wikipedia definition]

University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Open data and the polycrisis

- ODD'25 describes this year's theme more specifically as: **the concept of polycrisis, and more specifically the intersections between poverty and multiple inequalities, violence and conflict, and climate and disaster response**
- Given our short session, *we'll just focus on environmental datasets*, with the appreciation that any dataset on the environment would be useful for investigating the many interconnection between the above e.g. the climate and inequality/conflict

University of Reading

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Our ODD: a (relaxed) competition

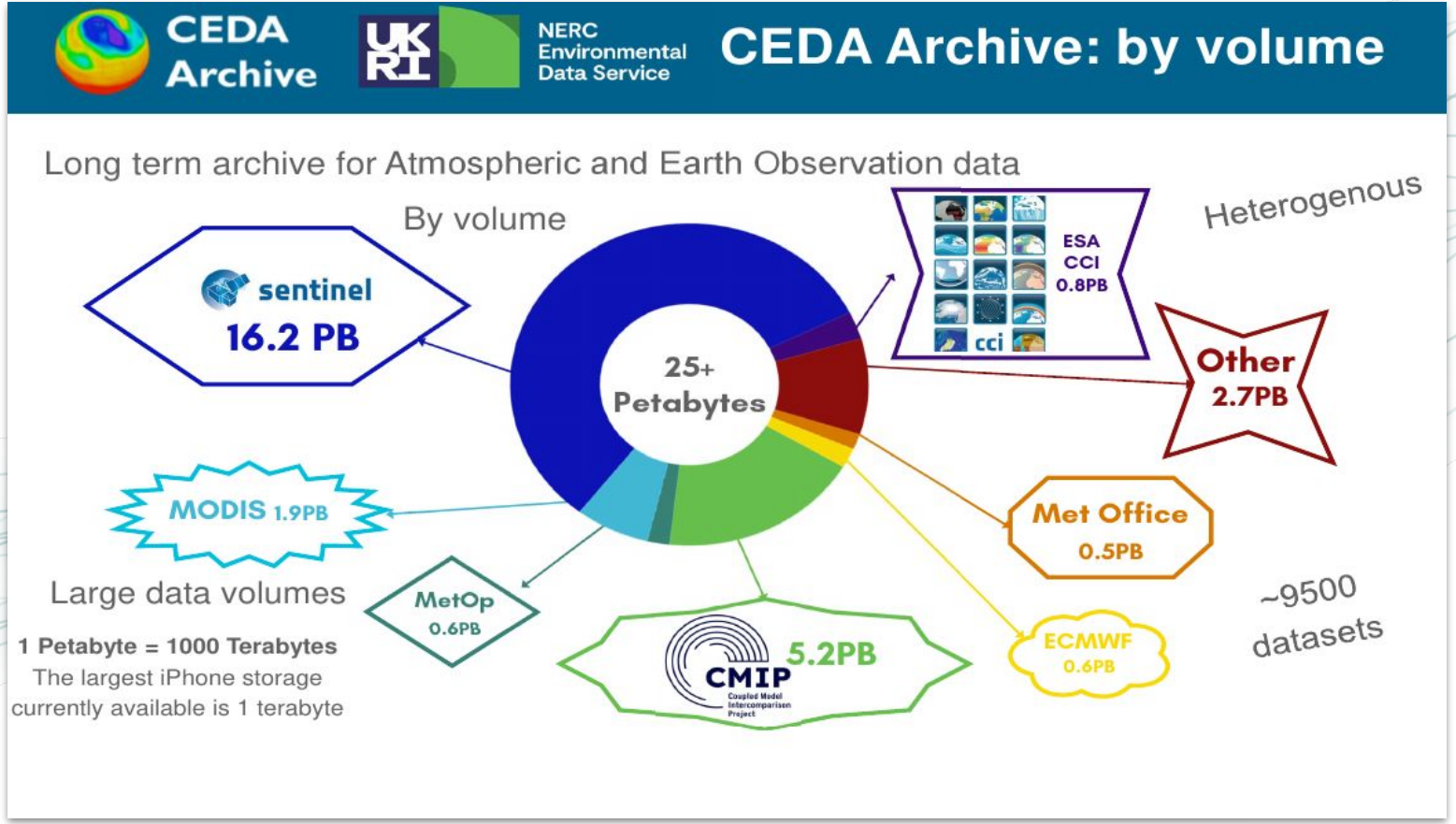
Summary: sources of Open Environmental Data

There are numerous sources of open environmental data, but I am going to provide some examples as a guide. You can use any you can find, or already know of/use, to do the challenge coming next!

