

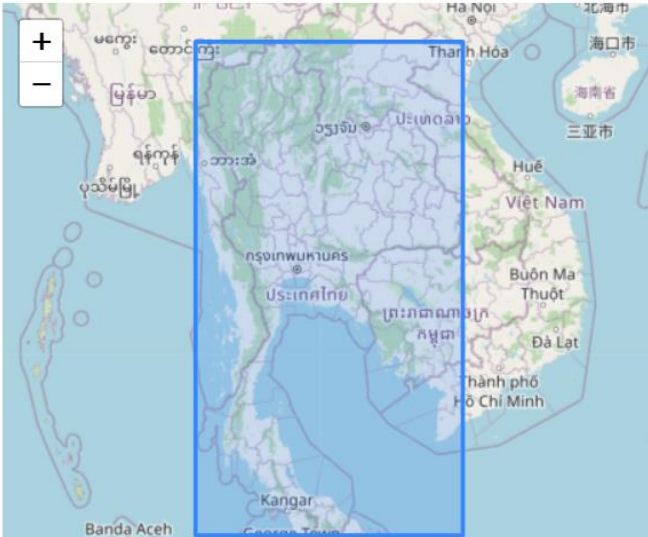
Datasets in WIS 2.0 and the Global Discovery Catalogue

Datasets in WIS2

A Dataset groups data items into a single, conceptual resource (collection):

- Helps to categorize and make easier to analyze
- Statements about the Dataset (metadata) apply to the entire collection
- The dataset has an identifier (metadata-id) associated to all data items

Hourly synoptic observations from fixed-land stations (SYNOP) (th-tmd)



Property	Value
id	urn:wmo:md:th-tmd:synop-hourly
type	dataset
title	Hourly synoptic observations from fixed-land stations (SYNOP) (th-tmd)
description	Thailand synop-hourly
keywords	observations temperature visibility precipitation pressure clouds evaporation radiation wind total sunshine humidity
themes	concepts: id: weather

Examples of datasets

- Synoptic observations from fixed-land stations
- Predictions from NWP models
- Archive of daily climate observations
- Real-time water levels and flow (discharge)

Datasets in WIS2

Datasets provide consistency to facilitate data processing:

- All the data should be of the **same type** (e.g., observations from weather stations).
- All the data should have the **same license and/or usage conditions**.
- All the data should be subject to the **same quality management** regime - which may mean that all the data is **collected or created using the same processes**.
- All the data should be **encoded in the same way** (i.e., using the same data formats and vocabularies).
- All the data should be **accessible using the same protocols**.



Datasets need metadata

- Description/documentation
- Identification
- Key words
- Temporal/Spatial Extents
- Data Provider Contact(s)
- Access control mechanism
- Rights and License



Discovery Metadata in WIS2 is defined by the WMO Core Metadata Profile 2 (WCMP2)

