

Discovering Earth Observation data

Despina-Athanasia Pantazi

**Knowledge Technologies
2022-23**

Using Google for dataset discovery

- Is there a land cover dataset produced by the European Environmental Agency covering the area of Chania, Crete, Greece?
- Google it!



Results

Google

Is there a land cover dataset produced by the European Environmental Agency

All Images News Videos Maps More Settings Tools

About 33,500 results (0.55 seconds)

(PDF) Creation of a land cover map of Crete, using spot satellite data
https://www.researchgate.net/publication/242238279_Creation_of_a_land_cover_map_of_Crete ... PDF | The aim of this work was to create a Land Cover map of Crete on a cartographic scale of 1:50000 ... Article (PDF Available) · January 2002 with 46 Reads ... environment for Member States of the European ... in the CORINE project and the Greek team produced ... covering the area of Rethymnon) was experientially.

[PPT] pptx
<cgi.di.uoa.gr/~koubarak/talks/manolis-koubarakis-talk-fraunhoferAIS.pptx> ▾
Lots of public sector data has been made open and freely available recently through various government ... Question: Is there a land cover dataset produced by the European Environmental Agency covering the area of Chania, Crete, Greece?

EEA land cover data to be ... - European Environment Agency
<https://www.eea.europa.eu/en/land-use/eea-land-cover-data-to-be-used-in-mobile-phone-maps> ▾
Dec 13, 2012 - Data on land use provided by the European Environment Agency ... The Corine dataset will improve mapping and navigation with its ... but at a later stage it may be used to also identify other land cover categories such as agricultural land ... Geographic coverage ... Nationally designated areas (CDDA) ...
Missing: ehania crete greece

Mapping sensitivity to desertification in Crete (Greece), the risk for ...
<https://waponline.com/jwcc/article/9/4/.../Mapping-sensitivity-to-desertification-in-Crete>
by GG Morianou - Cited by 1
Sep 3, 2018 - The Environmental Sensitivity Area (ESA) output is an indicator system producing ... It also influences the effects of chemical amendments, fertilizers, ... (2014) and the European Soil Database, soils in Crete are generally ... In terms of vegetation, Crete is mostly covered by natural grasslands and pastures.

Share Get App corine land cover greece download Download Link ...
<https://imgur.com/a/gCvWuUS/embed?pub=true> ▾
The Corine Land Cover project22 is an activity of the European Environment Agency ... the European Environment Agency that The land cover of Greece is available as an ... Greek Administrative Geography Dataset (download); CORINE Land ... Urban land covers producing the highest surface temperatures (hot spots) are ...

Let us pose a different query

- Is there a land cover dataset produced by the European Environment Agency?



Results



Is there a land cover dataset produced by the European Environment Agency? ✕ 🔍

[🔍 All](#) [🖼️ Images](#) [📰 News](#) [📺 Videos](#) [📍 Maps](#) [⋮ More](#) [Settings](#) [Tools](#)

About 393,000,000 results (0.71 seconds)

The Corine databases and several of its programmes have been taken over by the EEA. One of these is an inventory of **land cover** in 44 classes, and presented as a cartographic product, at a scale of 1:100 000. This **database** is operationally **available** for most areas of **Europe**.

www.eea.europa.eu > ... > CORINE Land Cover

[CORINE Land Cover — European Environment Agency](#)

[? About Featured Snippets](#) [Feedback](#)

www.eea.europa.eu > data-and-maps > data > copernicu... ▾

[Copernicus Land Monitoring Service - Corine Land Cover ...](#)

[European Environment Agency](#) Toggle navigation ... The [CORINE Land Cover](#) (CLC) inventory was initiated in 1985 (reference year 1990) to standardize **data** collection on land in Europe to support environmental policy development. Updates were ... Change layers were **produced** for 2000, 2006, 2012 and 2018. **Data** and ...

www.eea.europa.eu > data-and-maps > dashboards > la... ▾

[Land cover and change statistics 2000-2018 — European ...](#)

Oct 14, 2020 — Information on the **environment** for those involved in developing, ... This interactive **data** viewer provides an easy and comprehensive access to land ... The viewer facilitates the assessment of **land cover** consumed or **created** over a ... The understanding of the implications of changes in **land cover** and land ...

