Code sharing and knowledge exchange for users of Sentinel-1 and Sentinel-2 analysis-ready data

Webinar: Thursday 24th September 2020

Gwawr Jones and Paula Lightfoot, JNCC • Matthew McArthur, Defra • Ed Williamson, CEDA

Copernicus
Framework Partnership Agreement for Copernicus User Uptake

JNCC

Department for Environment Food & Rural Affairs

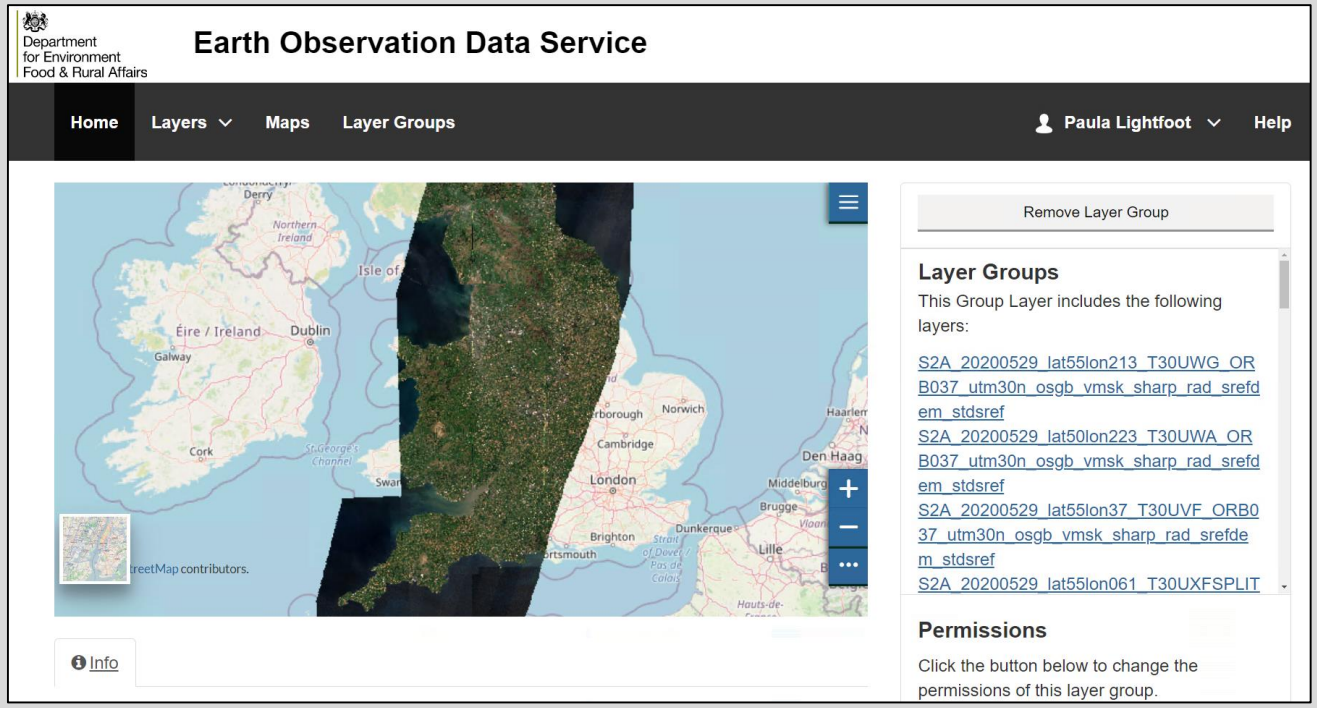
Centre for Environmental Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

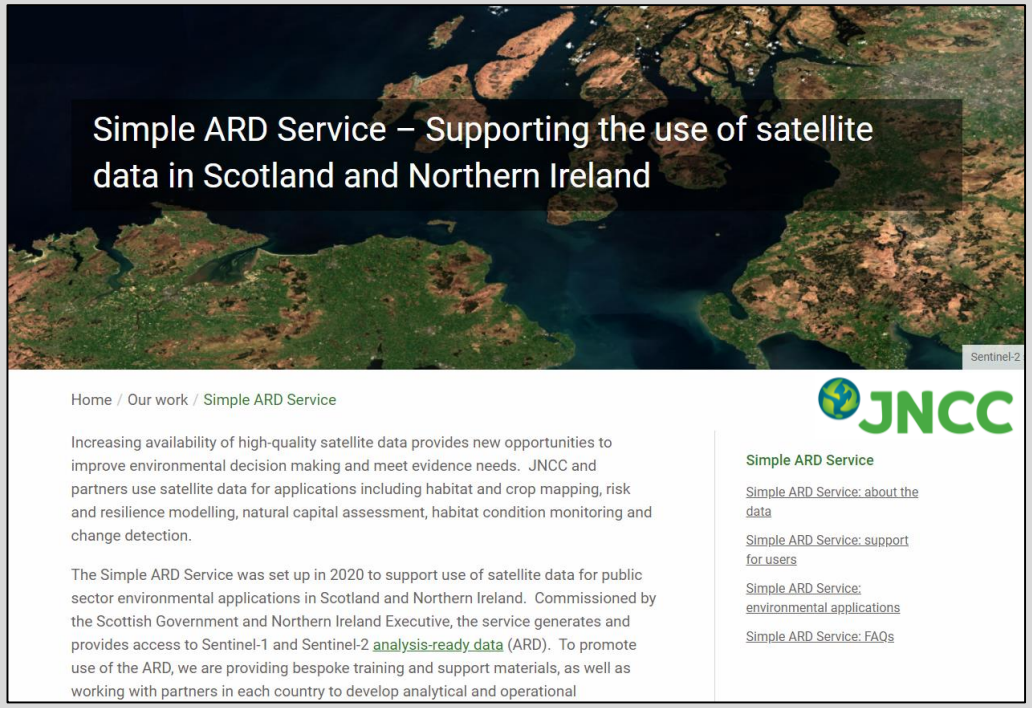
Webinar contents

- Background to project
  - Defra Earth Observation Data Service
    - Centre for Environmental Data Analysis
      - Code sharing - GitHub
      - Knowledge exchange - Slack
    - Next steps
  - Questions
    - ask via ‘question’ tab in GoToWebinar panel
    - webinar is recorded but questions anonymised
    - recording and slides will be available after the event

Background

- Defra Earth Observation Data Service and JNCC's Simple ARD Service launched in 2020.
- Routine production of Sentinel-1 and Sentinel-2 analysis-ready data (ARD) for England, Scotland and Northern Ireland.
- Data from both services available under an Open Government Licence via the Centre for Environmental Data Analysis (CEDA) Archive.



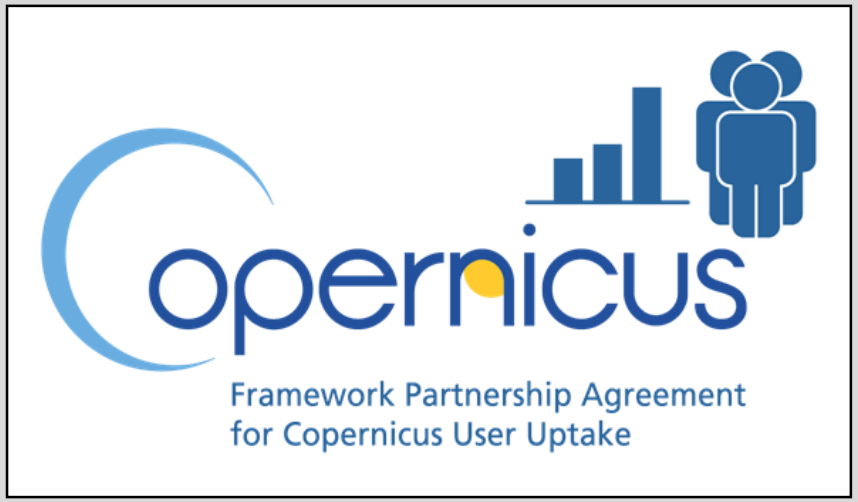


Background

- Both services offer data download, but it is more efficient for analytical users to access data programmatically via an application programming interface (API).
- JNCC produced a user guide, three example scripts, and ran a workshop and webinar on EO Data Service API use in February-March 2020.
- Feedback from participants identified a need for more example scripts and for code sharing and knowledge exchange platforms to enable peer-to-peer support.

Background

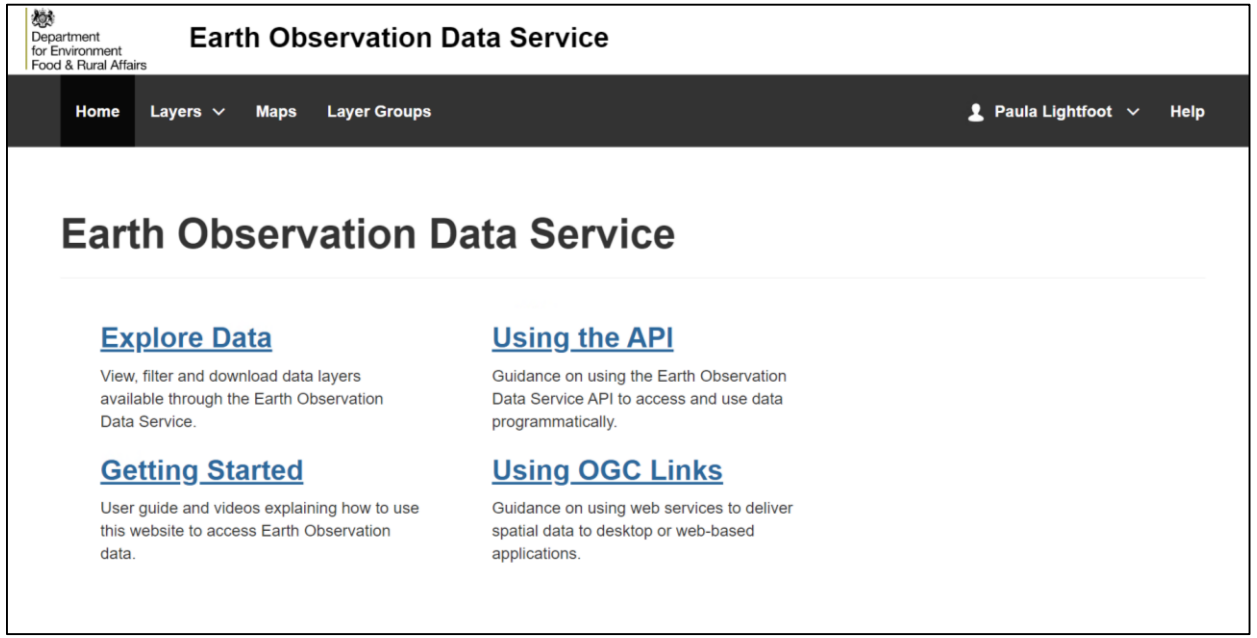
- Caroline Herschel Framework Partnership Agreement for Copernicus User Uptake funding to set up code sharing and knowledge exchange platforms and develop more example scripts.
- Focus on facilitating data use via EO Data Service and CEDA APIs, but wider aim of supporting all Sentinel-1 and Sentinel-2 ARD users.
- Survey in July 2020 to gain better understanding of user needs and inform choice of platforms.
- GitHub repository and Slack workspace set up September 2020.