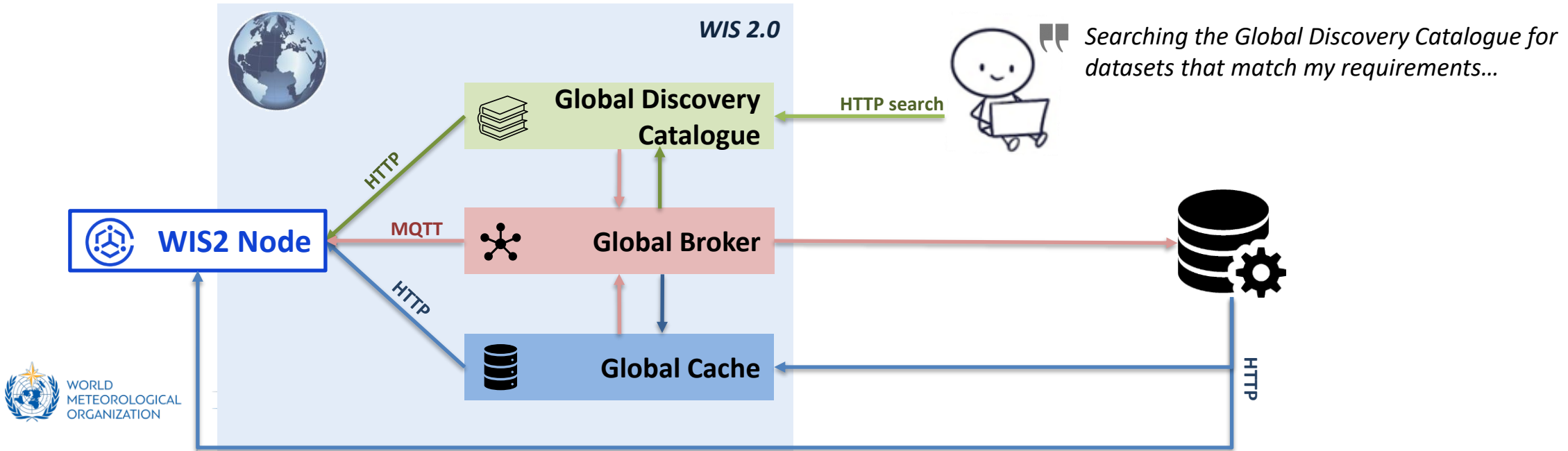
Discovering datasets in WIS2

Global Discovery Catalogues provide an OGC API – Records endpoint to search for data on WIS2

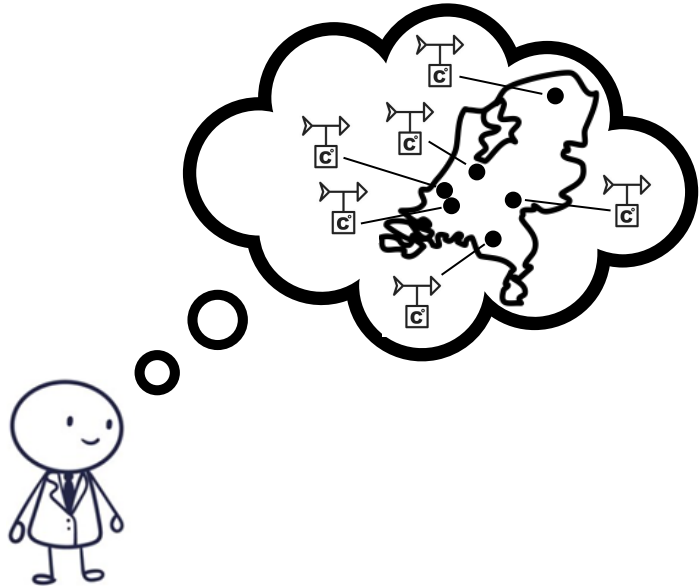
GDC ECCC, Canada: wis2-gdc.weather.gc.ca/collections/wis2-discovery-metadata/items