Google Dataset Search



Dataset Search

corine land cover greece



Try [coronavirus covid-19](#) or [education outcomes site:data.gov](#).

[Find out more](#) about including your datasets in Dataset Search.

Results

Google ✕ ? ! ☰ [Sign in](#)


▼ Last updated ▼ Download format ▼ Usage rights ▼ Topic Free Saved datasets

32 datasets found

CORINE Land Cover

Explore at the Datahub Explore at the Datahub Explore at data.wu.ac.at

5 scholarly articles cite this dataset ([View in Google Scholar](#))

 bz

Dataset updated Apr 7, 2014

Dataset provided by
TELEIOS


License
Other (Open)

Description

The CORINE Land Cover project is an activity of the European Environment Agency that collects data regarding the land cover of European countries. The project uses a hierarchical scheme with three levels to describe land cover. The first level indicates the major categories of land cover on the planet e.g., agricultural areas. The second level identifies more specific types of land cover, e.g., permanent crops, while the third level narrows down to a very specific characterization e.g., olive groves. The land cover of Greece is available as an ESRI shapefile that is based on this classification scheme. This shapefile is transformed in RDF as follows. Every land cover is represented with a class, e.g. ConiferousForest, and the hierarchy of land covers is expressed with the respective class taxonomy. For each specific area in the shapefile a unique URI is created and it is connected with an instance of the third level. Additionally, a data property of each area indicates its geometry. Some sample triples representing such an area are shown in the following triples: corine:Area_45 a corine:Area; strdf:hasGeometry "POLYGON((22.07 40.62, ..., 22.07 40.62))"^^strdf:WKT; corine:hasLandUse corine:coniferousForest.

CORINE Land Cover


datahub.io
datahub.ckan.io
+1 more

 bz

Updated Apr 7, 2014

Corine Land Cover 1990 (raster 100m) - version 2020_20u1, May...


sdi.eea.europa.eu
geodcat-ap.semic.eu

 +3

Updated Sep 8, 2017

Corine Land Cover

ai.di.uoa.gr

 7z

Google Dataset Search

- Datasets that are indexed using **schema.org**, as proposed by Google, show up.
- Enables users to find datasets stored across the Web by doing a simple keyword search.
- Uncovers information about datasets hosted in thousands of repositories across the web, making these datasets universally accessible and useful.

How does Dataset Search find datasets?

Authors need to add metadata in **schema.org** to each page that describes a dataset.

Schema.org:

- Founded by Google, Microsoft, Yahoo!, Yandex. Currently schema.org vocabularies are developed by an open community process.
- Provides a unique structured data markup schema to annotate a variety of topics.
 - Tags added to HTML as JSON-LD, Microdata, or RDFa.
- On-page markup allows search engines to understand information included in web pages.



Schema.org


```
<script type="application/ld+json">{
  "@context": "http://schema.org",
  "@type": ["ItemList", "Dataset"],
  "itemListOrder": "http://schema.org/ItemListOrderAscending",
  "numberOfItems": "7",
  "itemListElement": [{
    "@type": "ListItem", "position": 1,
    "item": {
      "@type" : "Dataset",
      "name": "GADM database of Global Administrative Areas",
      "alternateName": "GADM",
      "description": "GADM is a spatial database of the location ...",
      "author": "University of Athens",
      "sourceOrganization": "Robert Hijmans, in collaboration with ...",
      "copyrightYear": "2018",
      "keywords":["GADM", "Global Administrative Areas", "GADM 2015"],
      "spatialCoverage": "World", "temporalCoverage": "2015", "fileFormat": "7z",
      "isBasedOn": "GADM database of Global Administrative Areas 2015, Version 2.8",
      "isAccessibleForFree": true,
      "distribution": {
        "@type":"DataDownload", "encodingFormat":"7z",
        "contentUrl":"https://datahub.ckan.io/dataset/gadm"},
      "url": "https://datahub.ckan.io/dataset/gadm"}}}, ... ]} </script>
```


Google Dataset Search



Google ✕ i ! ☰ [Sign in](#)

▼ Last updated ▼ Download format ▼ Usage rights ▼ Topic Free [Saved datasets](#)

 **GADM database of Global Administrative Areas**
ai.di.uoa.gr
data.apps.fao.org
 7z


 **South Africa Admin Boundaries level I**
rcmr.d.africageoportal.com
Updated May 29, 2020

 **Kazakhstan District Boundary**
hub.arcgis.com
Updated Aug 7, 2016

 **Protected Areas and Deforestation: New Results from...**
datacatalog.worldbank.org
 excel
Updated Mar 12, 2018

GADM database of Global Administrative Areas
GADM

[Explore at ai.di.uoa.gr](#) [Explore at data.apps.fao.org](#)

 7z



Dataset provided by
Robert Hijmans, in collaboration with colleagues at the University of California, Berkeley Museum of Vertebrate Zoology (Julian Kapoor and John Wieczorek), the International Rice Research Institute (Nel Garcia, Aileen Maunahan, Arnel Rala) and the University of California, Davis (Alex Mandel), and with contributions of many others.

Authors
University of Athens

Time period covered
2015

Area covered
World

Description
GADM is a spatial database of the location of the world's administrative areas (or administrative boundaries) for use in GIS and similar software.

11

Google Dataset Search by the Numbers

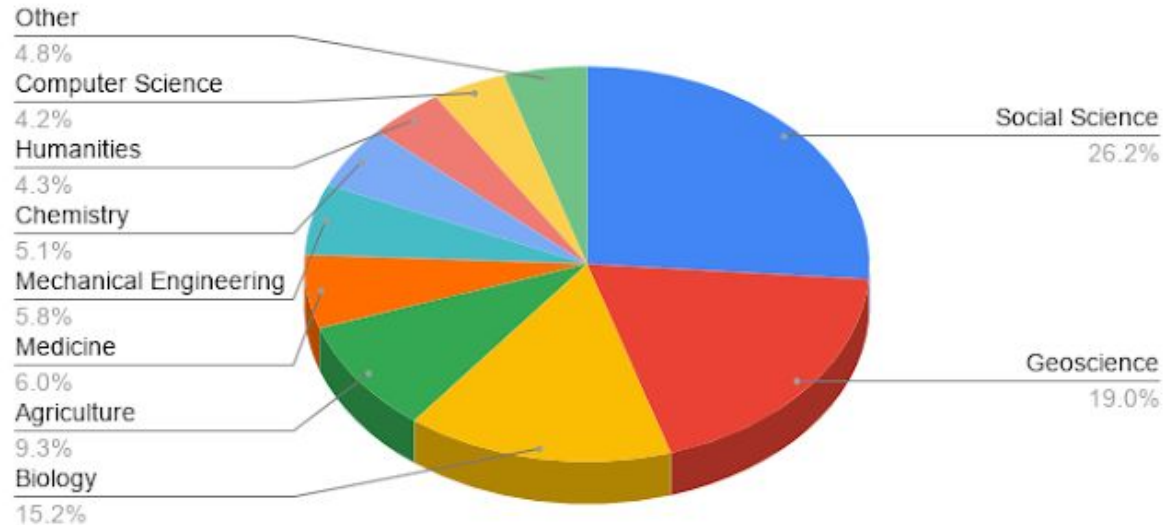
- ❖ O. Benjelloun, S. Chen, N. Noy: **Google Dataset Search by the Numbers** (Jun 2020)
<https://arxiv.org/abs/2006.06894>
- As of March 2020, the corpus contained 28 million datasets from more than 3,700 sites
- The corpus is a reasonably representative snapshot of the datasets published on the Web, but there is no way of measuring how well the corpus covers all the datasets available on the Web

Google Dataset Search by the Numbers

- ❖ O. Benjelloun, S. Chen, N. Noy: **Google Dataset Search by the Numbers** (Jun 2020)
<https://arxiv.org/abs/2006.06894>
 - **Licenses and access:** Only 34% of the datasets provide any licensing information
 - most of datasets available for free, almost always allowed reuse for both commercial and non-commercial purposes
 - **Linked Data:** **Fewer than 1%** of datasets in the corpus are in linked data formats
 - there is plenty of shared data that the Semantic Web community produces, but the final step of describing it appears to be less common

Google Dataset Search - An analysis of online datasets

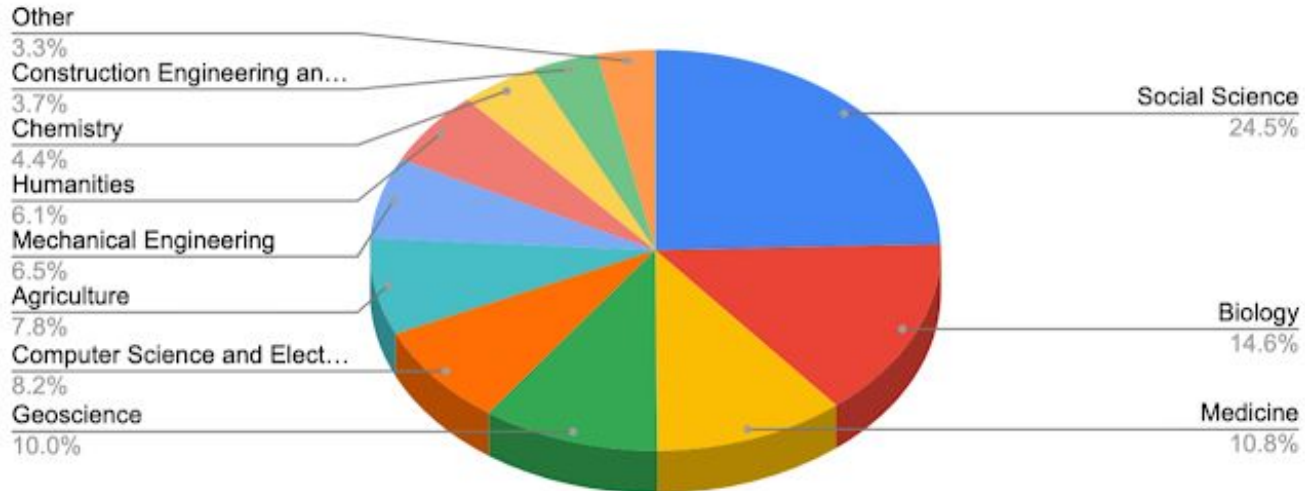
❖ Distribution of dataset topics



Reference: <https://ai.googleblog.com/2020/08/an-analysis-of-online-datasets-using.html>

Google Dataset Search - An analysis of online datasets

❖ What do users access?



Reference: <https://ai.googleblog.com/2020/08/an-analysis-of-online-datasets-using.html>

Google Dataset Search

- ❖ T. Alrashed, D. Paparas , O. Benjelloun, Y. Sheng, and N. Noy: **Dataset or Not? A Study on the Veracity of Semantic Markup for Dataset Pages** (October 2021) <https://research.google/pubs/pub50547/>
 - **Schema.org** has become **prevalent** on the Web as a way to express the semantics of Web page content
 - it is present on more than **30%** of Web pages
 - We cannot always take **Schema.org/Dataset** markup at face value
 - pages may include this markup **erroneously** or for the purposes of search-engine optimization

Let us pose our original query to Dataset Search

Google

Is there a land cover dataset produced by the Eur X



Sign in

Saved datasets

Your search - Is there a land cover dataset produced by the European Environment Agency covering the area of Chania, Crete, Greece? - did not match any datasets.

Suggestions:

- Make sure all words are spelled correctly.
- Try different keywords.
- Try more general keywords.
- Try fewer keywords.

[Learn](#) how you can add new datasets to our index.

EO dataset search - Copernicus Open Access Hub

- Provides complete, free and open access to Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P user products (<https://scihub.copernicus.eu>)

The screenshot displays the Copernicus Open Access Hub search interface. The top navigation bar includes the ESA and Copernicus logos, the text "Copernicus Open Access Hub", and user icons. Below the navigation bar is a search bar with the placeholder text "Insert search criteria...".

The left sidebar contains search filters:

- Sort By:** Ingestion Date (dropdown)
- Order By:** Descending (dropdown)
- Sensing period:** 2020/10/05 (calendar icon)
- Ingestion period:** 2020/10/29 (calendar icon)
- Mission:** Sentinel-1 (checked)
- Satellite Platform:** (dropdown)
- Polarisation:** (dropdown)
- Relative Orbit Number (from 1 to 175):** (input field)
- Mission:** Sentinel-2 (unchecked)
- Satellite Platform:** (dropdown)
- Relative Orbit Number (from 1 to 143):** (input field)
- Product Type:** (dropdown)
- Sensor Mode:** (dropdown)
- Collection:** (dropdown)
- Product Type:** (dropdown)
- Cloud Cover % (e.g.[0 TO 9.4]):** (input field)

The main area features a map of Europe with various cities and countries labeled. The map is overlaid with a grid and has a search bar in the top right corner. The map shows the search results for the specified criteria, with a focus on the European continent.

