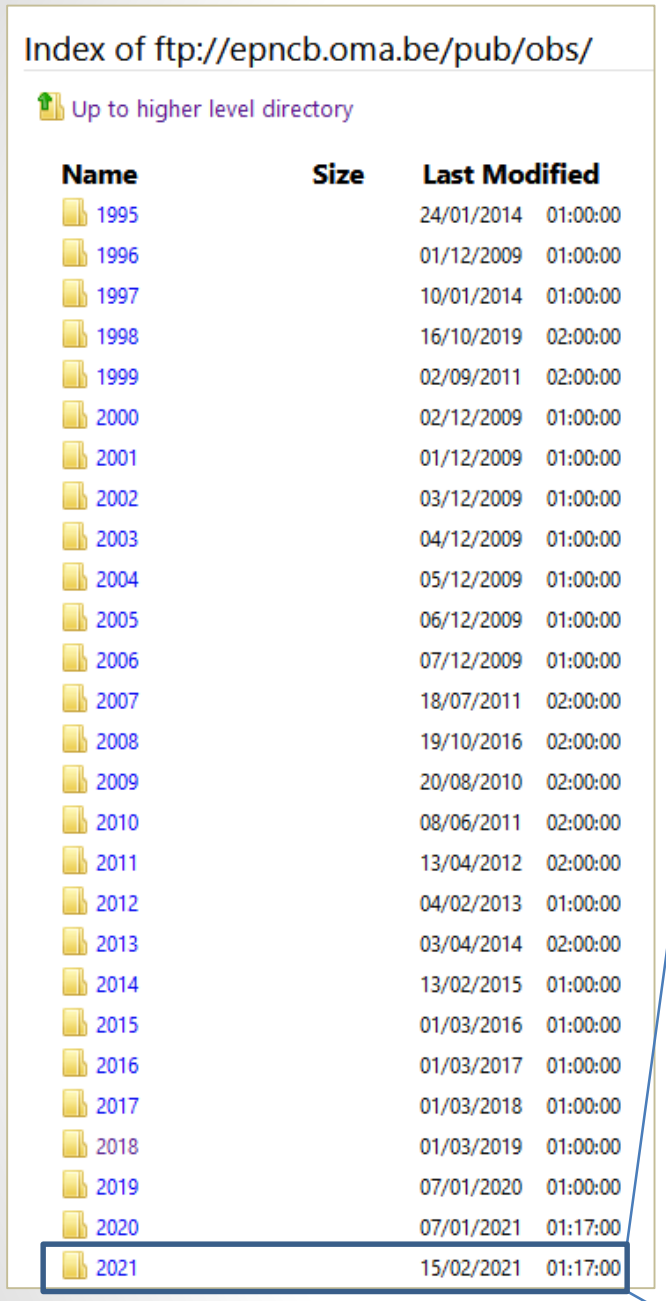
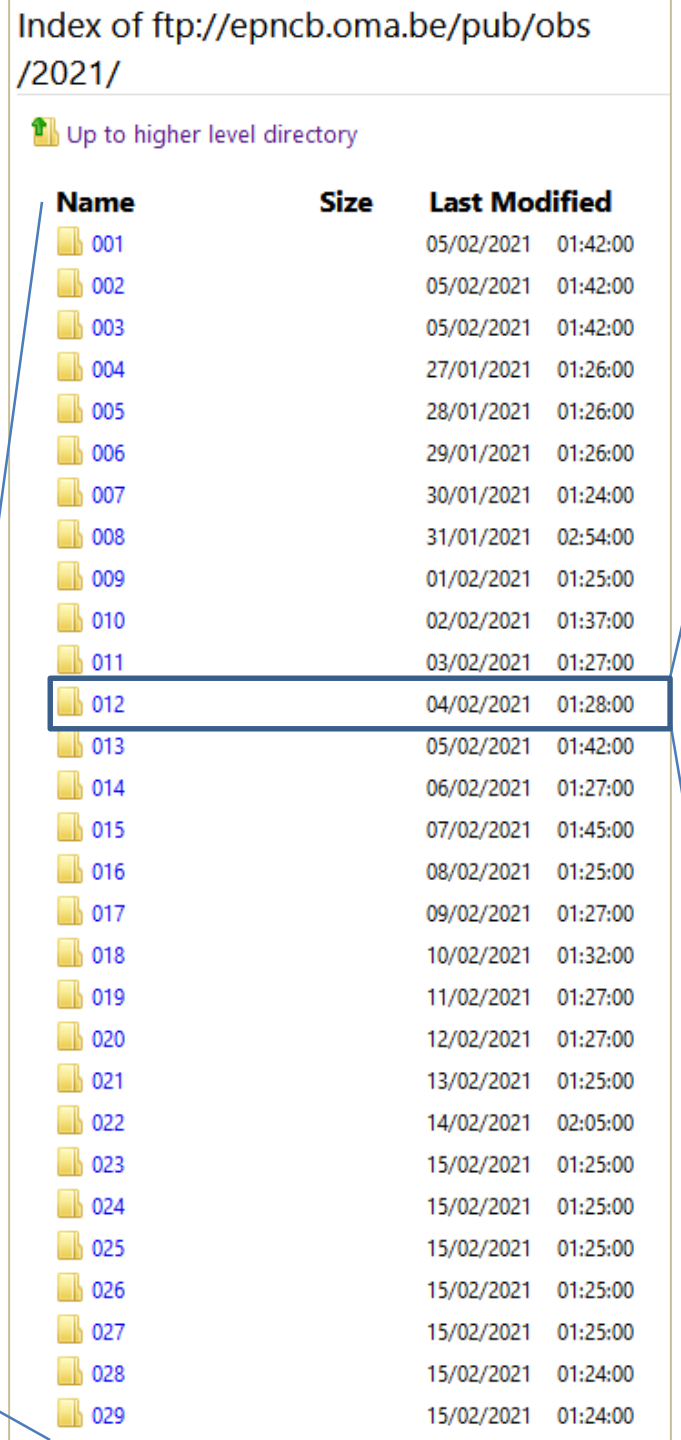
Towards FAIR GNSS data