GDC DWD, Germany: wis2.dwd.de/gdc/collections/wis2-discovery-metadata/items

GDC CMA, China: gdc.wis.cma.cn/collections/wis2-discovery-metadata/items



Sharing data in WIS 2.0



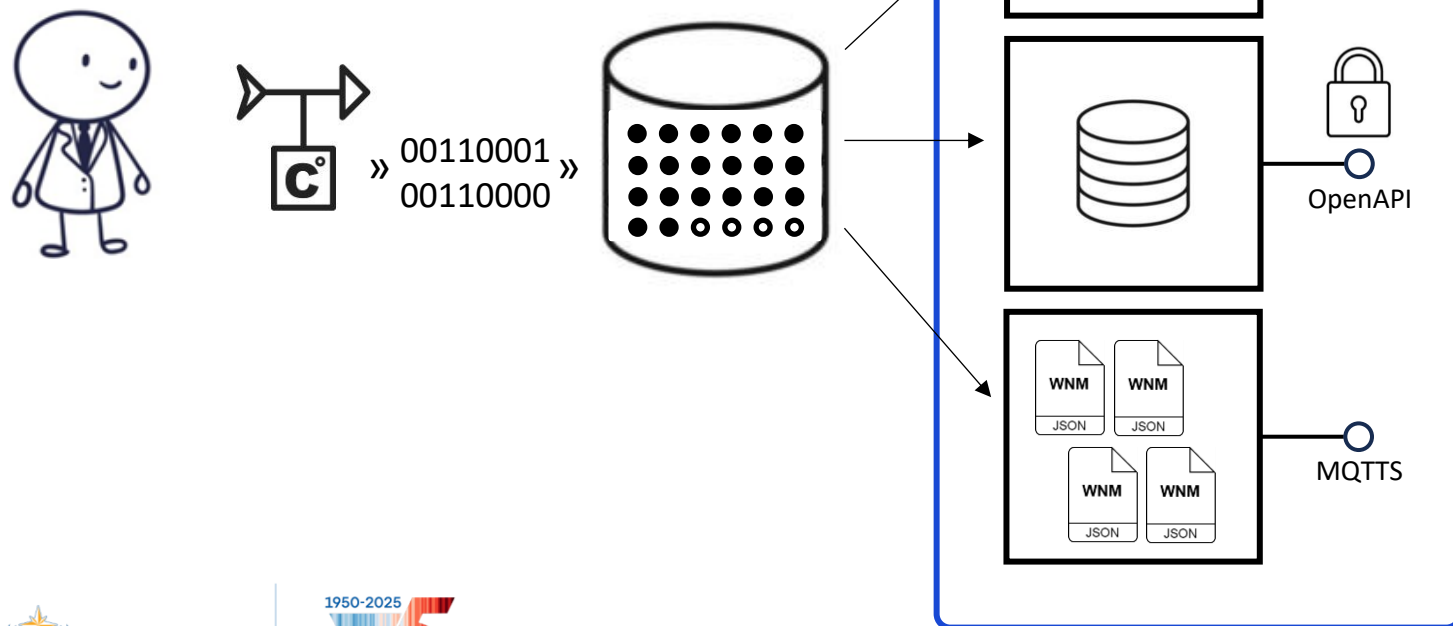
“ I want to share data from my national network of land-based, fixed observing stations



“ I need real-time weather data for my application

Sharing data in WIS 2.0

For WIS 2.0 I see that my collection of real-time observation data is called a **'Dataset'**¹
Following instructions in the WIS2 Guide², I establish a WIS2 Node to share my **Dataset**



¹ Guide to WIS (WMO No. 1061), Vol II, §1.1.4 Why are datasets so important [\[draft\]](#)

² Guide to WIS (WMO No. 1061), Vol II, §1.3.1 How to get started [\[draft\]](#)

³ Manual on WIS (WMO No. 1060), Vol II, Appendix E: Notification Message [\[draft\]](#)

⁴ Manual on WIS (WMO No. 1060), Vol II, Appendix D: WIS2 Topic Hierarchy [\[draft\]](#)

files in BUFR format
or other WMO approved data formats

interactive API (optional)
self-describing RESTful APIs, ideally with OpenAPI metadata

WIS2 Notification Messages³
GeoJSON published on standardised topic hierarchy⁴

But I still don't know anything
about your Dataset!

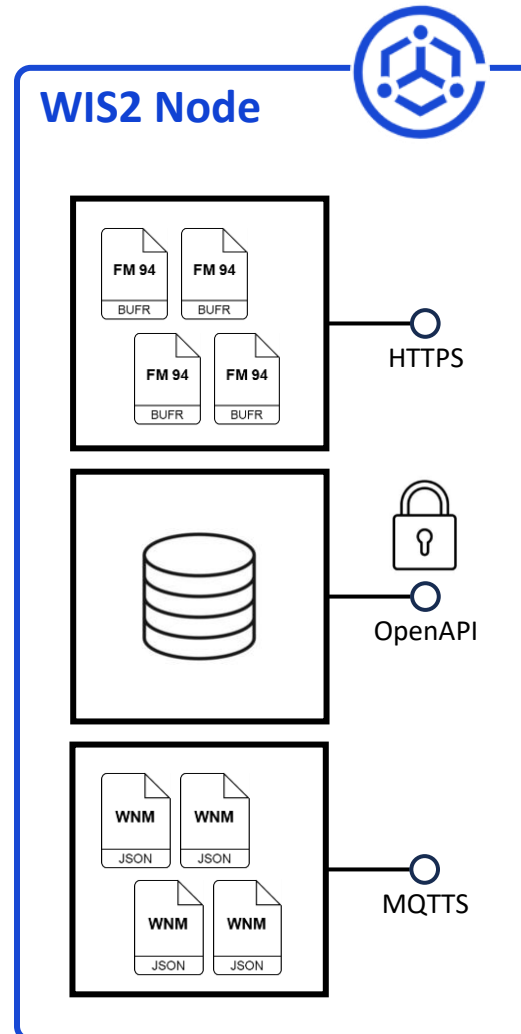


Sharing metadata in WIS 2.0

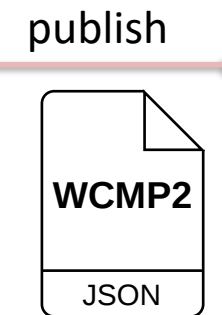
Ahh, good point! The next section in the WIS2 Guide⁵ says that I need to provide 'discovery metadata'⁶ about my Dataset



Description of the Dataset
Identifier
Title
Description
Keywords
Geometry (extent)
Time (extent)
Who to contact
Publisher contact
How to access
Data access (files)
Data access (API)
Data access (notifications)
Conditions of use
Data policy
Rights
License
How to attribute
Citation



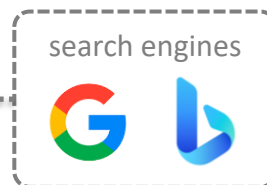
That's great. I can find the description of your Dataset. I can understand **what it contains** and **how to access** it. The **conditions of use** are clear, and I see how I can **attribute** you as the source of the data !



publish



index



⁵ Guide to WIS (WMO No. 1061), Vol II, §1.3.2 [How to provide discovery metadata to WIS2](#)

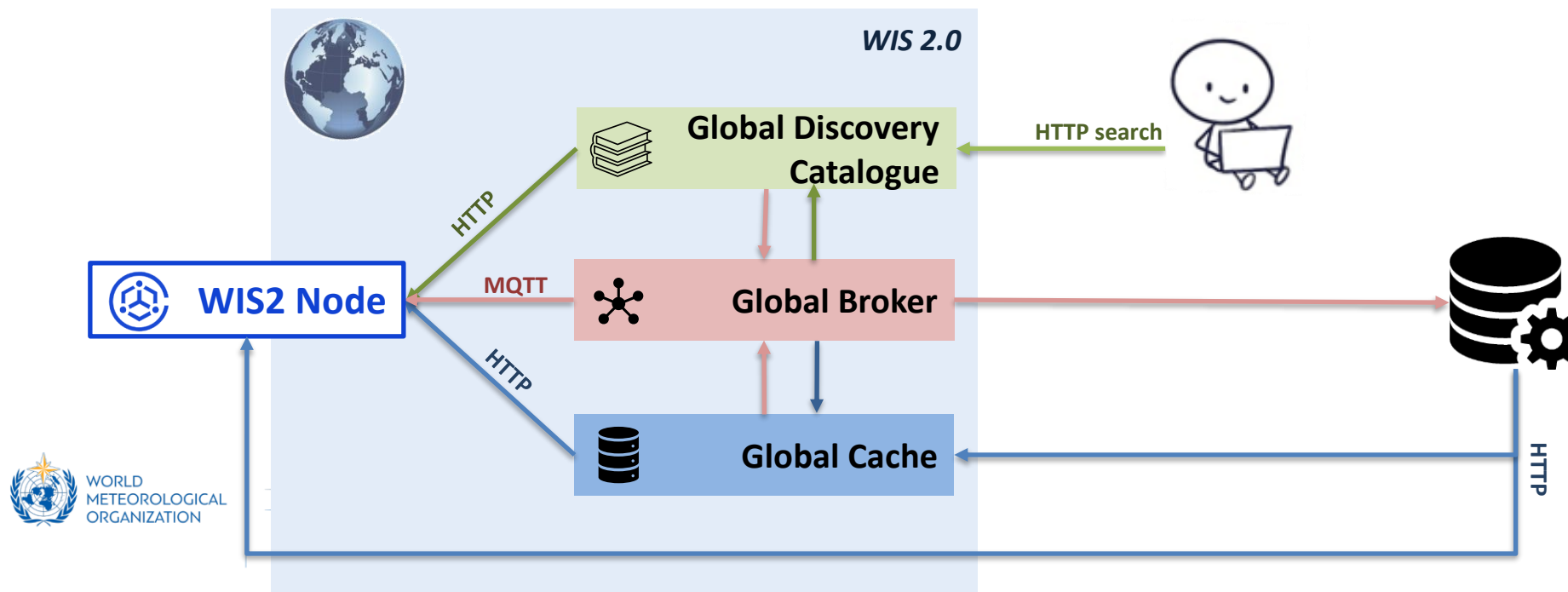
⁶ Manual on WIS (WMO No. 1060), Vol II, Appendix F: [WMO Core Metadata Profile 2](#)