EO dataset search - Copernicus Open Access Hub

- Maintains an online archive of at least the latest month of products for immediate download via HTTP
- Synchronous access to latest data, asynchronous access to historic data

The screenshot shows the Copernicus Open Access Hub interface. The top navigation bar includes the ESA and Copernicus logos, the title 'Copernicus Open Access Hub', and user icons. A search bar is present with the placeholder text 'Insert search criteria...'. Below the search bar, the results are displayed in a list format. The first result is for 'S1A_IW_GRDH_1SDH_20201029T202908_20201029T202925...'. Each result includes a small thumbnail image, the product ID, a download URL, and mission/instrument details. The right side of the page features a map of the world with several red rectangular overlays indicating the geographic locations of the data products. The map includes labels for various countries and regions such as Greenland, Iceland, France, Spain, Morocco, Senegal, Nigeria, Ethiopia, and South Africa. At the bottom of the results list, there is a pagination control showing 'Products per page: 150' and 'page: 1 of 640'.

EO dataset search - EOWEB[®] GeoPortal

- A multi-mission web portal for interactive access to the German Aerospace Center (DLR) Earth observation data holdings
 - Combines classic discovery and order services for data held in the German Satellite Data Archive (D-SDA) with browse and download features via interoperable, OGC-compliant visualization and download services (<https://eoweb.dlr.de/egp/>)

EO dataset search - EOWEB[®] GeoPortal

The screenshot displays the EOWEB[®] GeoPortal search interface. At the top, the logo 'EOC EOWEB[®] GeoPortal' is on the left, and 'Logged in as: guest' and 'Help' are on the right. A navigation bar includes 'Home', 'Collections', 'Products', and 'Maps'. Below this is a 'Filter Management' section with active filters for 'Keyword', 'Time', and 'Region'. The main filter area is divided into four panels: 'Filter by Keywords' (with 'Sentinel-2' entered), 'Filter by Region' (with a map of Europe and a 'Set Region' button), 'Filter by Time' (with 'Starttime' set to 2020-06-01 and 'Endtime' set to 2020-10-30), and 'Filter by Type' (with 'Dataset' selected). Below the filters, the results are sorted by 'Modification D...'. The results list includes:

- Sentinel-2 MSI - Level 3A (MAJA/WASP) - Germany**: This collection contains synthesized Sentinel-2 Level 3A surface reflectances for Germany on a monthly basis computed by the WASP processor (which utilizes L2A products derived from the MAJA processor).
- Sentinel-2 MSI - Level 2A (MAJA Tiles) - Germany**: This collection contains Sentinel-2 Level 2A surface reflectances, which are computed for the country of Germany using the time-series based MAJA processor. During the Level 2A processing, the data are
- Sentinel-2 MSI - Vegetation Indices (AGRO-DE) - Germany, Monthly Composites**: This product comprises monthly composites and temporal statistics of selected vegetation indices (VI) for all of Germany from 2015 to today in 10m resolution, which were calculated using the DLR TimeScan
- Sentinel-2 MSI - Vegetation Indices (AGRO-DE) - Germany, Yearly Composites**: This product comprises yearly composites and temporal statistics of selected vegetation indices (VI) for all of Germany from 2015 to today in 10m resolution, which were calculated using the DLR TimeScan

On the right side of the results, there is a map showing a red bounding box over Germany, with a zoomed-in satellite image of a region below it.