- NCAS's own **CEDA**: https://help.ceda.ac.uk/article/98-accessing-data, the CEDA catalogue search ise a way in: https://help.ceda.ac.uk/article/137-ceda-data-catalogue
- **NCEO**: https://www.nceo.ac.uk/data-facilities/datasets-tools/?dataset_type=datasets
- Various **NASA** sources: https://www.earthdata.nasa.gov/, https://data.nasa.gov/
- Further earth observation data e.g. from EU's **Copernicus**: https://dataspace.copernicus.eu/explore-data
- ODD lists various other environmental sources at https://opendataday.org/#resources
- And so many more! Use any good sources you know of, and please share them with us

CEDA as a great source of Open Environmental Data

Slide kindly provided by / courtesy of Ellie Fisher of CEDA:



University of Reading

CEDA as a great source of Open Environmental Data

Slide (also) kindly provided by / courtesy of Ellie Fisher of CEDA:

EO Open Access Data at CEDA

* access to some data restricted

NCEO

- MIPAS atmospheric profiles, **MIPAS clouds**
- Southampton TOPEX data, Plymouth plankton products
- **Biomass maps**
- Gome O3 profiles
- Leicester LST data
- Swansea GLAS vegetation height
- **Weighing Trees with Lasers**

Sentinel data

- S1A+**B, S2A+B, S3A, S3B, S5P products, most global**
- Defra/JNCC S1/S2 ARD (Analysis Ready Data) - UK (except Wales)

ESA – ERS and Envisat*

- (A)ATSR, MIPAS, MERIS,SCIAMACHY
- SAR (limited)

Eumetsat – MetOp A+B*

- IASI, AVHRR-3, GOME-2

NASA/NOAA

- CALIOP, AVHRR_GAC, **Landsat8**
- **MODIS – many products**

NERC facilities/projects

- ARF 1978-present
- ARC SST
- NCAVEO (Network for Calibration & Validation of EO data)
- Globolakes

CCI

- **Data products for all ECVs**

(SST, OC, GHG, Cloud, Aerosol, Landcover, Glaciers, Ice Sheets, Ozone, Sea Level, Soil Moisture, Sea Ice, Fire)

Other

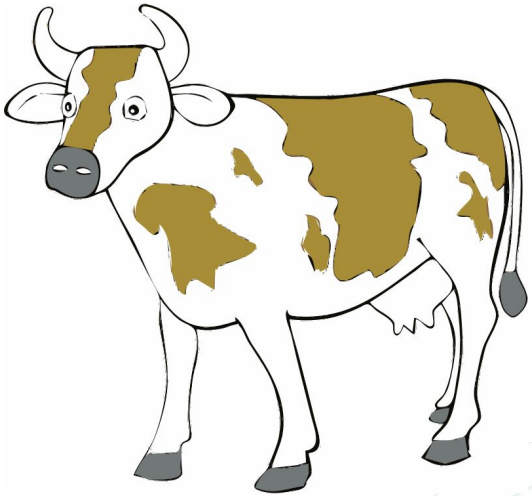
- AATSR Validation data (ISAR, SISTER)
- **SLSTR Calibration data**
- MTCI (MERIS Terrestrial Chlorophyll Index)
- SEVIRI FRP (Fire Radiative Power)
- Globsnow
- COMET Continuous GPS
- GERB (Geostationary Earth Radiation Budget)

The challenge itself

Challenge: find datasets that you fulfil these criteria as best as you can find, and save the link to them so we can compare answers a group. Notes/clarifications:

- To be eligible, datasets must be **open** and **concern some aspect of the environment in their nature**.
- You do not need to download the dataset for it to count, please save a link to the page which provides the basic information about it (metadata)
- Don't worry about registration needs for the workshop - as long as the data is open pending source registration and we can see the description, that is fine (counts)

Pre-challenge example



Let's find a dataset that concerns cows, as an example!

- [demo on screen]

Challenge items

Challenge: find datasets that you fulfil these criteria as best as you can find, and save the link to them so we can compare answers a group. To be eligible, datasets must be **open** and **concern some aspect of the environment in their nature**.

Objective cases:

- *Oldest* dataset, with two cases:
  - oldest by age/time of data represented
  - oldest by dataset creation *or* editing
- Dataset covering the *smallest* area of the globe
- *Largest* single data file (no need to actually download it!)
- Non-terrestrial dataset *furthest away* from Earth (but still in the solar system!)

Subjective (or hard to tell) cases, to be judged as a group on a most-votes/agreement basis:

- Most unusual!
- (Likely to) have the *most variability* (i.e. range) across the data
- (Likely to) have the *least variability* (i.e. range) across the data

Let's compare our findings!


University of
Reading


National Centre for
Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL