

Linking and Finding Earth Observation (EO) Data on the Web

MACS-G20 Workshop: Linked Open Data in Agriculture

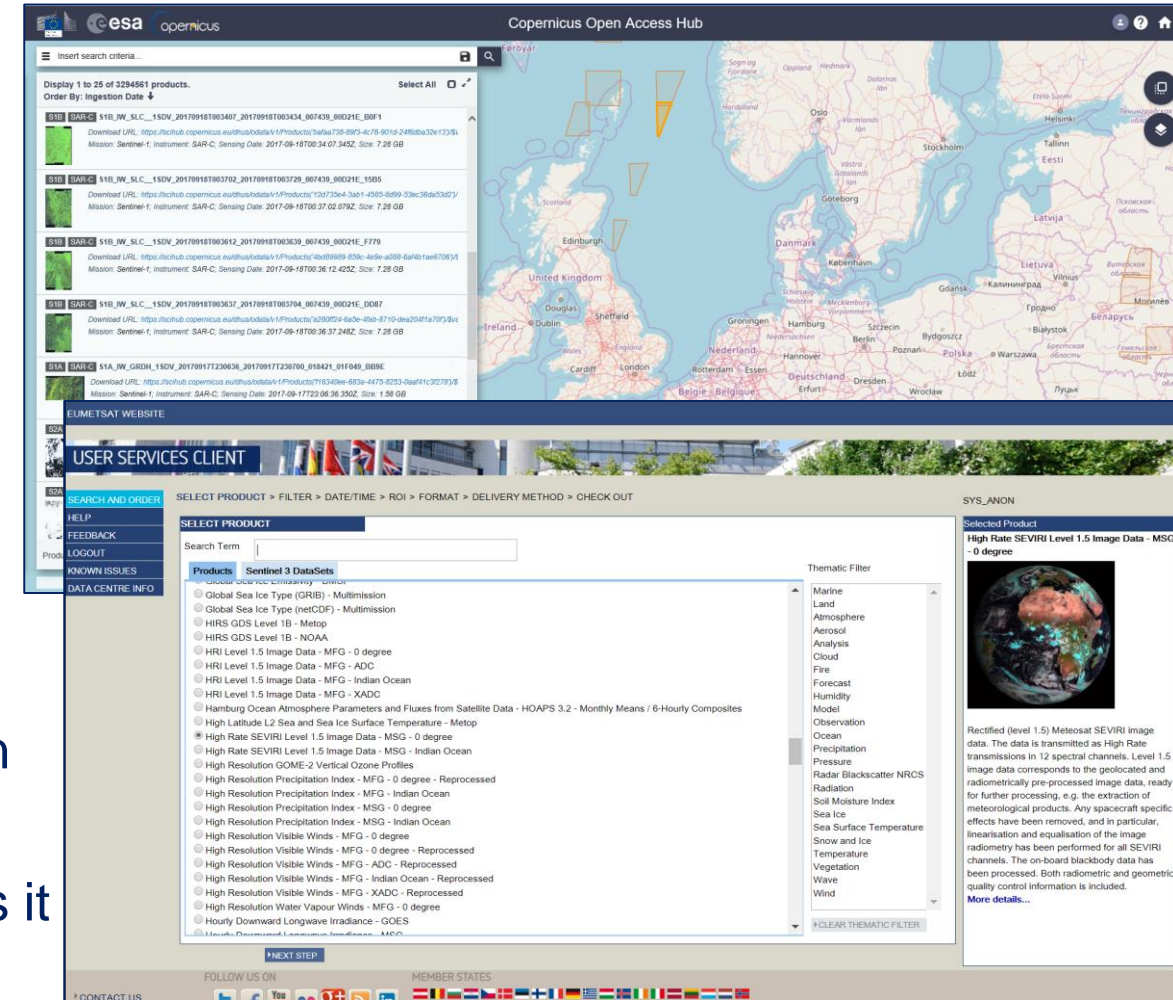
Berlin, September 27-28, 2017

Dr. Uwe Voges

u.voges@conterra.de

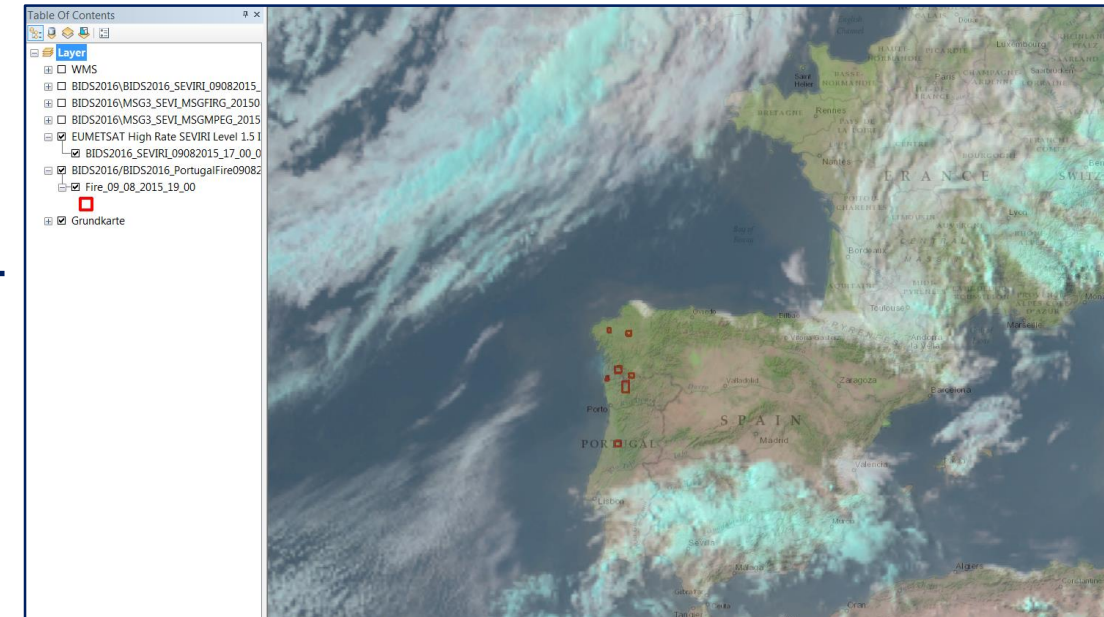
Introduction

- **Earth Observation (EO) Datasets (incl derived Products)** play an important role for decision making processes
 - > New satellite programs recently setup: **EU/ESA (Copernicus), EUMETSAT (Meteosat, MetOp)...**
 - > worldwide observations with high temporal repetition rates -> real **Big Data** !
- EO Datasets **published as Open Data** - but often through **special Portals** - stored in **Archives**
- For users not knowing about the different portals it is **hard to find, discover** and **get access** to it



Introduction

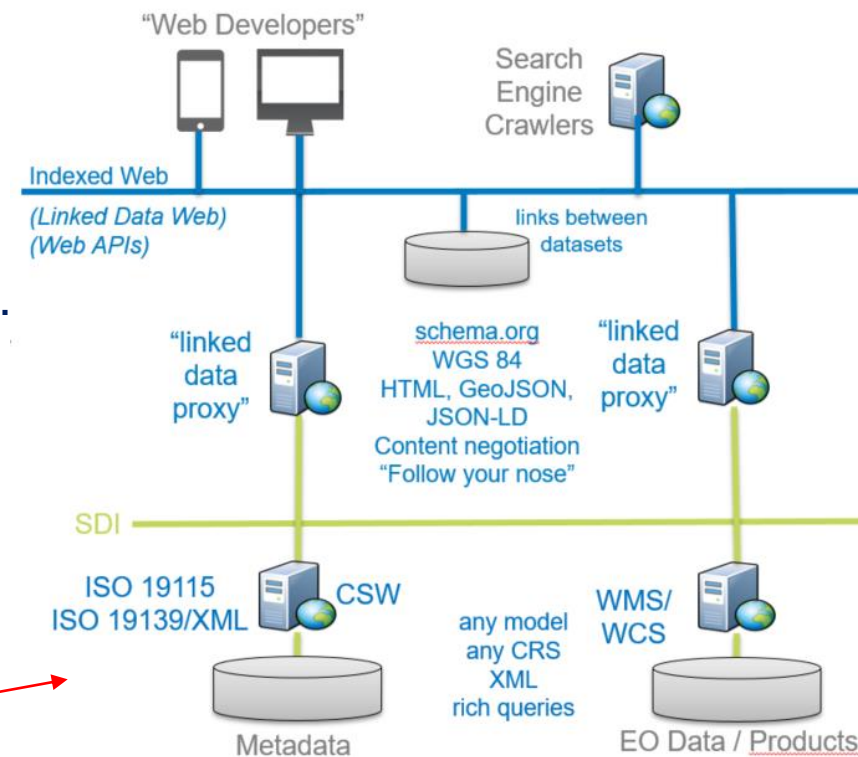
- Definition: **EO Collections** represent **set of EO Datasets** sharing a common specification
- For usage of product it can either be **downloaded...**
 - > Requires: **register, order, download, process** (format conversions, re-projections,...) the data
- OR - for an **easier access** - data is provided by **Access and View Services**, e.g. OGC WMS →
- **Collections, Products and Services** are usually **described and indexed** based on **specific metadata models** and encodings (e.g. ISO 19115 or 19119)



EUMETSAT SEVIRI data & Forest Fires WMS' used with [Esri ArcMap](#)

Introduction – Linked Open Data

- Metadata indexed and searched via **specific Catalogue Services** (e.g. **OGC CSW**)
 - > to Google & Co: **metadata hidden**
 - > Require **specific clients**
 - > Access to metadata just **following web links** is mostly **not possible**.
- For improving the situation: **technologies** related to **Linked Open Data** are promising
 - > starting at a piece of Linked Data, and **follow links** to other pieces
 - > **link EO data and mainstream web data**
 - > Can be used **on top of existing SDI**



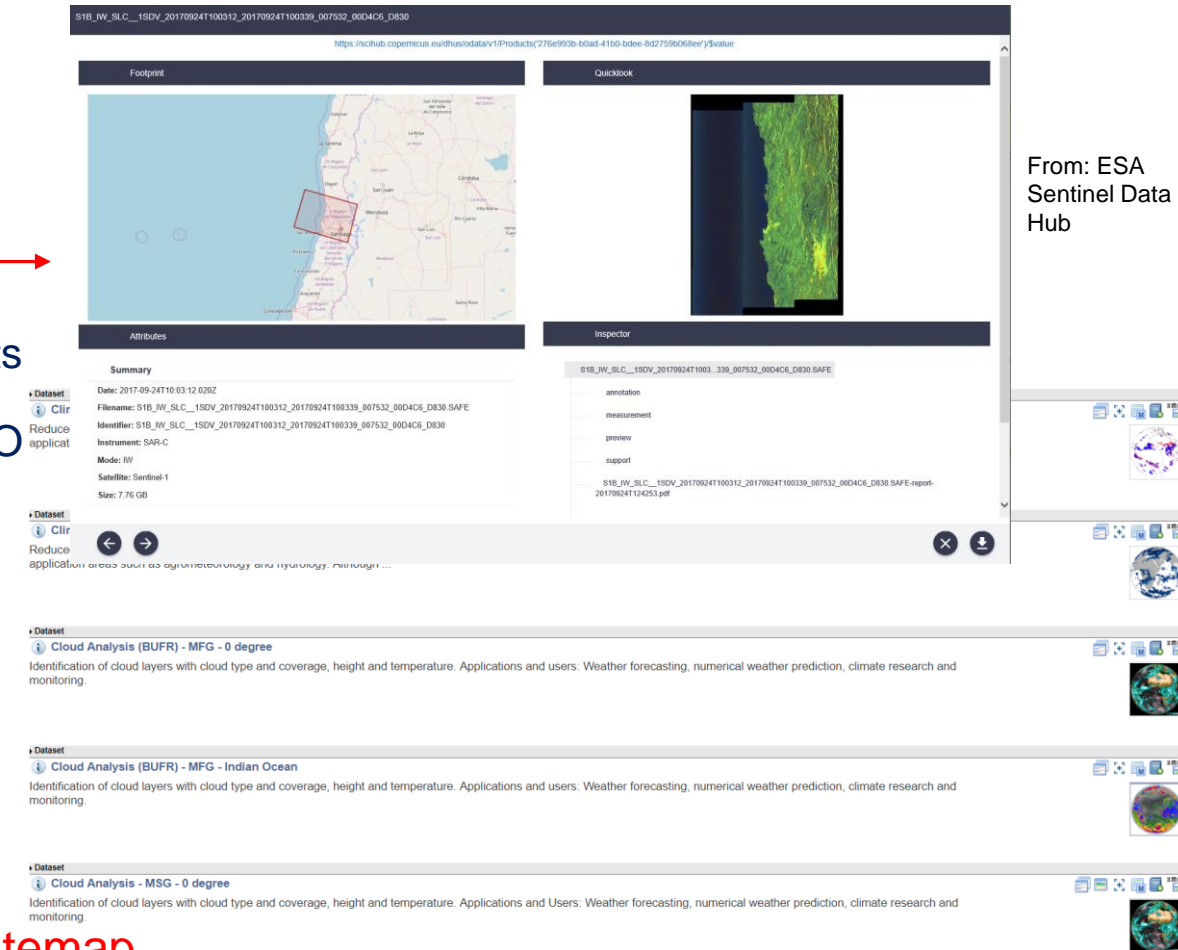
Modified from a figure provided by Geonovum

Linked Open Data for Earth Observation

- Basic technologies and best practices for LOD in EO:
 - > unique persistent HTTP URIs for EO Datasets
 - > provision of EO data (upon de-referencing) in different formats
 - > ATOM / (Geo)JSON metadata
 - > using well known (e.g. GeoNames) or well-designed vocabularies
 - > HTML page
 - > binary data
 - > maps in tiled form suitable for display in zoomable maps (WMTS)
 - > Link EO data to other resources to integrate with the web of data.

Searching with mainstream Search Engines

- Alternative to getting to data via links: **searching with Google & Co**
 - > Search engines **index web pages (HTML)** →
 - > should **provide metadata** and **links** to service endpoints
 - > **schema.org annotated**: to improve search results /SEO
 - > **landing page**: to enable crawlers to find web pages
 - > Set of EO collections as landing page: →
 - > point to EO datasets
 - > **paging** should be provided
 - > Set of links to web pages can also be **added to a sitemap**



From: ESA
Sentinel Data
Hub

From: EUMETSAT Product Navigator

Searching with mainstream Search Engines

- But in case of **Mio of similar datasets** (new permanently coming in) ... **approach is no more an option:**
 - > **not possible** to force Google to index **data as needed**
 - > **indexing takes too long:** months / years or not possible...
 - > **Domain specific vocabulars** to index/search **not supported**
- **Approach:**
 - > let just **collections** be indexed by Google
 - > use **specialized search engine** for indexing / filtering **Datasets**

The screenshot shows a Google search interface with the query 'fire monitoring collection eumetsat'. The search results page displays several links to EUMETSAT resources, including a PDF titled 'Towards an improved Active Fire Monitoring Product for ... - Eumetsat' and a page about 'The SEVIRI Active Fire Monitoring Product — EUMETSAT'. A sidebar on the left contains a navigation menu with categories like DATA, DATA DELIVERY, PRODUCTS, SERVICE STATUS, METEOSAT SERVICES, METOP SERVICES, REGIONAL DATA SERVICE / EARS, GLOBAL DATA SERVICE, CLIMATE SERVICE, COPERNICUS SERVICES, THIRD PARTY, TRAINING, TRAINING LIBRARY, TRAINING TEAM, SCIENCE ACTIVITIES, and TECHNICAL DOCUMENTS. Below the menu is a 'FILTER BY' section with dropdown menus for Category, Topic, and Keywords. The main content area on the right shows the details of the selected document, including its title, URL, and a brief description of the active fire monitoring product.

Google

fire monitoring collection eumetsat

Alle News Bilder Shopping Videos Mehr Einstellungen Tools

Ungefähr 167.000 Ergebnisse (0,57 Sekunden)

PDF Towards an improved Active Fire Monitoring Product for ... - Eumetsat
<https://www.eumetsat.int/cs/idcplg?...GET...> ▼ Diese Seite übersetzen
von S Joro - Zitiert von: 1 - Ähnliche Artikel
The Active Fire Monitoring (FIR) product displays information on the presence of fire within a collected on 18 August 2008 11:30 UTC over Western Africa.

The SEVIRI Active Fire Monitoring Product — EUMETSAT
https://www.eumetsat.int/website/home/.../DAT_2043077.html?... ▼ Diese Seite übersetzen
04.09.2009 - The active fire monitoring product is a fire detection product indicating ... The lecture describes the fire detection algorithm and presents results ...

MTG Services — EUMETSAT
<https://www.eumetsat.int/website/.../FutureSatellites/.../MTGServi...> ▼ Diese Seite übersetzen

The active fire monitoring product is a fire detection product indicating the presence of fire within a pixel.

Published: 4 September 2009

The underlying concept of the algorithm takes advantage of the fact that SEVIRI channel IR3.9 is very sensitive to hot spots which are caused by fires. The algorithm distinguishes between potential fire and active fire.

The lecture describes the fire detection algorithm and presents results over Africa. It also discusses some critical issues like 'hot' clouds (when working without cloud mask), digital filter and sunglint effects.

In the end, the lecture proposes possible future improvements of the product, and presents the imager mission (FCI) of the third generation of Meteosat (MTG).

Pre-requisites: basics in Remote Sensing SEVIRI channels, in particular the IR3.9 channel.

CATEGORY	LANGUAGE	DIFFICULTY	AUDIO	DURATION	AUTHOR
Atmosphere, Land	English	Intermediate		60 min	Hans-Joachim Lutz

LINKS

- SEVIRI Active Fire Monitoring Product presentation (PDF, 3 MB)
- The GOES 3.9 um Channel Tutorial (CIRA)
- Real-time images and ► products (EUMETSAT)

OpenSearch

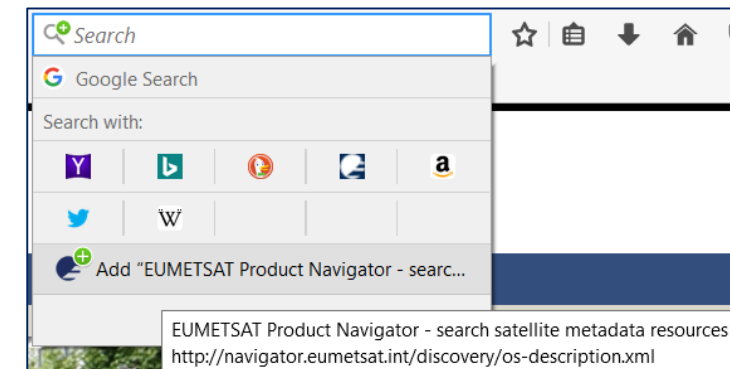
- Possible Solution: OpenSearch (OS)

- > REST-Style, web compatible API
- > providing key-value parameters to constrain the search
- > OS engines provide description documents (OSDD) used by clients
 - > include URL template per response format with parameters to be replaced by client
 - > usable by browser plugins via Autodiscovery

```
<OpenSearchDescription xmlns="http://a9.com/-/spec/opensearch/1.1/"...>
  <ShortName>EO Portal EOP OpenSearch (EOPOS) Adaptor</ShortName>
  <Url type="application/atom+xml" template="http://l...?pi={eop:parentIdentifier}&bbox={geo:box?}&dtstart={time:start?}&...">
    <param:Parameter name="dtstart" value="{time:start}" minimum="0" minInclusive="1970-01-01T00:00:00Z".../>
    <param:Parameter name="bbox" value="{geo:box}" title="Spatial extent overlapping this bounding box" minimum="0"/>
    <param:Parameter name="pi" value="{eop:parentIdentifier}" minimum="1" title="Collection Identifier">
    <param:Option value="urn:ogc:def:EOP:EUM:acronym:AVHAMV02:fileid:EO:EUM:DAT:METOP:AVHRAMV"
  ...
</OpenSearchDescription>
```

```
<link rel="search" type="application/opensearchdescription+xml"
href="http://...osdd.xml" title="EUMETSAT Product Navigator..." />
```

- > ...or by Search clients which can easily be created from OSDD



OpenSearch – Responses

- > **Results:** can be returned as HTML, Atom, RDF, KML, JSON....

- > **ResultSet**: container with search entries
- > **Container** includes information about current search and means for pagination
- > **Entries** represent **EO Datasets**
 - element **values mapped** from original metadata
 - **links** provide means for linking to external metadata details (@rel="via")

- > Results can be aggregated by clients and e.g. rendered in browser

[illegible]

OpenSearch – OGC Extensions for Geo/EO

- OpenSearch is extendable:
 - > by query parameters (in addition to those provided in OS 1.1 namespace, e.g. searchTerm)
 - > by additional search response formats
- OGC 10-032r8: OGC OpenSearch Geo and Time Extensions
 - > parameters and operators to define geographic and temporal constraints
 - > Example URL template: `http://server...?q={searchTerms}&g={geo:geometry?}&r={geo:relation}`
 - > Example request: `http://...?q=wind&g=POLYGON((0.582 40.496, ..., 0.582 40.496))&r=disjoint`
 - > Atom response format defining constraints for the elements



OpenSearch – OGC Extensions for Geo/EO

- OGC 13-026r9: OGC OpenSearch Extension for Earth Observation (OpenSearch-EO)

- > Search for EO Collections and EO Datasets (build upon 10-032 - last slide)

- > EO specific parameters, e.g. platform, orbitType, sensorType, cloudCover,..

- > semantic defined by mappings to EO MD models (e.g. OGC 10-157)

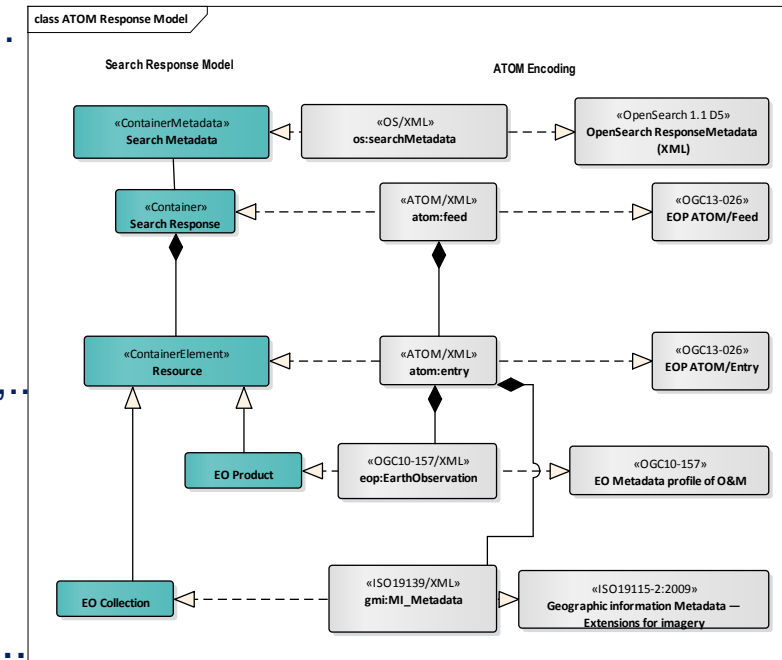
- > Defines Atom response format with constraints - plus:

- > links with defined semantic (“rel” attribute):

- > e.g. 2-step-search, source of entry, alternate MD formats, docu, images,..

- > service or online content offering for entries targeting OGC clients

- > Used by Aggregators: CEOS/CWIC, ESA/FEDEO, EUM/EOP, NASA...



GeoJSON / JSON-LD

- Problem: Atom not ideal for web clients and developers, poor mass market adoption

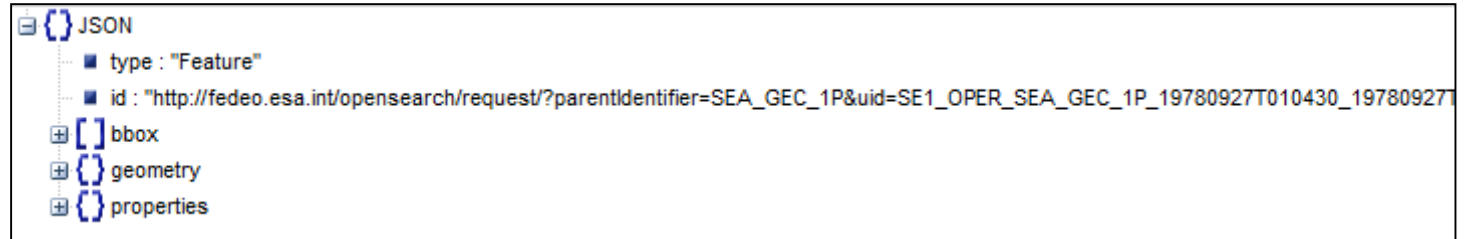
- Better: GeoJSON 

- > JSON: text-based, lightweight, key/value pairs, derived from JavaScript

- > GeoJSON: format to encode simple geogr. features using JSON

- > supported by many js-libraries

- > But: No commonly agreed vocabulary and not LD



- Solution: JSON-LD (JSON-based format to serialize Linked Data)

- > way to combine web-based environments and linked data



- > implements: 5 star SDW-BP

GeoJSON / JSON-LD

- > In addition to JSON, JSON-LD introduces:
 - > Semantically enriched properties => identified by an IRI
 - > defines the vocabulary
 - > @id keyword: refer to JSON object on different site by IRI
- > JSON-LD introduces a @context
 - > addresses issue of verbose documents by mapping terms to IRIs
 - > a way to associate datatypes with values (e.g. dateTime)
- > serializable as RDF: can be used with SPARQL
- > next generation REST for hypermedia driven systems (HATEOAS)

```
{
  "@context": {
    "gj": "https://purl.org/geojson/vocab#",
    "dct": "http://purl.org/dc/terms/",
    "gda": "http://schemas.google.com/g/2005",
    ...
    "geometry": "gj:geometry",
    "coordinates": "gj:coordinates",
    "Polygon": "gj:Polygon",
    "properties": "gj:properties",
    "updated": {
      "@id": "dct:modified",
      "@type": "http://www.w3.org/2001/XMLSchema#dateTime"
    },
    "bbox": {
      "@container": "@list",
      "@id": "gj:bbox"
    }
  },
  "features": {
    "@container": "@set",
    "@id": "gj:features"
  },
  "type": "@type",
  "id": "@id",
  "Feature": "gj:Feature",
  "FeatureCollection": "gj:FeatureCollection"
},
"type": "FeatureCollection",
"id": "http://rs211980.rs.hosteurope.de/.../eopos?dtstart=...",
"bbox": [ -67.5, -67.5, 67.5, 67.5 ],
...
"features": [ {
  "type": "Feature",
  "id": "urn:ogc:def:EOP:EUM:acronym:MSGCSKR:...",
  "geometry": {
    "type": "Polygon",
    "coordinates": [ [ [ -2.682513, 63.261372 ], ... ] ]
  },
  "properties": {
    "updated": "2016-08-04T00:45:00Z",
    ...
  }
} ]
}
```

OpenSearch – OGC GeoJSON / JSON-LD response formats for EO

- **OGC 17-074**: OGC OpenSearch-EO GeoJSON (-LD) Response Encoding Standard
- **OGC 17-003**: OGC EO Dataset Metadata GeoJSON(-LD) Encoding Standard
 - > both define **GeoJSON encoding** as **compaction** through a **context** of the proposed JSON-LD encoding
 - > **define properties** with domain, range, namespace,...
 - > **widely adopted vocabularies** used DC, OGC, Atom,...
 - > **17-003: EO Vocabulary** defined with **RDF Schema**:

```
eop:statusSubType a owl:DatatypeProperty ;
  rdfs:label "statusSubType" ;
  dcterms:source "OGC 17-003"^^xsd:string ;
  rdfs:comment ""Refines the status of a product when the "status" is set to "ARCHIVED".
Possible values:
- "ON-LINE"
- "OFF-LINE".
"" ;
  rdfs:domain eop:ProductInformation ;
  rdfs:isDefinedBy eop: ;
  rdfs:range xsd:string .
```

```
{
  "@context": "http://www.opengis.net/spec/opensearcheo/geojson/1.0",
  "@type": "FeatureCollection",
  "@id": "http://rs211980.rs.hosteurope.de/eo-csw-umarf/...",
  "bbox": [ -67.5, -67.5, 67.5, 67.5 ],
  "properties": {
    "type": "Properties",
    "title": "EUMETSAT EO Portal - EOPOS search response (geojson)",
    "subject": "http://www.eumetsat.int/concept/theme/climatologyMeteorologyAtmosphere",
    "totalResults": 3755,
    ...
  },
  "features": [
    {
      "type": "Feature",
      "id": "MSGCSKR:fileid:EO:EUM:DAT:MSG:CSR:20160804004500MSG3MSGCSKR02EF00PE",
      "geometry": {
        "type": "Polygon",
        "coordinates": [[[ -2.682513, 63.261372], [-2.69574, 61.997604] ...]]
      },
      "properties": {
        "type": "Properties",
        "kind": "EarthObservation",
        "updated": "2016-08-04T00:45:00Z",
        "identifier": "MSGCSKR:fileid:EO:EUM:DAT:MSG:CSR:20160804004500MSG3MSGCSKR02EF00PE",
        ...
      },
      "links": {
        "@type": "Links",
        "@context": { "type": "atom:type" },
        "alternates": [ {
          "@type": "Link",
          "href": "http://rs211980.rs.hosteurope.de/eo-csw-umarf/servlet/csw?service=CSW-ebRIM&...",
          "type": "application/gml+xml;profile=http://www.opengis.net/spec/EOMPOM/1.1",
          "title": "Metadata (OGC O&M EOP Format)"
        } ],
        "data": [ {
          "@type": "Link",
          "href": "http://rs211980.rs.hosteurope.de/eo-download/servlet/download?service=DSEO&ve...",
          "type": "application/binary",
          "title": "Download"
        } ],
        "previews": [ {
          "@type": "Link",
          "@context": { "type": "atom:type" },
          "id": "http://navigator.eumetsat.int/smartEditor/preview/msgcsr.jpg",
          "type": "image/jpeg",
          "expression": "sample",
          "category": "http://www.opengis.net/spec/EOMPOM/1.0#THUMBAIL"
        } ]
      }
    }
  ]
}
```

OpenSearch – Outlook

- Problems with OpenSearch OSDD...
 - > XML based
 - > URL template not structured, string based
 - > Limited options especially for response definitions
- Need for a better concept for a service description:
 - > Better aligned with web clients and developers
 - > structured and more flexible
 - > existing tool support
- Possible Solution: OpenAPI 3.0
 - > Some points open: e.g. how to link parameters to predefined EO parameter definitions...



con terra – Gesellschaft für Angewandte Informationstechnologie mbH

Martin-Luther-King-Weg 24

48155 Münster

Telefon +49 89 207 005 2200

info@conterra.de

Thank you !