The Framework Partnership for Copernicus User Uptake is financed by the European Commission under the FPA no.: 275/G/GRO/COPE/17/10042

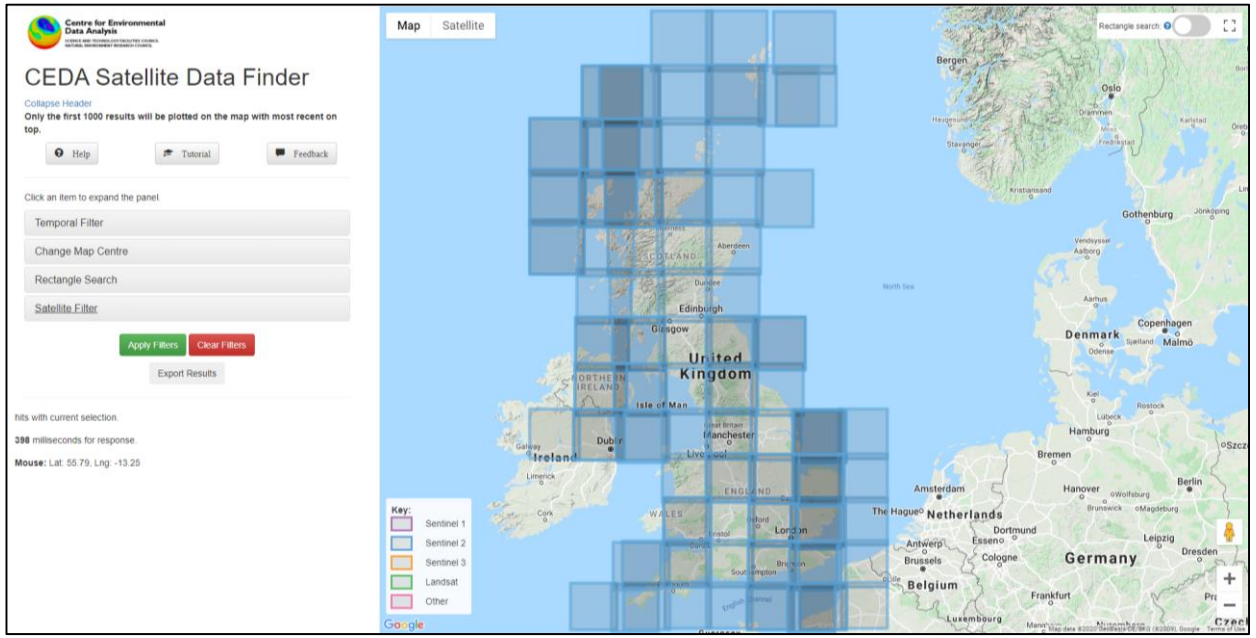
Department
for Environment
Food & Rural Affairs

Defra Earth Observation Data Service



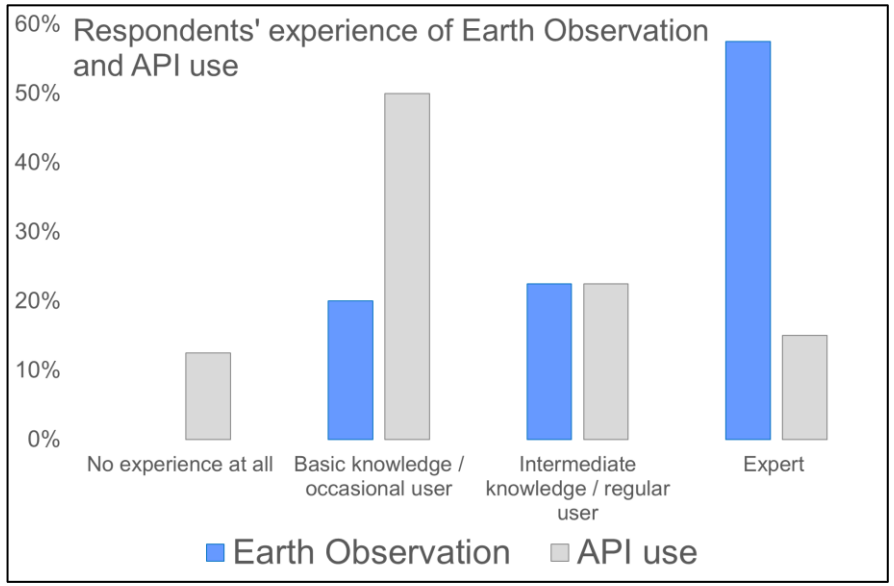
Centre for Environmental
Data Analysis
SCIENCE AND TECHNOLOGY FACILITIES COUNCIL
NATURAL ENVIRONMENT RESEARCH COUNCIL

Centre for Environmental Data Analysis (CEDA) Archive

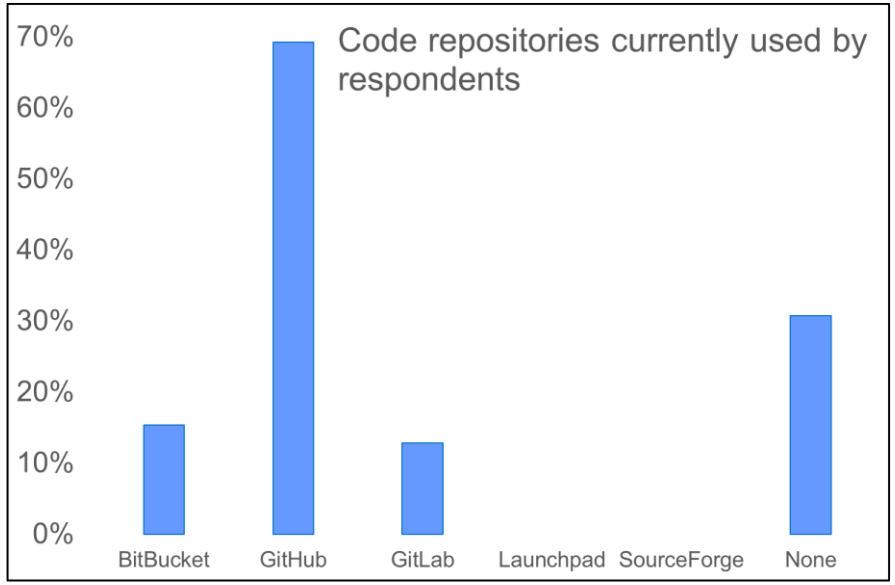


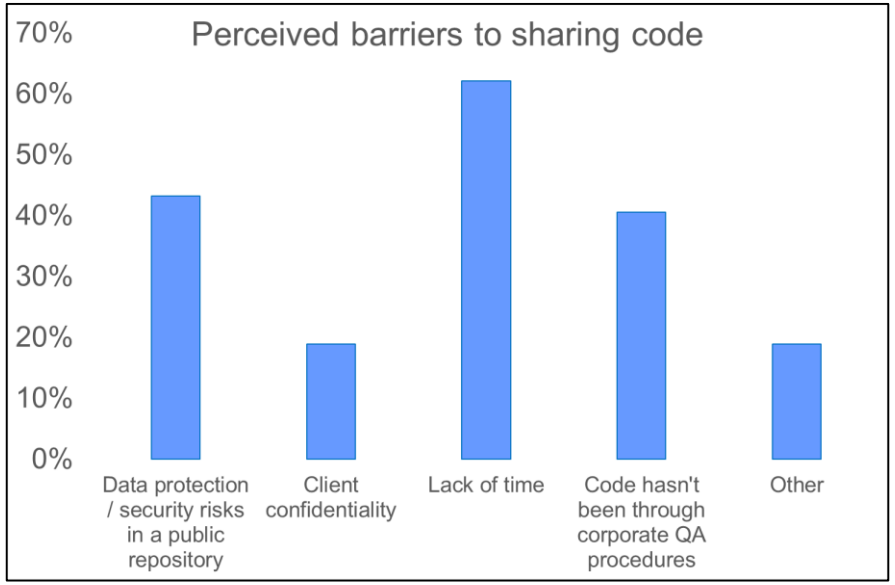
Code sharing

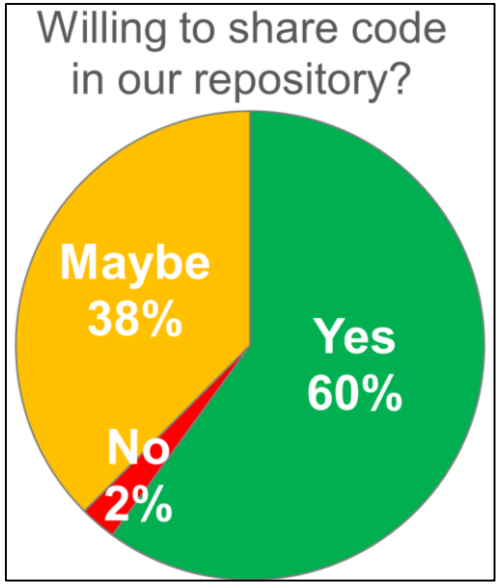
User consultation received 40 responses from at least 24 organisations



Importance of features	Not important	Fairly important	Very important	Essential
Version control	0%	16%	32%	51%
Issue / bug tracking	5%	14%	49%	32%
Ease of commenting your own code	3%	16%	49%	32%
Code review tools	8%	24%	49%	19%
Security features	14%	24%	51%	11%
Upload speed	8%	39%	47%	6%
Wiki	13%	42%	24%	21%
Community forum	16%	43%	27%	14%
Organisation accounts	19%	50%	19%	11%
Project management	14%	57%	22%	8%







Code sharing

The decision was taken to set up a **curated list on GitHub**:

- Already widely used by potential contributors
- Provides the features required by users
- Easy to update and document
- Easy to add contributions
- Free to set up a public repository
- Avoid duplication by linking to relevant code in public repos

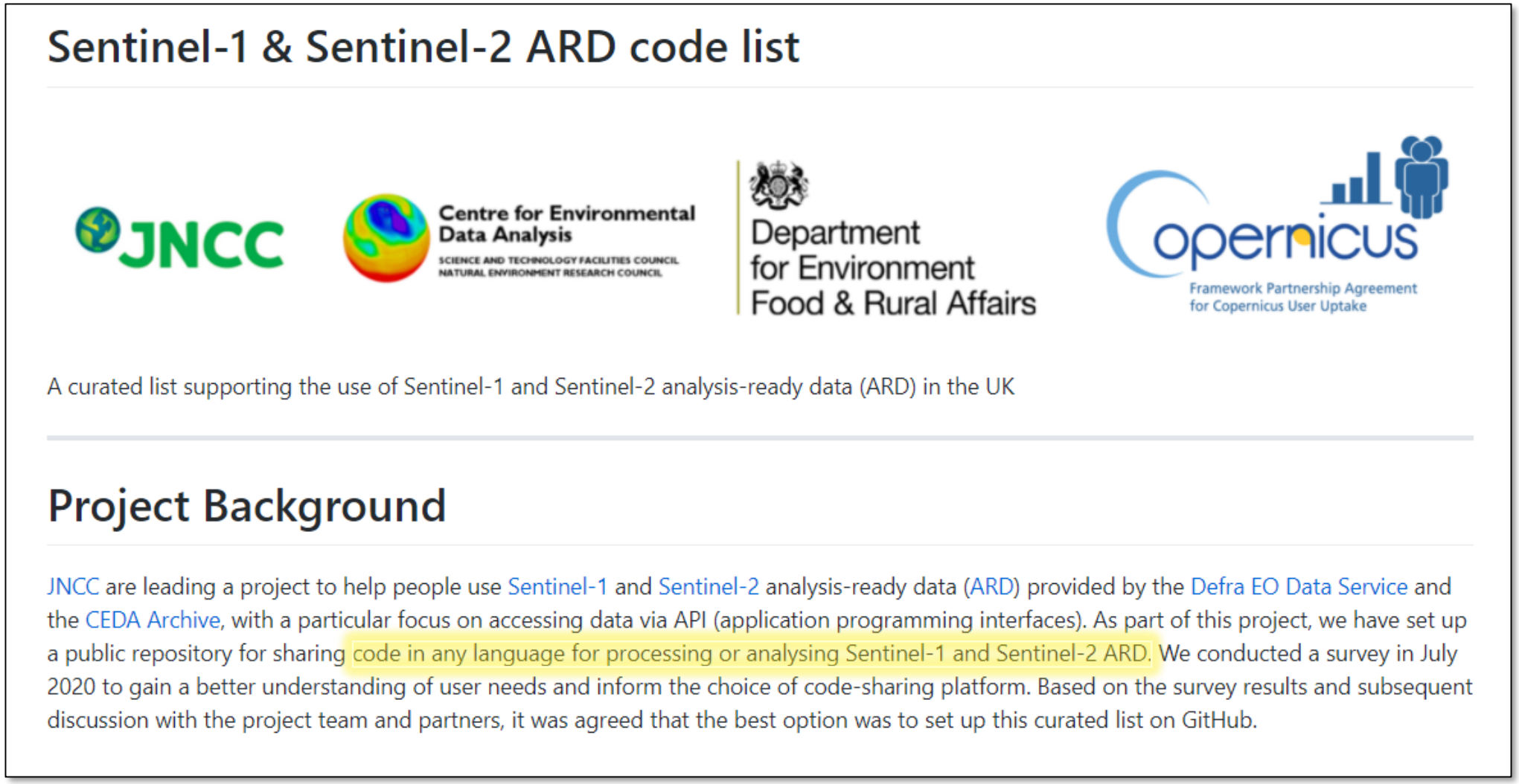
This decision was inspired by the Awesome-EarthObservation-Code list created by Andrew Cutts (acgeospatial)

https://github.com/acgeospatial/awesome-earthobservation-code

“A curated list of awesome tools, tutorials, code, helpful projects and links about Earth Observation and Geospatial stuff!”

Code sharing

You can find the list here:



https://github.com/jncc/s1-s2-ard-code-list

Code sharing

Content and structure

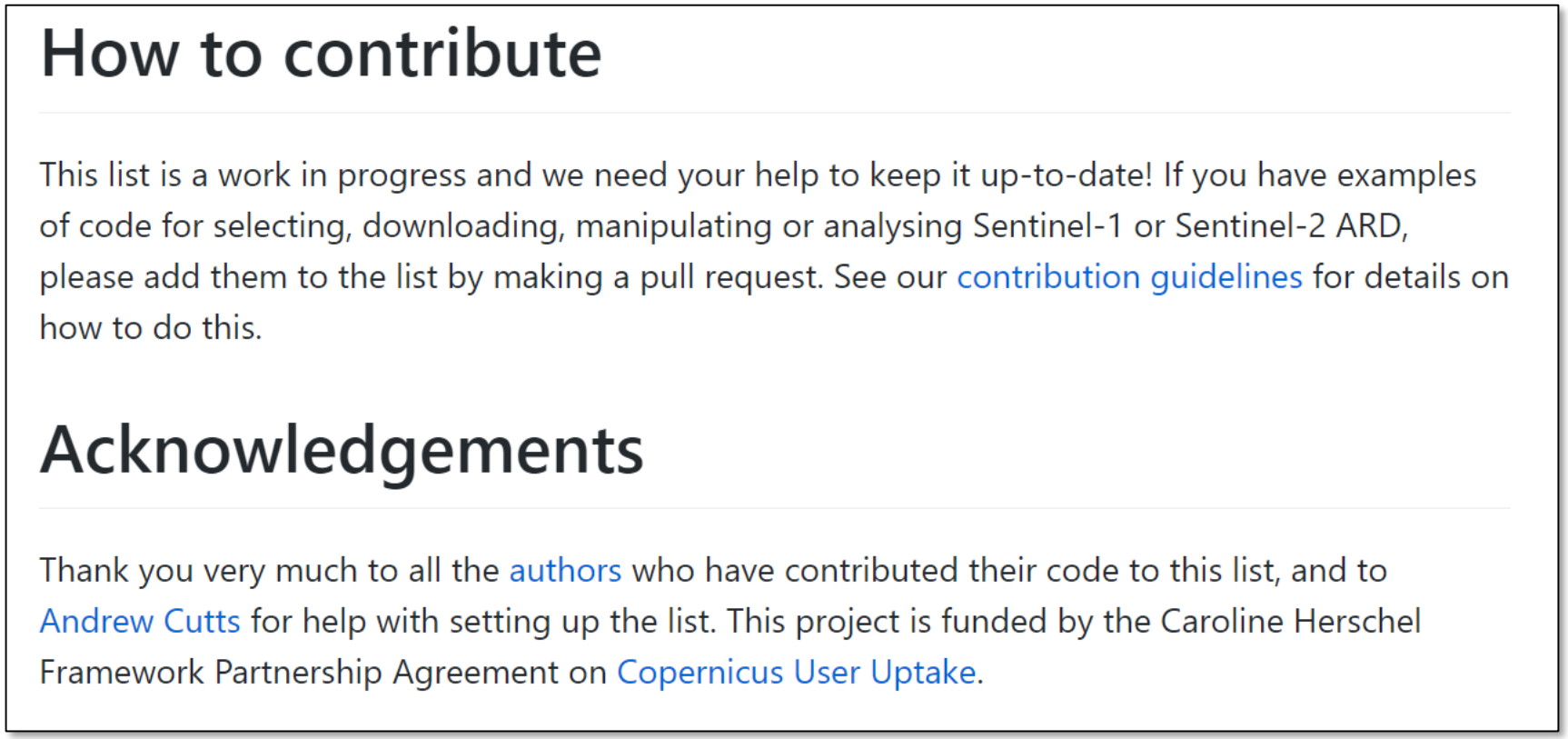
- The list of code is split into three main sections with sub-sections
- Can add more sections as needed
- Tutorials and functions: example scripts used for training and a selection of useful generic processes
- Applications: complete workflows

Contents

Code sharing

Contributing

- We are grateful for any contributions!
- The best way to contribute is via a pull request – full details on how to do this are provided.
- Information is also provided on how to edit and commit files, online and offline using GitHub Desktop and Git command line.
- Contributors are acknowledged in the authors list (and in list of contributors if done via pull request).



How to contribute

This list is a work in progress and we need your help to keep it up-to-date! If you have examples of code for selecting, downloading, manipulating or analysing Sentinel-1 or Sentinel-2 ARD, please add them to the list by making a pull request. See our contribution guidelines for details on how to do this.

Acknowledgements

Thank you very much to all the authors who have contributed their code to this list, and to Andrew Cutts for help with setting up the list. This project is funded by the Caroline Herschel Framework Partnership Agreement on Copernicus User Uptake.

Code sharing

Example 1

- Two new Jupyter Notebooks demonstrating use of the EO Data Service API created as part of this project.
- Brief description and keyword tags for each entry.
- Ancillary files are provided.

- 'EODS_API' Download Sentinel-2 data via API from the Defra EO Data Service, then create a mosaic, clip to area of interest and create NDVI. `Python` `JNCC` `EODS-API` `WPS`
- 'EODS_API_Best_pixel'. Download Sentinel-2 data via API from the Defra EO Data Service, then create best-pixel composite using the pixels with least cloud in a stack of imagery from different dates `RSGISLib` `Python` `JNCC` `EODS-API` `WPS`
- config.py. Config file for use with the above Jupyter Notebooks, e.g. to input your authentication token.
- environment.yml. Environment file for use with the Jupyter Notebooks in this section and the section below.

Code sharing

Example 1

- Readme.md file in the source repo explaining the context and purpose of the code and anything else users will need to know in order to use the code.

Accessing Sentinel-2 satellite imagery via the Defra EO Data Service API

These Jupyter Notebooks were created to demonstrate how to access Sentinel-2 analysis-ready data (ARD) using the application programming interface (API) provided by the Defra Earth Observation Data Service. The EO Data Service is available to members of staff at Defra, its agencies and arms length bodies. In order to run these Notebooks, you need to know your username and authentication token for the EO Data Service and you will need to use a computer on your organisation's IT network.

Installing RSGISLib

For the next two notebooks to work correctly the RSGISLib library package for Jupyter Notebook is needed.

To make a conda install of this follow the instructions below:

In Anaconda prompt:

1. Create the conda environment and install the two libraries needed (rsgislib and Jupyter) with:

```
$ conda create --name rsgislibenv -c conda-forge rsgislib jupyter
```

2. activate the environment you just created with:

```
$ conda activate rsgislibenv
```

3. start the Jupyter notebook server which should launch in a browser or give the localhost URL:

```
$ jupyter notebook
```

Code sharing

Example 1

- Script commented to explain to users step-by-step what the code does.
- Personal data (e.g. staff names) and corporate data (e.g. file paths) removed from code.

In [ ]:
```
#Importing the neccesary libraries for downloading data
import requests
import time
from datetime import datetime
import pandas as pd
from pathlib import Path
import xmltodict
from zipfile import ZipFile
```

The following cell will create a folder, where the notebook is located, and all the downloaded images will be stored there

In [ ]:
```
# get input for "test name" eg 'PRD_STANDARDISED_TEST_WPS_S1x10_RUN1'
RUNNAME = input()
```

In [ ]:
```
output_fmt='%Y%m%dT%H%M%S'
pretty_fmt='%Y-%m-%d %H:%M:%S'
headers = {'Content-type': 'application/xml','User-Agent

wps_test_config = {
        'gs:Download':{
            'template_xml':'gsdownload-template.xml',
            'mimetype':'application/zip',
            'output':'result.zip',
        }}
```

In [ ]:
```
# Load list of granules from dowloaded csv
#(Those are the resulting csv from eods-api-example-gene
col_list = ["typename"]
min_cloud_per_granule = pd.read_csv(r"C:\Users\.......\min_cloud_per_granule.csv")

min_cloud_per_granule_per_orbit = pd.read_csv(r"C:\Users\.........\min-cloud-per-granule-per-orbit.csv")
```

```
### USER INPUT
### Enter your token in PRD_DM =

wps_job_request_list = [
     {'tool':'gs:Download','layer_list':min_cloud_per_granule_per_orbit["typename"],'dl_bool':True}
]

PRD_DM = "Your token goes here"
ACCESS_TOKEN = PRD_DM

ENV_PRD = 'https://earthobs.defra.gov.uk'
wps_env = ENV_PRD

wps_server = wps_env + '/geoserver/ows/?access_token=' + ACCESS_TOKEN + '&SERVICE=WPS&VERSION=1.0.0'
```

Code sharing

Example 2

- R scripts to create indices from Sentinel-1 and 2 ARD to detect change in habitats and historic landscape features. Example of an analytical workflow.
- Brief description and keyword tags for each entry.
- Uses functions which are included in the ‘tutorials and functions’ part of the list, so links to that.

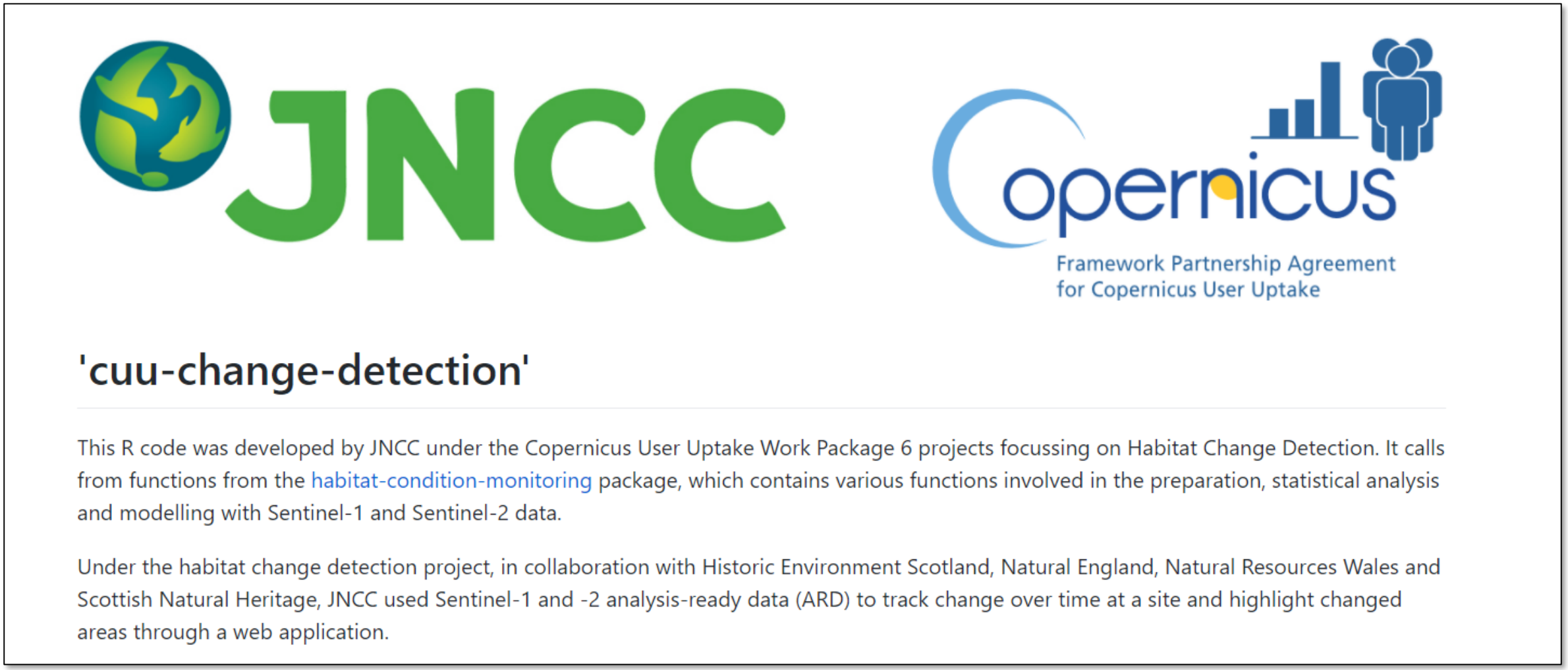
Habitat Change Detection

- Workflow for processing Sentinel-1 and Sentinel-2 with habitat map shapefiles to produce input data for change detection RShiny app, using functions provided in the 'habitat-condition-monitoring' package. There are separate workflows for English sites, Welsh sites and Scottish sites. `R` `JNCC` `copernicus-user-uptake`
- 'change-analysis-examples' A short analysis of NDVI statistics generated from Sentinel-2 and how they can be used in conjunction with a habitat map shapefile to identify polygons which deviate from mean values by more than set thresholds. `R` `JNCC` `copernicus-user-uptake`
- 'change-statistics-analysis'. An interactive document demonstrating use of NDVI derived from Sentinel-2 in conjunction with a habitat map shapefile to identify polygons which deviate from mean values by more than set thresholds. `R` `JNCC` `copernicus-user-uptake`
- Change detection RShiny App 2020 pilot - coming soon `R` `JNCC` `copernicus-user-uptake`

Code sharing

Example 2

- Readme.md file in the source repo explaining the context and purpose of the code and anything else users will need to know in order to use the code.



JNCC

Copernicus

Framework Partnership Agreement
for Copernicus User Uptake

'cuu-change-detection'

This R code was developed by JNCC under the Copernicus User Uptake Work Package 6 projects focussing on Habitat Change Detection. It calls from functions from the habitat-condition-monitoring package, which contains various functions involved in the preparation, statistical analysis and modelling with Sentinel-1 and Sentinel-2 data.

Under the habitat change detection project, in collaboration with Historic Environment Scotland, Natural England, Natural Resources Wales and Scottish Natural Heritage, JNCC used Sentinel-1 and -2 analysis-ready data (ARD) to track change over time at a site and highlight changed areas through a web application.

Code sharing

Example 2

- Script commented to explain to users step-by-step what the code does.
- Personal data (e.g. staff names) and corporate data (e.g. file paths) removed from code.

````
## Example site - Insh Marshes

```{r, include=F,eval=F}
#load in the data and flag change
#plot differences
dirpath <- 'Scotland/'
hab_stat <- read.csv(paste0(dirpath,"Statistics/InshMarshes_EUNIS/Seasonal_statistics/InshMarshes_NDVI_seasonal_changestats.txt"))
poly_stat <- read.csv(paste0(dirpath,"Statistics/InshMarshes_EUNIS/Seasonal_statistics/InshMarshes_NDVI_seasonal_stats.txt"))
polygons <- sf::st_read(paste0(dirpath,"Data/Habitat_maps/Scotland/ELCS_Scottish_site_1.shp")) %>% sf::st_drop_geometry()

#join the data
poly_hab <- poly_stat %>% dplyr::left_join(polygons[,c('Id','EUNIS_DESC')], by=c('ID'='Id'))
hab_stat <- hab_stat %>% dplyr::select(-X) %>% dplyr::rename(EUNIS_DESC = 'get.habclass.')
poly_all <- poly_hab %>% dplyr::left_join(hab_stat,by=c('seasonyear','EUNIS_DESC'))

#order factor levels - again
##get unique years
all_lev <- stringr::str_split(levels(poly_all$seasonyear),"_",simplify=T) %>% data.frame()  %>% dplyr::mutate_if(is.character,as.factor)

#order by season and year
all_lev$X1 <- factor(all_lev$X1, levels = c('Spring','Summer','Autumn','Winter'),ordered=T)
lev_sort <- all_lev %>% arrange(X1) %>% dplyr::arrange(X2) %>% tidyr::unite(yearseason, sep="_")
##sort levels in data
poly_all$seasonyear <- factor(poly_all$seasonyear, levels = lev_sort$yearseason)

# flag those above or below 2SD as change
````
````

Code sharing

Example 3

- Python library for interacting with the EO Data Service API.
- Very useful functionality which is not available via the portal, e.g. finding Sentinel-2 data with the lowest cloud cover for a given area and date range, removing Sentinel-2 ‘split granules’ from search results.
- Four Jupyter Notebooks to demonstrate use of the library functions.

EODS API Python library

A module of library functions for programmatic interaction with the EO Data Service developed by Sam Franklin at CGI, together with four Jupyter Notebooks demonstrating applications of the library for filtering, downloading and manipulating Sentinel-2 data.