EUREF Symposium – May 27, 2021

C. Bruyninx, A. Miglio, A. Fabian, J. Legrand


Index of ftp://epncb.oma.be/pub/obs/
Up to higher level directory
Name Size Last Modified
1995 24/01/2014 01:00:00
1996 01/12/2009 01:00:00
1997 10/01/2014 01:00:00
1998 16/10/2019 02:00:00
1999 02/09/2011 02:00:00
2000 02/12/2009 01:00:00
2001 01/12/2009 01:00:00
2002 03/12/2009 01:00:00
2003 04/12/2009 01:00:00
2004 05/12/2009 01:00:00
2005 06/12/2009 01:00:00
2006 07/12/2009 01:00:00
2007 18/07/2011 02:00:00
2008 19/10/2016 02:00:00
2009 20/08/2010 02:00:00
2010 08/06/2011 02:00:00
2011 13/04/2012 02:00:00
2012 04/02/2013 01:00:00
2013 03/04/2014 02:00:00
2014 13/02/2015 01:00:00
2015 01/03/2016 01:00:00
2016 01/03/2017 01:00:00
2017 01/03/2018 01:00:00
2018 01/03/2019 01:00:00
2019 07/01/2020 01:00:00
2020 07/01/2021 01:17:00
2021 15/02/2021 01:17:00


Index of ftp://epncb.oma.be/pub/obs/2021/
Up to higher level directory
Name Size Last Modified
001 05/02/2021 01:42:00
002 05/02/2021 01:42:00
003 05/02/2021 01:42:00
004 27/01/2021 01:26:00
005 28/01/2021 01:26:00
006 29/01/2021 01:26:00
007 30/01/2021 01:24:00
008 31/01/2021 02:54:00
009 01/02/2021 01:25:00
010 02/02/2021 01:37:00
011 03/02/2021 01:27:00
012 04/02/2021 01:28:00
013 05/02/2021 01:42:00
014 06/02/2021 01:27:00
015 07/02/2021 01:45:00
016 08/02/2021 01:25:00
017 09/02/2021 01:27:00
018 10/02/2021 01:32:00
019 11/02/2021 01:27:00
020 12/02/2021 01:27:00
021 13/02/2021 01:25:00
022 14/02/2021 02:05:00
023 15/02/2021 01:25:00
024 15/02/2021 01:25:00
025 15/02/2021 01:25:00
026 15/02/2021 01:25:00
027 15/02/2021 01:25:00
028 15/02/2021 01:24:00
029 15/02/2021 01:24:00