EO dataset search - Earthdata Search

- Earthdata Search enables data discovery and access to more than 33,000 EO data collections from NASA's Earth Observation System Data and Information System (EOSDIS), U.S. and international partner agencies (search.earthdata.nasa.gov/)
 - Users can:
 - **Search for EO data:** Earthdata Search uses the Common Metadata Repository (CMR) for sub-second search
 - **Preview EO data:** Using Global Imagery Browse Services (GIBS), enables high-performance data visualization
 - **Download and access EO data:** In addition to direct download, surfaces OPeNDAP services for simpler spatial and parameter subsetting

EO dataset search - Earthdata Search

- Many of the imagery layers are updated daily & are available within 3 hours of the observation

The screenshot displays the Earthdata Search web application. On the left is a navigation sidebar with the NASA Earthdata Search logo and various filter categories: Spatial (with a 'Rectangle' selection), Features (Map Imagery, Near Real Time, Customizable), Keywords, Platforms, Instruments, Organizations, Projects, and Processing Levels. The main content area shows search filters: 'Sort by: End Date', 'Only include collections with granules' (checked), 'Include non-EOSDIS collections' (checked), and 'Advanced Search'. A tip suggests adding collections to a project for comparison. Below the filters, it indicates 'Showing 40 of 4,350 matching collections'. A list of four datasets is shown, each with a thumbnail, title, description, and 'MAP IMAGERY' and 'LP DAAC' tags. The datasets are: 1) MODIS/Terra Surface Reflectance Daily L2G Global 1km and 500m SIN Grid V006 (7,481 Granules, 2000-02-24 ongoing); 2) MODIS/Terra Surface Reflectance Daily L2G Global 250m SIN Grid V006 (7,481 Granules, 2000-02-24 ongoing); 3) MODIS/Terra Surface Reflectance 8-Day L3 Global 500m SIN Grid V005 (950 Granules, 2000-02-24 ongoing); 4) MODIS/Terra Geolocation Fields 5-Min L1A Swath 1km (19,346 Granules, 2000-02-23 ongoing). On the right, a map of the Mediterranean region is shown with a blue rectangle highlighting the area around Athens, Greece. The map includes labels for countries like Serbia, Bulgaria, Greece, and cities like Sofia, Thessaloniki, and Athens. A scale bar and 'Earthdata Login' button are also visible.

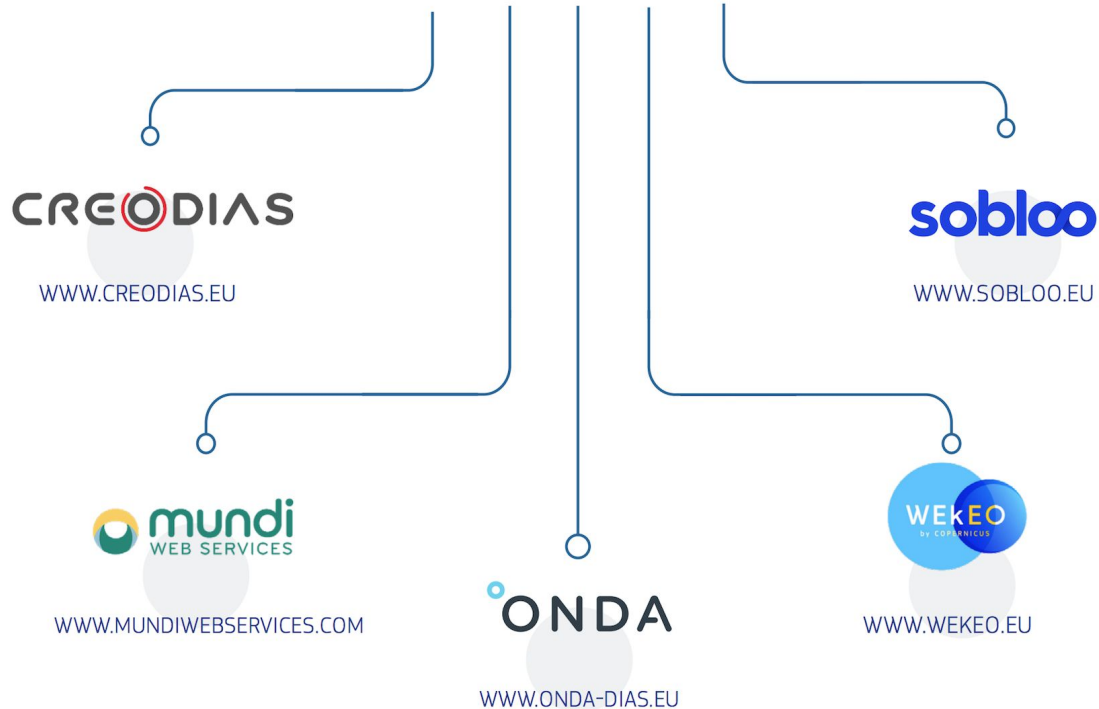
Copernicus Data and DIAS

Data and Information Access Services (DIAS):

- Five competitive platforms for quick access to a huge resource of Earth Observation data (satellite imagery) and Copernicus themed services.
- Easy and quick search, processing and sharing of satellite data.
- Quick access to satellite imagery via virtual machines.
- Data is free but computing power on a DIAS platform needs to be paid for.