The Dataset concept

¹ Guide to WIS (WMO No. 1061), Vol II, §1.1.4 [Why are datasets so important](#)

The Dataset concept is central for WIS 2.0:

- **Dataset** groups data items into a single, conceptual resource
- Dataset properties provided by **Discovery Metadata** using **WMO Core Metadata Profile 2 (WCMP2)**
- Search for Datasets that using the **Global Discovery Catalogue (GDC)**
- **Subscribe** to notifications about **updates** for a **Dataset** via a **Global Broker (GB)**
- **Access the data** that comprises a Dataset using **HTTP/HTTPS**



GDC Overview

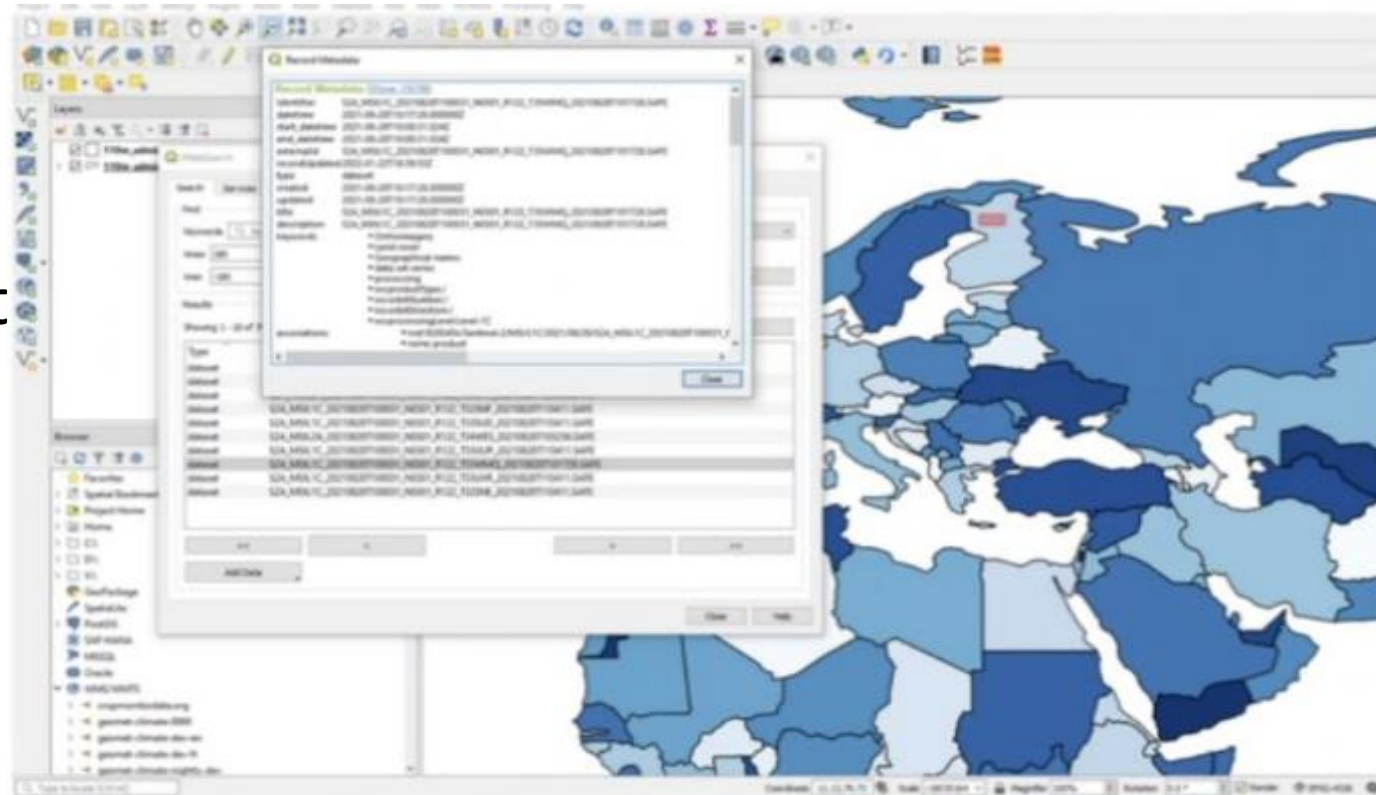
- Provides cataloguing and discovery capability of WMO dataset collections
- Web-based API facilitating search/browse data published via WIS
- Harvests WIS 2 discovery metadata from WIS2 Nodes
- Yellow pages (discovery metadata) gateway into WIS data and services
- Provides indexing capability to mass market search engines
- Provides WCMP2 validation and quality assessment “as a service” in support of WIS2 compliance and continuous improvement in alignment with WIS 2 metadata Key Performance Indicators (KPIs)

Searchable Catalogue

Predicate	Example(s)
Spatial	<code>bbox=minx,miny,maxx,maxy</code>
Temporal	• <code>datetime=t1</code> • <code>datetime=t1/t2</code> • <code>datetime=t1/..</code>
Freetext	<code>q=air+temperature</code>
Attributes	• <code>title=air+temperature</code> • <code>nested.attribute=ozone</code> • <code>wmo:dataPolicy=core</code>
Sorting/paging	• <code>sortBy=title&limit=100&startindex=11</code> • <code>"prev" / "next" link relations</code>
Filter returnable properties	• <code>properties=title,description</code>

GDC Catalogue Search

- Better metadata results in better findability/use
- The WIS2 Global Discovery Catalogue will work with multiple applications and environments (mobile, web application, workstation, GIS, etc.)
- Example: using QGIS, users will be able to search the GDC, view results and choose to subscribe and/or download data of interest



WIS2 GDC: OpenAPI/Swagger

Execute

Clear

Responses

Curl

```
curl -X GET "https://api.weather.gc.ca/collections/wis2-discovery-metadata/items?f=json&lang=en-CA&limit=10" -H "accept: application/geo+json"
```

Request URL

```
https://api.weather.gc.ca/collections/wis2-discovery-metadata/items?f=json&lang=en-CA&limit=10
```