- eodslib.py. Functions for interacting with the EO Data Service. Examples include: keyword arguments for filtering data; finding the least cloudy Sentinel-2 granules; creating XML files for use in Web Processing Service (WPS) requests; submitting WPS requests; removing split Sentinel-2 granules from a dataframe; processing data downloaded via WPS (unzipping and renaming files, deleting zip files). `Python` `Defra` `EODS-API` `WPS`
- eodslib.py example 1: simple query. Apply filter parameters and return results as a dataframe. `Python` `Defra` `EODS-API`
- eodslib.py example 2: query and download. Apply filter parameters and download returned data. `Python` `Defra` `EODS-API` `WPS`
- eodslib.py example 3: query, download, optimise and mosaic. Apply filter parameters, download results, convert downloaded files to 'optimised' GEOTIFFs and mosaic the imagery by creating a virtual raster. `Python` `Defra` `EODS-API` `WPS`
- eodslib.py example 4: query, download clipped area, calculate ndvi. Apply filter parameters, download the returned data clipped to an area of interest, calculate and plot NDVI from the downloaded data. `Python` `Defra` `EODS-API` `WPS`

```
def find_minimum_cloud_list(df):
    """
    eods query "special" keyword function
    + takes an input dataframe, groups by the granule-reference,
    + sorts by cloud cover and takes the lowest cloud per granule
    + returns a new dataframe
    """

    df_min_cloud_per_granule = df.sort_values("ARCSI_CLOUD_COVER").groupby(["granule-ref"],

    return df_min_cloud_per_granule

def ignore_split_granules(df):

    new_df = df[df["alternate"].str.contains("SPLIT") == False]

    return new_df
```

Overcomes two current issues with EODS API:

- unzips and renames downloaded data
- creates XML files for use with WPS requests.

Code sharing

Other information

- Links to relevant GitHub repos and lists.
- Others will be added, suggestions are welcome!

GitHub accounts of relevant organisations

- Nature.Scot - Scotland's Nature Agency homepage
- defra - UK government department responsible for safeguarding our natural environment, supporting our food & farming industry, and sustaining a thriving rural economy. homepage
- CefasRepRes - Cefas is a world leader in marine science and technology, providing innovative solutions for the aquatic environment, biodiversity and food security homepage
- cedadev - Centre for Environmental Data Analysis Developers homepage

Other useful Earth Observation GitHub links

- Awesome-EO-Code - A curated list of awesome tools, tutorials, code, helpful projects, links, stuff about Earth Observation and Geospatial stuff!
- Awesome-SAR - A curated list of awesome synthetic aperture radar (SAR) software, libraries and resources.

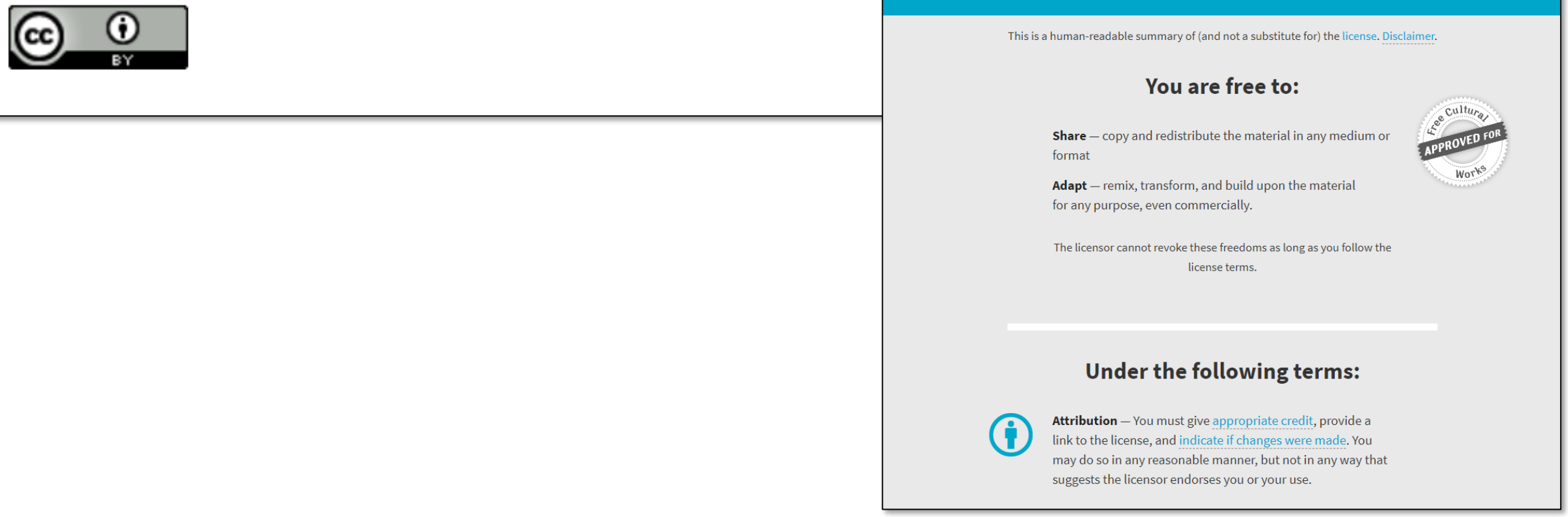
Code sharing

License

- Data shared under Creative Commons Attribution license (CC BY 4.0).
- Can be used or adapted for any purpose, including commercially, but please give credit.

This work is licensed under a Creative Commons Attribution 4.0 International License.

License CC BY 4.0



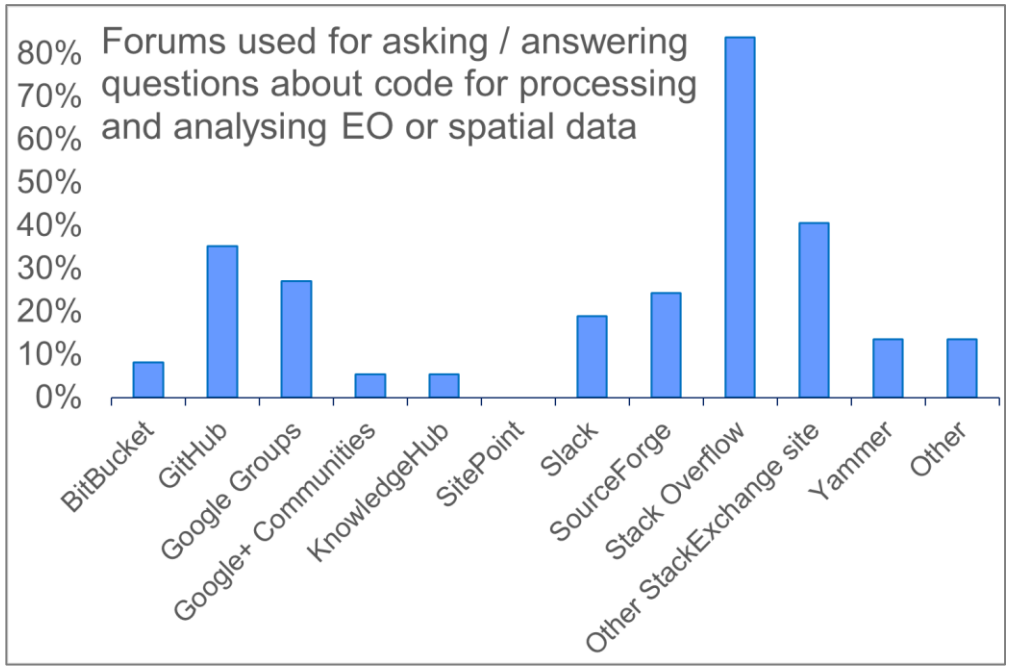
Code sharing

Recap of best practice

- Contribute to the list via pull request.
- Add code to the relevant section of the list. Create a new sub-section if necessary.
- Ensure code is well commented and provide a readme file to explain the context and purpose.
- Include any necessary ancillary files, e.g. environment.yml files for Jupyter Notebook.
- Ensure you have permission to share code if you are not the sole author.
- Remove personal data (e.g. staff names) and corporate data (e.g. file paths).
- If you use code, please acknowledge its creators.
- Please be constructive and courteous if you wish to let an author know about an issue with their code.

User Forum

User consultation results



Importance of features	Not important	Fairly important	Very important	Essential
Search function	0%	8%	32%	60%
Assign tags to questions	5%	37%	45%	13%
Send private messages to other users	27%	30%	32%	11%
Upvote and downvote posts	22%	41%	32%	5%
Build reputation points	44%	42%	14%	0%
Obtain privileges linked to reputation	72%	22%	6%	0%

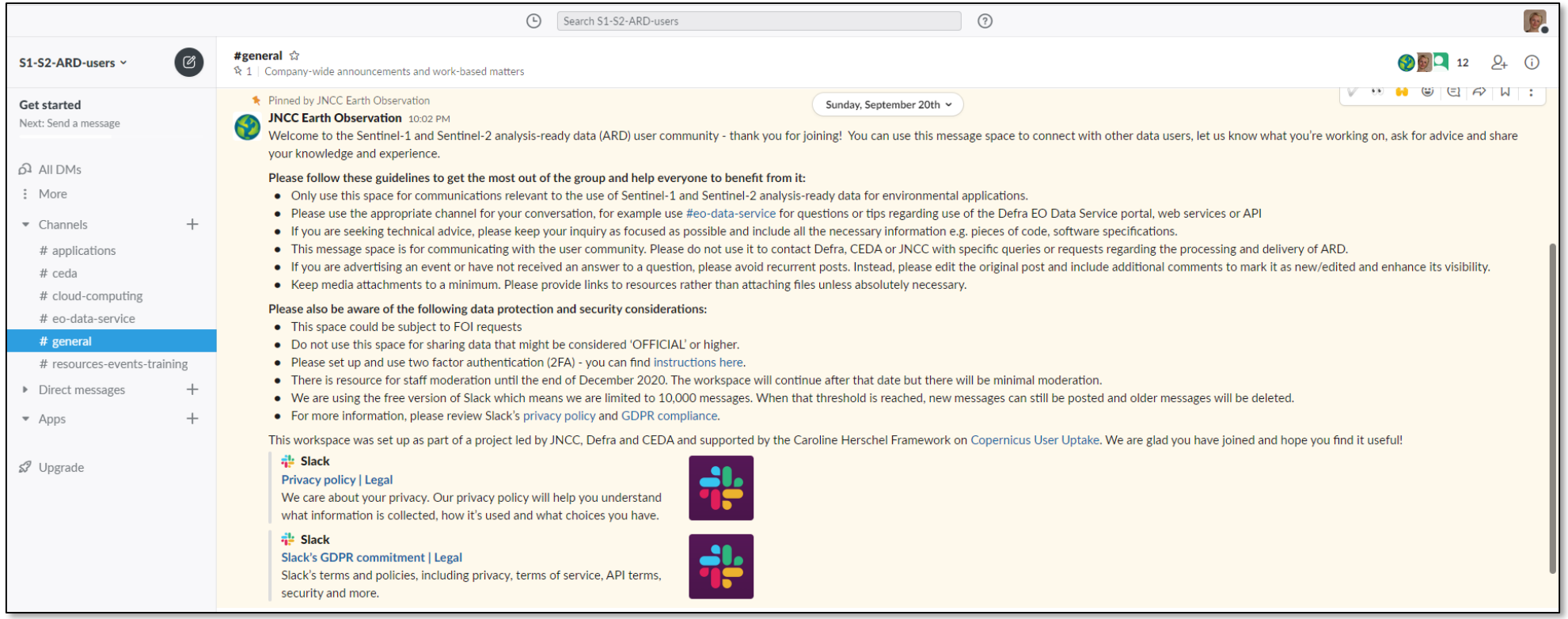
User Forum

The decision was taken to set up a **workspace on Slack**:

- Widely used, especially within Government / public sector
- Free plan available
- Easy to set up and invite members
- ‘Discoverable’ via URL
- Browser based or app
- Channels act as chat rooms for conversations on particular topics
- Allows private messaging as well as group chat
- Posts are archived and accessible e.g. search by keywords
- Use emojis and keywords as tags
- User controls notification settings

User Forum

You can find the workspace here:



If you have a Slack account:

https://S1-S2-ARD-users.slack.com/

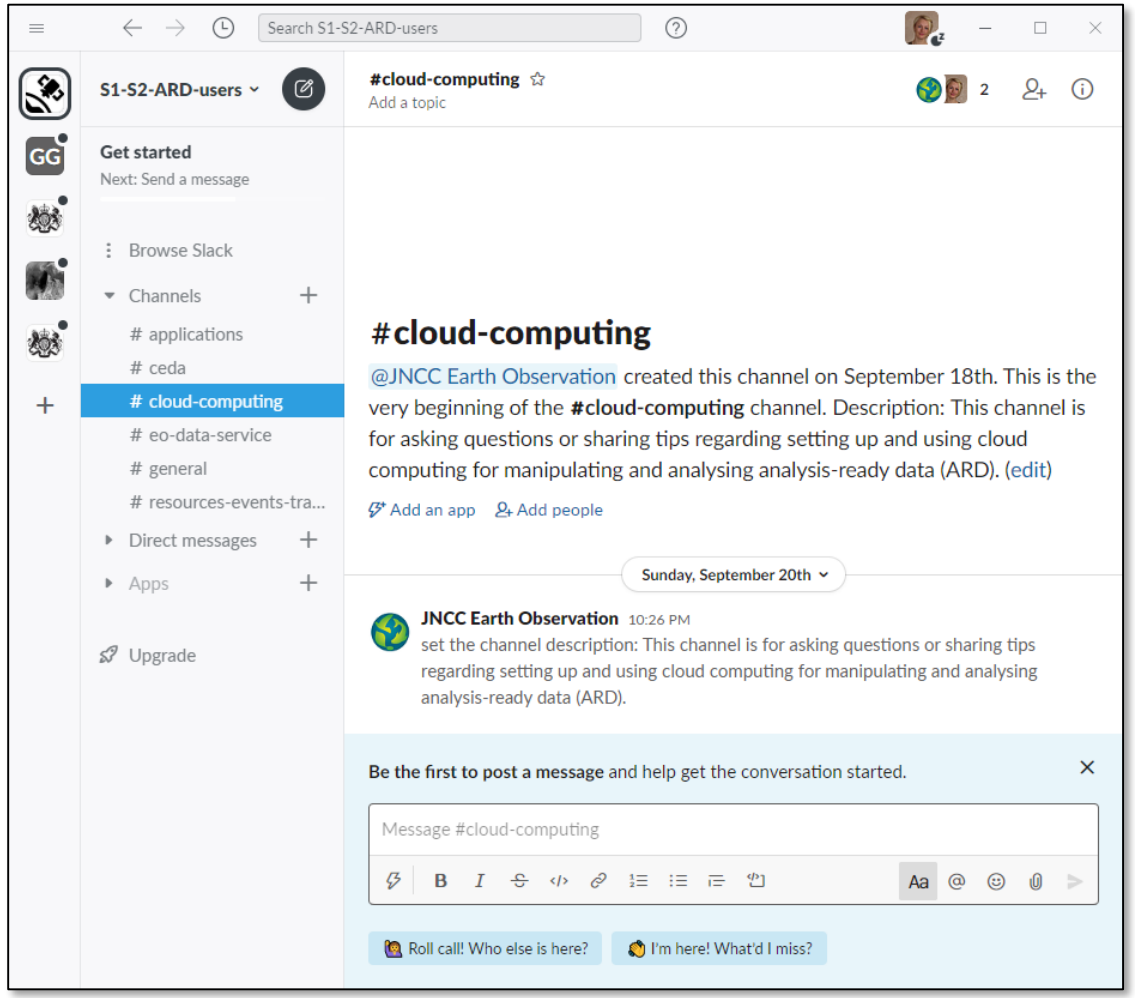
If you do not yet have a Slack account:

https://join.slack.com/t/s1-s2-ard-users/signup

User Forum

Launching Slack

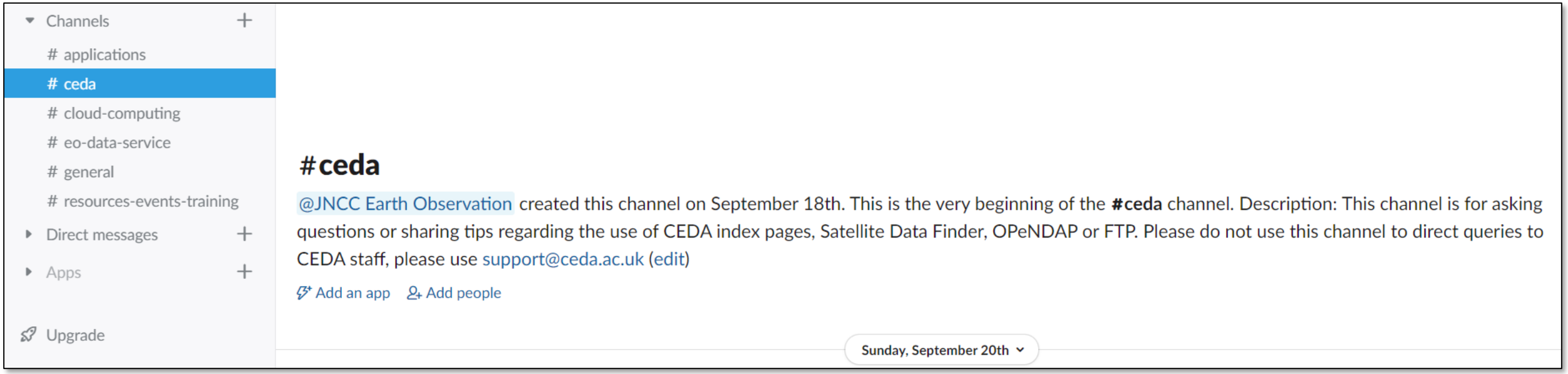
- Open in your browser (left) or use the Slack app (right). Please note IE11 is not supported.
- Slack app lets you have several workspaces open and switch between them.



User Forum

Structure

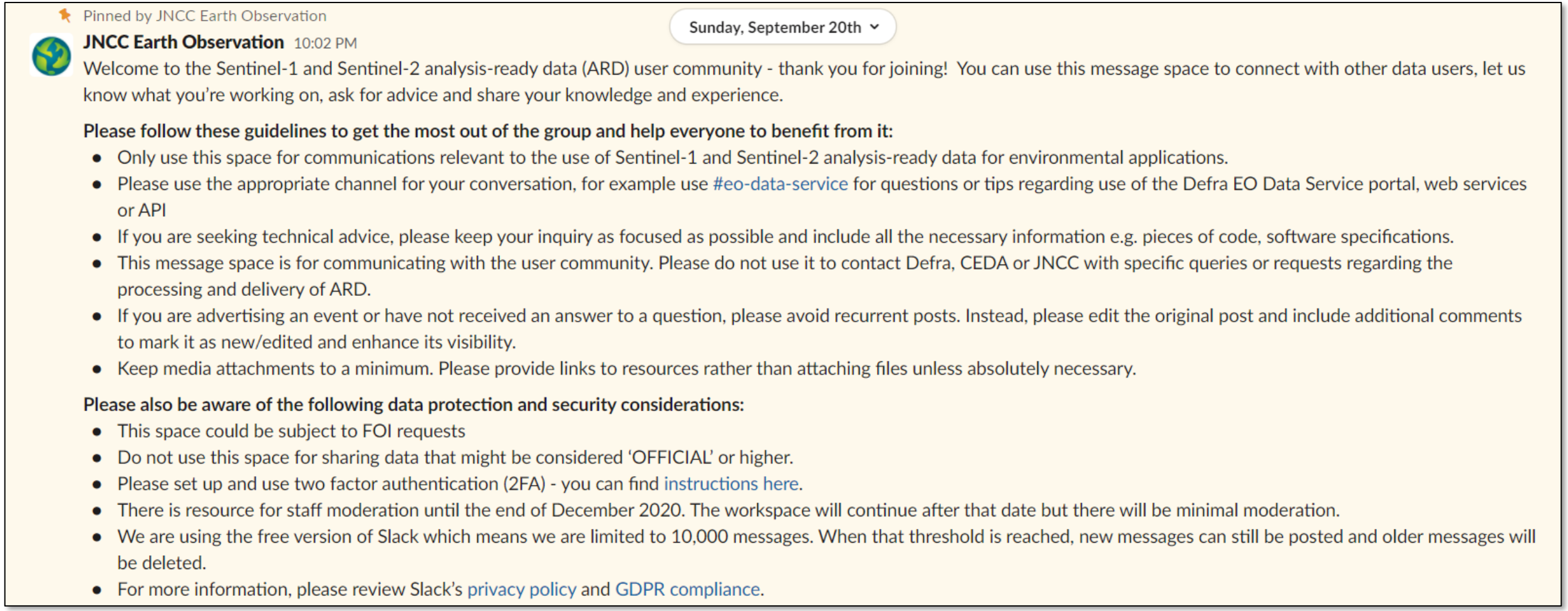
- There are several channels. Each channel has a description saying what it can be used for.
- Every new member is added to the **#general** channel, and you can sign up to others.
- We can create new channels if needed or delete any that aren't being used.
- We can pin posts at the top of channels.



User Forum

Guidelines on use and data security considerations

- Please read the pinned post in the #general channel before you start.



Pinned by JNCC Earth Observation

Sunday, September 20th

JNCC Earth Observation 10:02 PM

Welcome to the Sentinel-1 and Sentinel-2 analysis-ready data (ARD) user community - thank you for joining! You can use this message space to connect with other data users, let us know what you're working on, ask for advice and share your knowledge and experience.

Please follow these guidelines to get the most out of the group and help everyone to benefit from it:

- Only use this space for communications relevant to the use of Sentinel-1 and Sentinel-2 analysis-ready data for environmental applications.
- Please use the appropriate channel for your conversation, for example use #eo-data-service for questions or tips regarding use of the Defra EO Data Service portal, web services or API
- If you are seeking technical advice, please keep your inquiry as focused as possible and include all the necessary information e.g. pieces of code, software specifications.
- This message space is for communicating with the user community. Please do not use it to contact Defra, CEDA or JNCC with specific queries or requests regarding the processing and delivery of ARD.
- If you are advertising an event or have not received an answer to a question, please avoid recurrent posts. Instead, please edit the original post and include additional comments to mark it as new/edited and enhance its visibility.
- Keep media attachments to a minimum. Please provide links to resources rather than attaching files unless absolutely necessary.

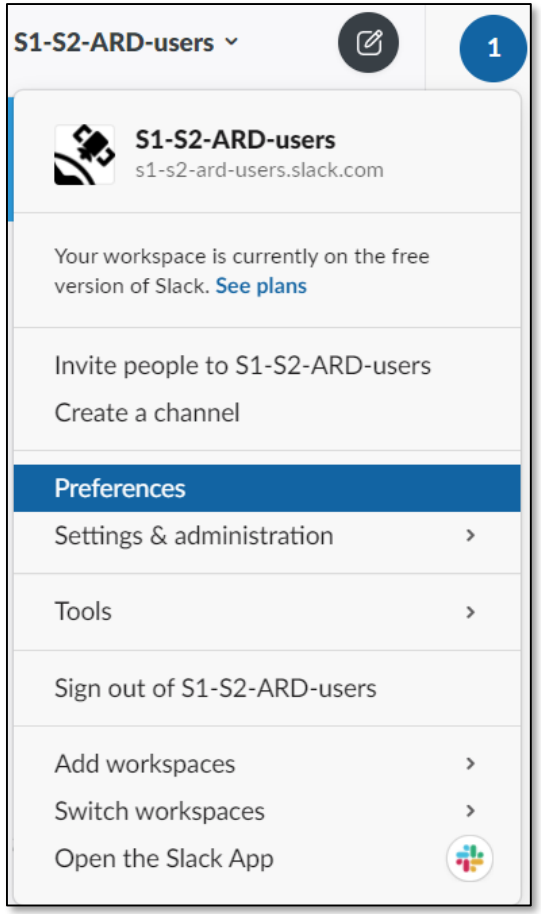
Please also be aware of the following data protection and security considerations:

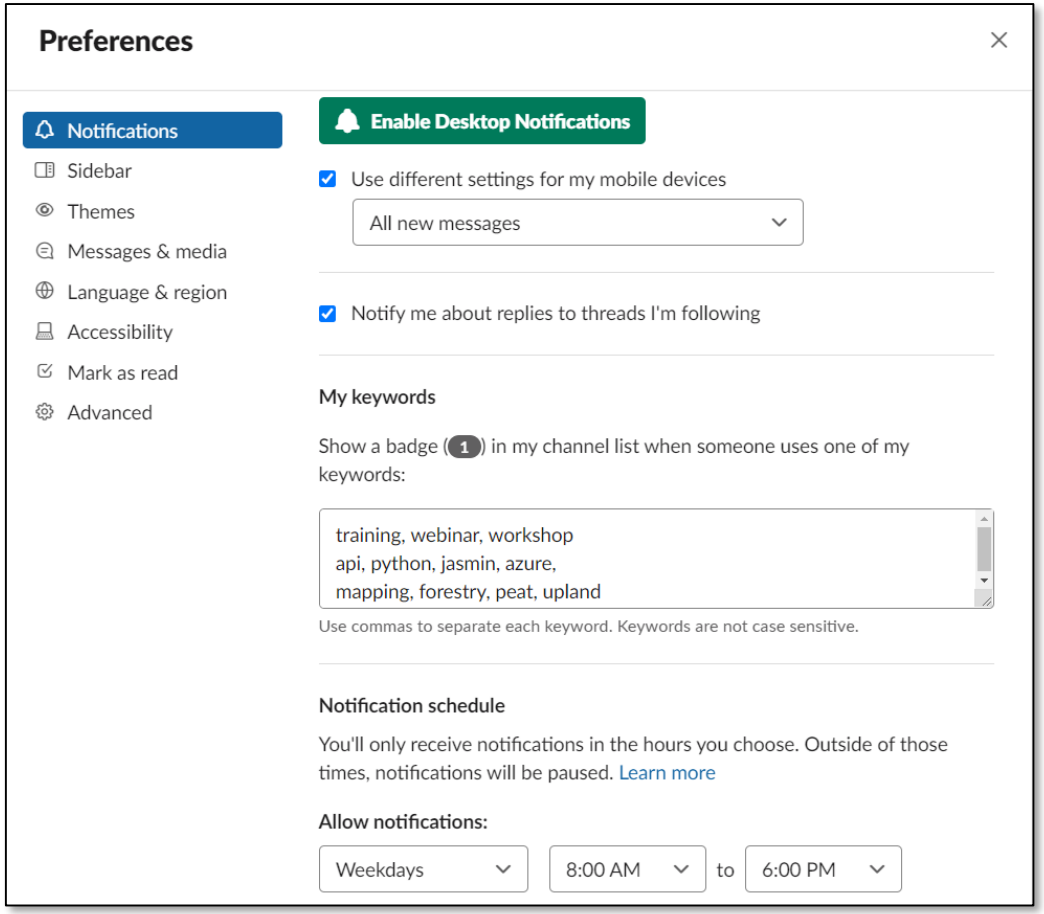
- This space could be subject to FOI requests
- Do not use this space for sharing data that might be considered 'OFFICIAL' or higher.
- Please set up and use two factor authentication (2FA) - you can find instructions here.
- There is resource for staff moderation until the end of December 2020. The workspace will continue after that date but there will be minimal moderation.
- We are using the free version of Slack which means we are limited to 10,000 messages. When that threshold is reached, new messages can still be posted and older messages will be deleted.
- For more information, please review Slack's privacy policy and GDPR compliance.

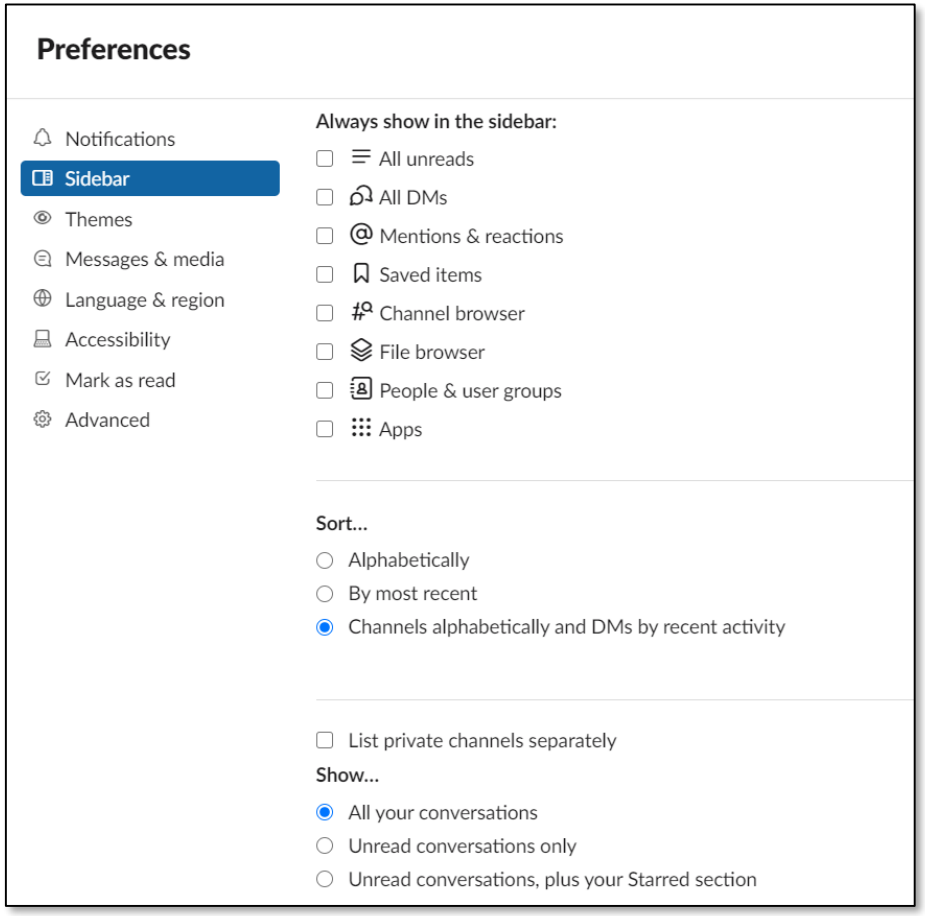
User Forum

Set your preferences

- Choose what type of notifications you receive, and when you receive them.
- Set keywords for topics which interest you.
- Choose what to display on the sidebar and how to organise it.