Copernicus Data and DIAS

THE DIAS & WHERE TO REACH THEM



CREODIAS - Uses linked data in its catalogue!

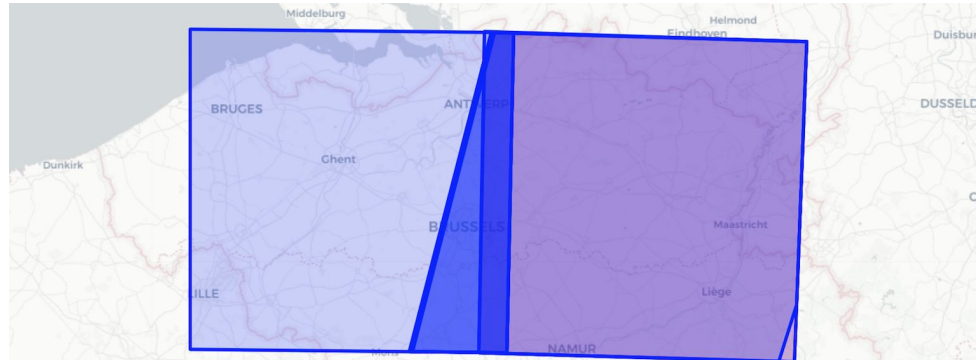
SPARQL interface: provides extended search capabilities for linking metadata of all products stored in the repository with various information from the Internet

Example Query: Find all Sentinel-2 images in the area of Brussels

SPARQL Query:

```
SELECT DISTINCT ?title ?geom {
  ?feature <http://ws.eodias.eu/metadata/attribute#title> ?title.
  ?feature <http://ws.eodias.eu/metadata/attribute#mission>
<http://ws.eodias.eu/metadata/mission/ Sentinel-2> .
  ?hex <http://ws.eodias.eu/metadata/attribute#feature> ?feature.
  ?hex <http://ws.creodias.eu/metadata/object/settlement >
    <http://dbpedia.org/resource/ Brussels> .
  ?feature <http://ws.eodias.eu/metadata/attribute#geometry> ?geom .
} LIMIT 100
```

CREODIAS



Found 75 results

<< page 1 of 8 >>

title	geom
"S2B_MSIL1C_20180403T104019_N0206_R008_T31UFS_20180403T132754.SAFE" << http://www.w3.org/2001/XMLSchema#string >>	"POLYGON ((4.435748354912912 51.33997008701231, 4.468636264611875 51.4195692531676, 4.477748734669471 51.441613138005, 6.017025008525815 51.41234265506691, 5.953869713140706 50.4262936061962, 4.408715109637645 50.45527216066188, 4.435748354912912 51.33997008701231))" << http://www.opengis.net/ont/geosparql#wktLiteral >>
"S2B_MSIL1C_20180423T104019_N0206_R008_T31UES_20180423T110230.SAFE" << http://www.w3.org/2001/XMLSchema#string >>	"POLYGON ((4.073553537257442 50.45666472371114, 4.091206001706374 50.50070662033345, 4.149881201906933 50.64667849413536, 4.208763333989855 50.79266510091305, 4.268072654378468 50.93859130689168, 4.327401261270531 51.08455377042205, 4.387143604179749 51.23043754970608, 4.446993597869525 51.37627878694308, 4.473671406372481 51.44125427635581, 4.579544021469633 51.44054116582309, 4.546436441462649 50.45352327891901, 4.073553537257442 50.45666472371114))" << http://www.opengis.net/ont/geosparql#wktLiteral >>

Current work

- We are developing an **extension to schema.org for Earth Observation data (eo.schema.org)**.
- Main idea:
 - Use OGC 17-003: EO Dataset Metadata GeoJSON(-LD) Encoding Standard
- We are also developing an **annotation tool** that can be used to annotation Earth Observation datasets using this extension.