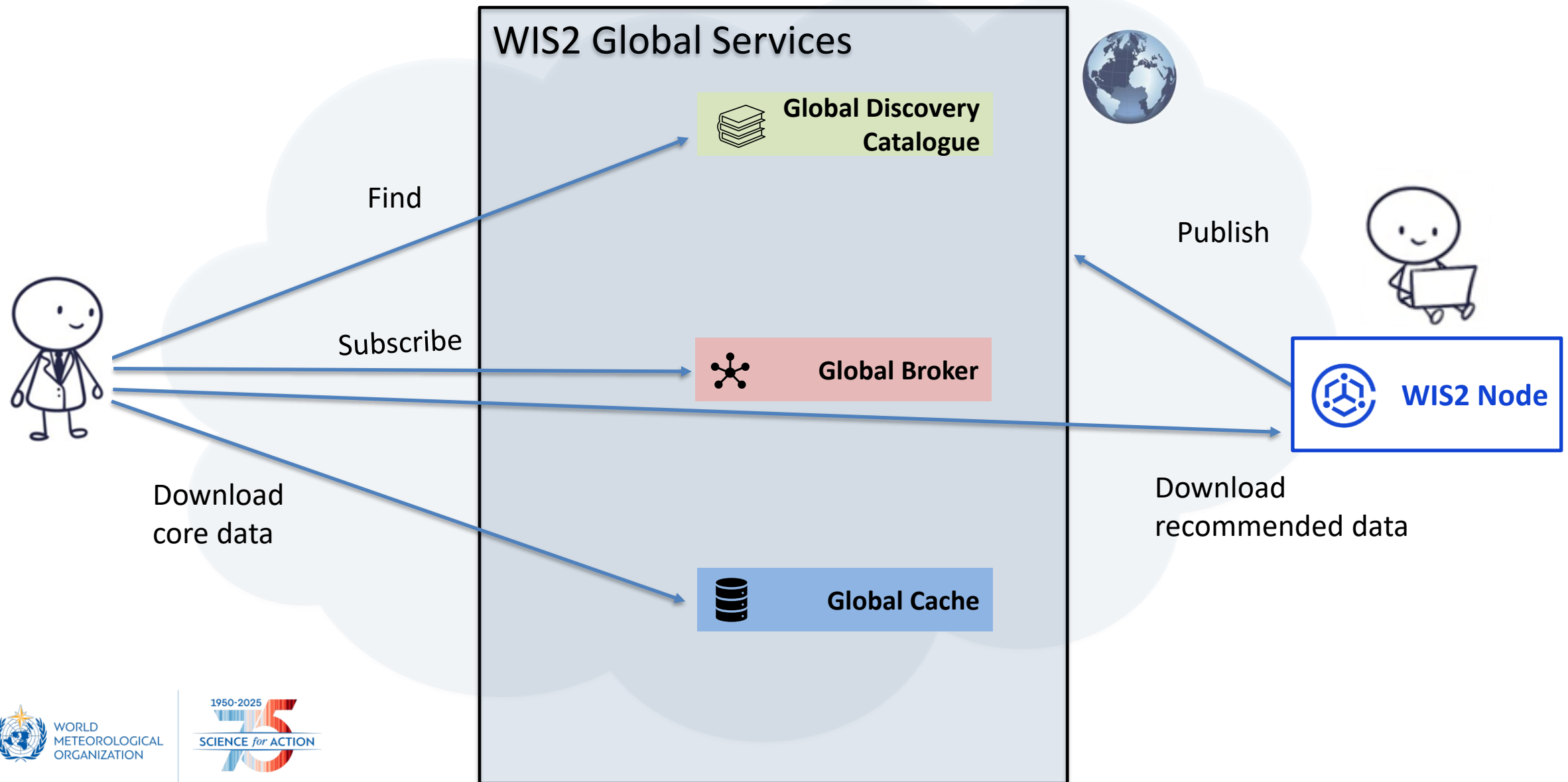
Server response

Code	Details
200	Response body <pre>{ "type": "FeatureCollection", "features": [{ "id": "urn:x-wws:sd:ca.gc.ec.msc.weather.observations.swob-realtime", "conformsTo": ["http://www.opengis.net/spec/ogcapi-records-1/1.0/req/record-core"], "type": "Feature", "geometry": { "type": "Polygon", "coordinates": [[[-142, 28], [-142, 82], [-52, 82], [-52, 28]]] } }] }</pre> Download

Response headers

```
access-control-allow-headers: OriginX-Requested-WithContent-TypeAcceptAuthorization
access-control-allow-origin: *
connection: Keep-Alive
content-encoding: gzip
content-language: en-CA
content-type: application/json
date: Mon05 Dec 2022 00:49:34 GMT
keep-alive: timeout=5max=100
server: Apache
strict-transport-security: max-age=63072000; preload
transfer-encoding: chunked
vary: Accept-Encoding
x-powered-by: pygeosapi 0.12.0
```

Publish, Find, Subscribe, Download



Thank you

wmo.int