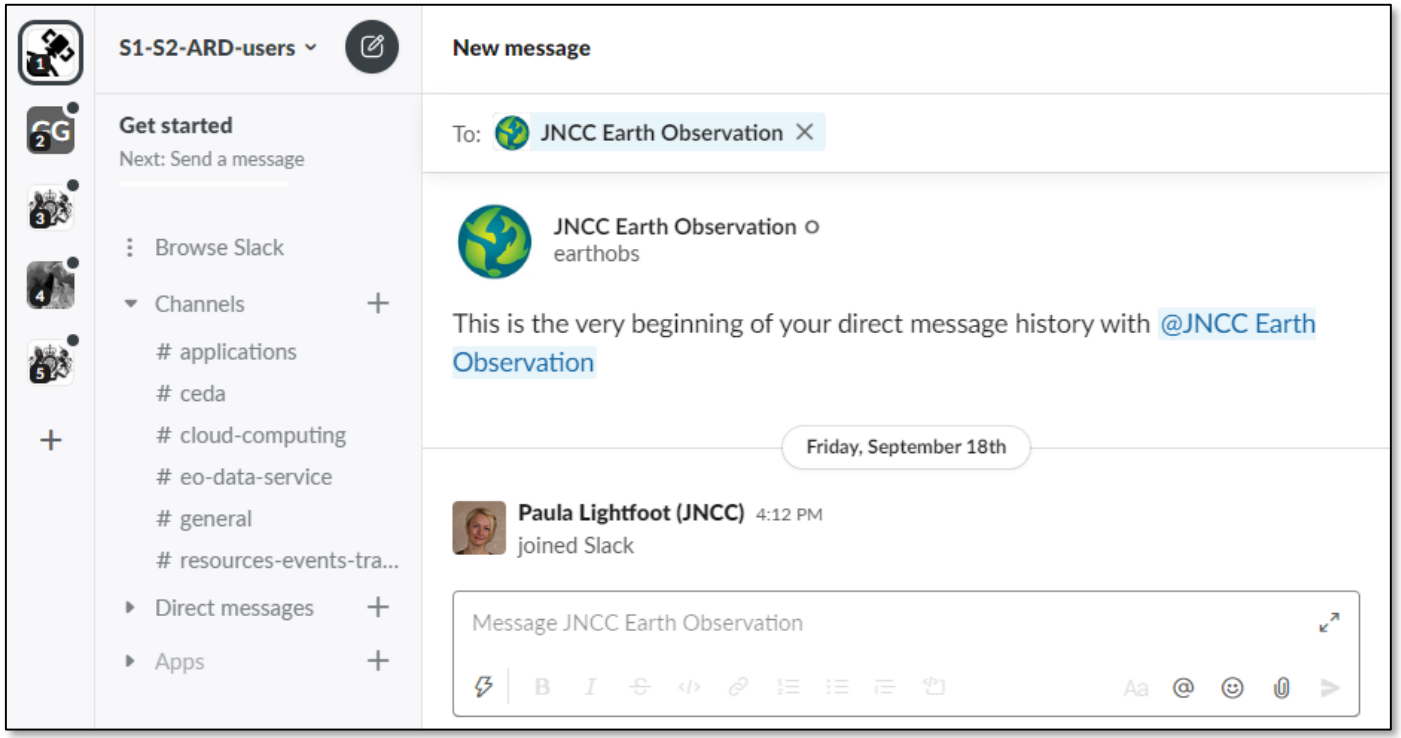


User Forum

Messaging

- Post in a channel. Use @ to tag someone. Press enter to send, shift+enter for new line.

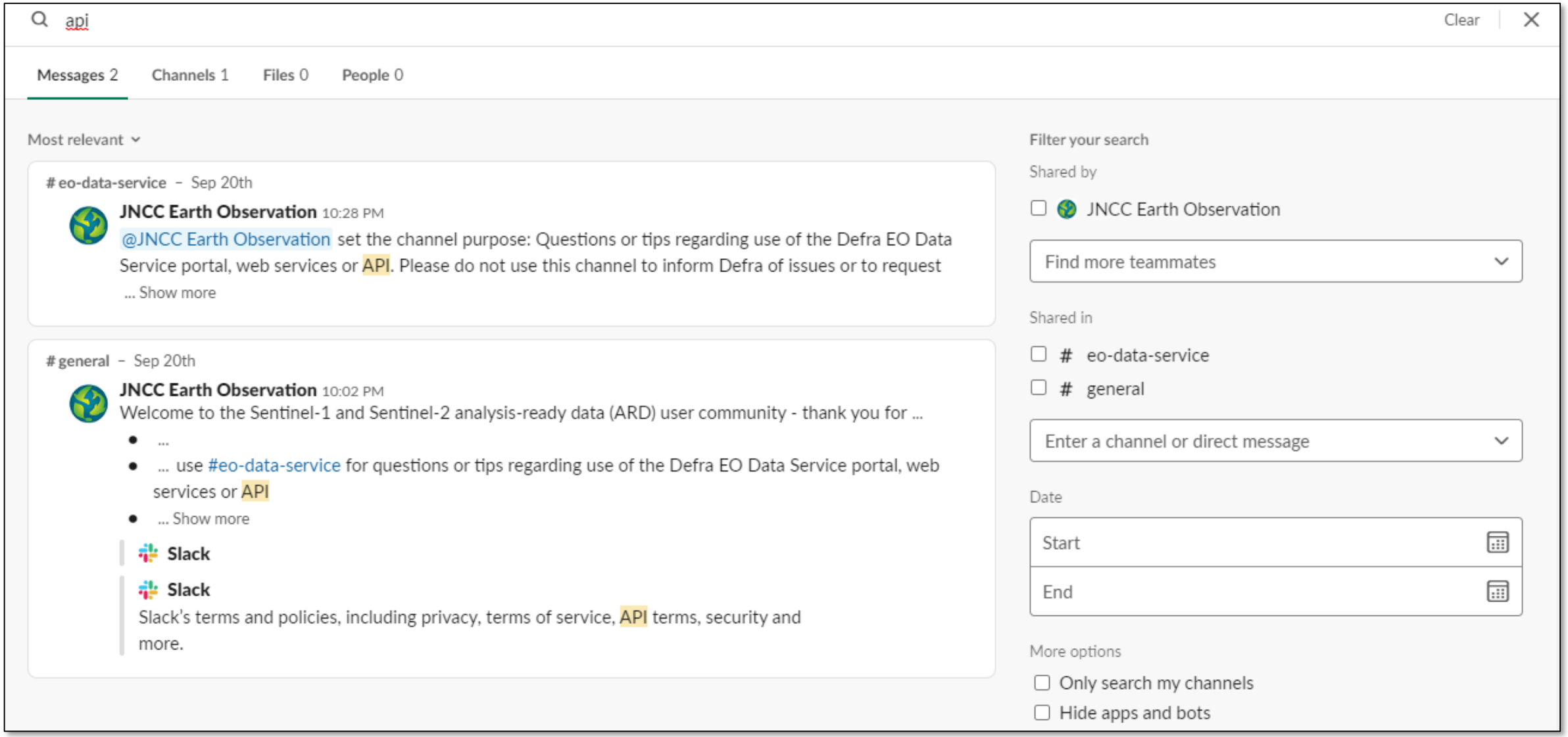
- Send a direct message to one or more members.



User Forum

Searching Slack

- Enter a term in the search box at the top of the page. Apply advanced filters if required.



User Forum

Membership

- Anyone with an e-mail address from the domains below can join the group without invitation.
- Others can join by invitation.
- Any member of the group can invite others to join (from any organisation).

ahdb.org.uk
apha.gov.uk
ceda.ac.uk
cefas.co.uk
cyfoethnaturiolcymru.gov.uk
daera-ni.gov.uk
defra.gov.uk
environment-agency.gov.uk
forestresearch.gov.uk
forestry.gov.scot
gov.scot
gov.wales
hes.scot,
historicengland.org.uk
jncc.gov.uk
marinemanagement.org.uk
naturalengland.org.uk
nature.scot
rpa.gov.uk
sepa.org.uk

Other domains can be added – just ask!

earthobs@jncc.gov.uk

User Forum

Important information

- The Slack free plan allows 10,000 posts and 5 GB file storage per workspace.
- When the 10,000 threshold is reached, the oldest messages will be deleted as new messages are posted.
- We can provide staff resource until the end of December 2020 for user engagement and light touch moderation. We hope it will be largely self-sustaining after that.
- If we can demonstrate impact, maybe we can make the case for resourcing the Slack standard plan and/or providing ongoing staff support if necessary.

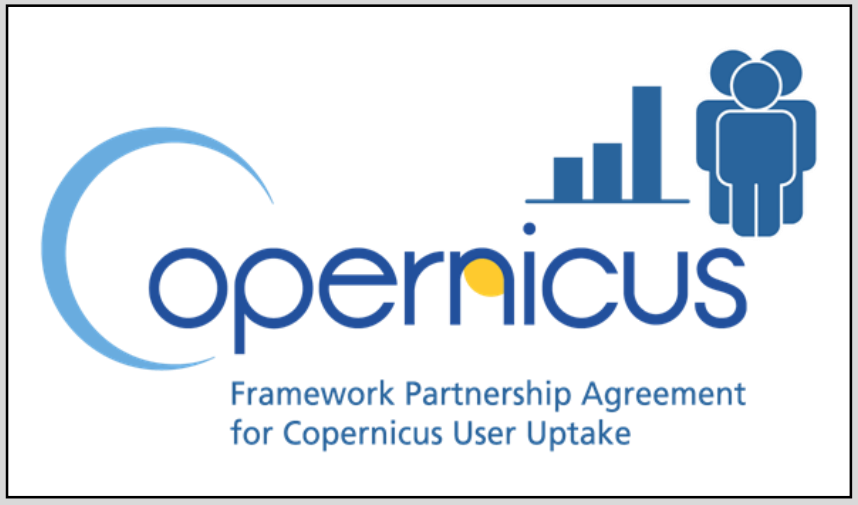
User Forum

Recap of best practice

- This is for peer-to-peer communication and support.
- Please be constructive and courteous in your messages.
- Ensure messages are on-topic.
- Join relevant channels and use the appropriate channel for your conversation.
- Include all necessary information if requesting technical advice.
- Be aware of data protection and security considerations and relevant policies.
- Do not post anything that might be considered ‘OFFICIAL’ or higher.
- Set your user preferences to manage notifications.
- Avoid recurrent posts and keep attachments to a minimum.

Next steps

- Promote both platforms more widely.
- Encourage and support participation.
- Monitor use and measure impact.
- Please help us by:
  - Joining the Slack workspace and starting/contributing to discussions.
  - Telling us about examples of code to add to the GitHub list, and sharing your own code.
  - Promote both platforms to colleagues and contacts.
  - Tell us if you have benefited from either platform e.g. used code, made new contacts etc.



Acknowledgements

- User survey respondents
- Andrew Cutts, acgeospatial
- All code contributors, especially Kostas Sideris and Becky Trippier (JNCC) and Sam Franklin (CGI)
- JNCC Data and Digital Services teams
- Defra EOCoE Implementation Group
- Scottish Remote Sensing Working Group

Sentinel-1 satellite © ESA

Thank you!

Any questions?

Sentinel-2 satellite © ESA