Cefas/DASSH OGC EDR API Pilot Project

Kate Collingridge & Rebecca Millard | Marine Data Scientists



Together we are working for **a** sustainable blue future

Centre for Environment Fisheries & Aquaculture Science



25 YEARS of Cefas 120 years of science

Pilot project

Aims:

- Test use of Cefas and DASSH OGC EDR APIs to access data programmatically
- Write R functions to query the APIs
- Create an interactive application to visualise data accessed via the APIs
- Provide feedback on how easy the APIs were to use



Querying the API

- Can use swagger to test queries
- Then learn how to construct query URLS.



OGC ENVIRONMENT DATA RETIEVAL API

/swagger/ogc/swagger.json

An Environmental Data Retrieval (EDR) API provides a family of lightweight interfaces to access Environmental Data resources, as per the standard defined by the Open Geospatial Consortium. Only a limited number of Cefas recordsets are available via this API.

OGC Environment Data Retrieval Standard Terms and Conditions Cookies Policy

Contact Cefas Data Manager

e.q.

https://www.dassh.ac.uk/edr/collections/dbossh/instances/DASSHDT00000078/ area?coords=POLYGON%28%28-33.31958%20-7.43087%2C-33.31958%2068.91161%2C57.73653%2068.91161%2C57.73653%20-7.43087%2C-33.31958%20-7.43087%29%29

eristics of the information available from the API

Functions

- R code to return usable data given basic inputs
- Used in Rshiny application
- E.g:
 - To construct a query url
 - To take a shape drawn

 on a map and convert it
 into a polygon query to
 return data within that
 shape.
 - To extract the spatial information for a dataset to plot it onto a map

```
retrieve_Cefas_OGC_data <- function(collectionId, guery_type = "Whole dataset", polygon_url)
 # define API link
 url <- "https://data-api.cefas.co.uk/ogc/"
 # create collection url and retrieve info
 url_collection <- pasteO(url, "collections/", collectionId)
 dat_collection <- httr::GET(url_collection)
 # get content and convert from json
 content_collection <- httr::content(dat_collection, as = "text")
 content collection <- isonlite::fromJSON(content collection)
 # get extent of spatial data (bounding box)
 bbox <- content_collectionSextentSspatialSboundingBox
 if(query_type == "Whole dataset") {
   # make the bounding box a polygon guery url for collection
   # create query URL with bounding box coordinates
   # note: list coordinates from the bottom left in clockwise order, repeat bottom left at the end
   dat_collection_all <- httr::GET(paste0(url_collection,</pre>
                                           "/area?coords=POLYGON",
                                          RCurl::curlEscape(paste0(
                                            bbox[1],
                                                          bbox 2
                                            bbox[1].
                                                          bbox[4]
                                            bbox[3].
                                                          bbox[4],
                                            bbox[3],
                                                          bbox[2].
                                                     11 11
                                            bbox[1],
                                                        , bbox[2],
                                            ")", ")")))))
 # get content and convert from json
 content_collection_all <- httr::content(dat_collection_all, as = "text")
 content_collection_all <- jsonlite::fromJSON(content_collection_all)</pre>
 # return features from data
 return(content_collection_all$features)
```

R Shiny App

• Shiny (rstudio.com)

- App allows users to query APIs and visualise data without needing technical expertise.
- Hosted on Cefas' rstudioconnect infrastructure.

C7862 API Demo

This app allows you to query datasets from the Cefas Data Hub and DASSH using the OGC ENV API

Cefas

This app uses 6 datasets from the Cefas Data Hub which have had the OGC ENV API applied. These datasets are stored as *collections*

Collection ID

11460 Seafloor Litter 🔹

Query type

Draw polygon •

Click on the pentagon on the left to draw a polygon on the map

Query Cefas API Your query returned 3807 rows of data

DASSH

DASSH is a data portal from the MBA. DASSH data also has the OGC ENV API applied. DASSH data is all stored in one collection called *dbossh*. It is not possible to query the whole collection at once, so select individual surveys (called instances) which are in the same area as your polygon.

-		9	Ogio	Sverige
C Î		LUnited Kingo Eire / Irela	Iord Danmark Imain Hamburg Londor Netierland @ Deutschla	Berlin ® Polska
how 2: area Greater	5 v e statio	ntries on fldHaulLatD 51	Leaflet © OpenStreetMap or Search: egrees fldHaulLatM 44.3	Slovensk Magyarorsz © Zagreb ontributors, CC-BY-SA inutes fldHa
North				
North Sea Greater North Sea	2	51	36.1	2
North Sea Greater North Sea Greater North Sea	2 3	51	36.1 49.4	2 3
North Sea Greater North Sea Greater North Sea Greater North Sea	2 3 4	51 51 52	36.1 49.4 49.4	2 3 2

 Differing Implementation – made writing standardised functions difficult

Cefas

Collection1 data Collection2 data Collection3 data

DASSH

Collection-dbossh Instance-survey1 data Instance-survey2 data Instance-survey3 data



- Timeouts data not returned, why?
- Data volume DASSH has a large amount of data so most queries take a long time or timeout.







• Blind querying – difficult to use an API if you don't already know what data it serves!

VS







• API not setup to return complete datasets easily

"Return the entire dataset in collection ID x" **is not** possible

... but there are workarounds!

"Get the spatial extent of dataset in collection ID x and use this to return the entire dataset in collection ID x" is possible



Thank you for listening

Follow @CefasGovUK



- Visit cefas.co.uk
- Subscribe to our newsletters

data.manager@cefas.co.uk (soon to be @cefas.gov.uk)

Together we are working for **a sustainable blue future**

Centre for Environment Fisheries & Aquaculture Science