Index of ftp://epncb.oma.be/pub/obs/2021/012/
Up to higher level directory
Name Size Last Modified
File: ACOR00ESP_R_20210120000_01D_30S_MO.crx.gz 1695 KB 19/01/2021 01:27:00
File: ADAR0120.21D.gz 1143 KB 19/01/2021 01:17:00
File: AJAC00FRA_R_20210120000_01D_30S_MO.crx.gz 3747 KB 19/01/2021 01:27:00
File: AJAC0120.21D.gz 2465 KB 19/01/2021 01:17:00
File: ALAC00ESP_R_20210120000_01D_30S_MO.crx.gz 1378 KB 19/01/2021 01:27:00
File: ALBA00ESP_R_20210120000_01D_30S_MO.crx.gz 1347 KB 19/01/2021 01:27:00
File: ALME00ESP_R_20210120000_01D_30S_MO.crx.gz 1722 KB 19/01/2021 01:27:00
File: ANK20120.21D.gz 753 KB 19/01/2021 01:17:00
File: ANKR00TUR_R_20210120000_01D_30S_MO.crx.gz 1803 KB 19/01/2021 01:27:00
File: ANKR0120.21D.gz 1355 KB 19/01/2021 01:17:00
File: AQUI00ITA_R_20210120000_01D_30S_MO.crx.gz 2063 KB 19/01/2021 01:27:00
File: ARGI00FRO_R_20210120000_01D_30S_MO.crx.gz 2305 KB 19/01/2021 01:27:00
File: ARGI0120.21D.gz 1006 KB 19/01/2021 01:17:00
File: ARIS0120.21D.gz 1152 KB 19/01/2021 01:17:00
File: ARJ600SWE_R_20210120000_01D_30S_MO.crx.gz 3111 KB 19/01/2021 01:27:00
File: AUBG00DEU_R_20210120000_01D_30S_MO.crx.gz 2616 KB 19/01/2021 01:27:00
File: AUBG0120.21D.gz 1262 KB 19/01/2021 01:17:00
File: AUT10120.21D.gz 343 KB 19/01/2021 01:17:00
File: AUTN00FRA_R_20210120000_01D_30S_MO.crx.gz 1808 KB 19/01/2021 01:27:00
File: AUTN0120.21D.gz 1614 KB 19/01/2021 01:17:00
File: AXPV00FRA_R_20210120000_01D_30S_MO.crx.gz 1997 KB 19/01/2021 01:27:00
File: AXPV0120.21D.gz 2039 KB 19/01/2021 01:17:00
File: BADH00DEU_R_20210120000_01D_30S_MO.crx.gz 2596 KB 19/01/2021 01:27:00
File: BAIA00ROU_R_20210120000_01D_30S_MO.crx.gz 2591 KB 19/01/2021 01:27:00
File: BAIA0120.21D.gz 637 KB 19/01/2021 01:17:00
File: BAUT00DEU_R_20210120000_01D_30S_MO.crx.gz 5600 KB 19/01/2021 01:27:00
File: BAUT0120.21D.gz 1597 KB 19/01/2021 01:17:00
File: BBYS00SVK_R_20210120000_01D_30S_MO.crx.gz 2080 KB 19/01/2021 01:28:00
File: BCLN00ESP_R_20210120000_01D_30S_MO.crx.gz 1300 KB 19/01/2021 01:28:00
File: BELL00ESP_R_20210120000_01D_30S_MO.crx.gz 2391 KB 19/01/2021 01:28:00
File: BOGE00POL_R_20210120000_01D_30S_MO.crx.gz 2582 KB 19/01/2021 01:28:00
File: BOGE0120.21D.gz 1981 KB 19/01/2021 01:17:00

GNSS data repositories

"I am not able to download the RINEX data from your website and I am interested in the data from all the stations in a given country."

"All stations come with RINEX header coordinates. How frequently are they updated and hence how reliable are they?"

"Are data free to use?"

"I need data from EPN stations providing BeiDou-3 observations"

"I need the data of the following three points.
- NOA point in Greece
- TUC Point in Greece
- DYNG Point in Greece"

"Can I develop commercial service based on EPN data?"

..need for a service that, unlike FTP,

- allows users to **easily search for and download data** (and metadata) from multiple stations
- clearly provides **data usage conditions**
- allows to **acknowledge data provider**

EPN CB & M³G

FTP

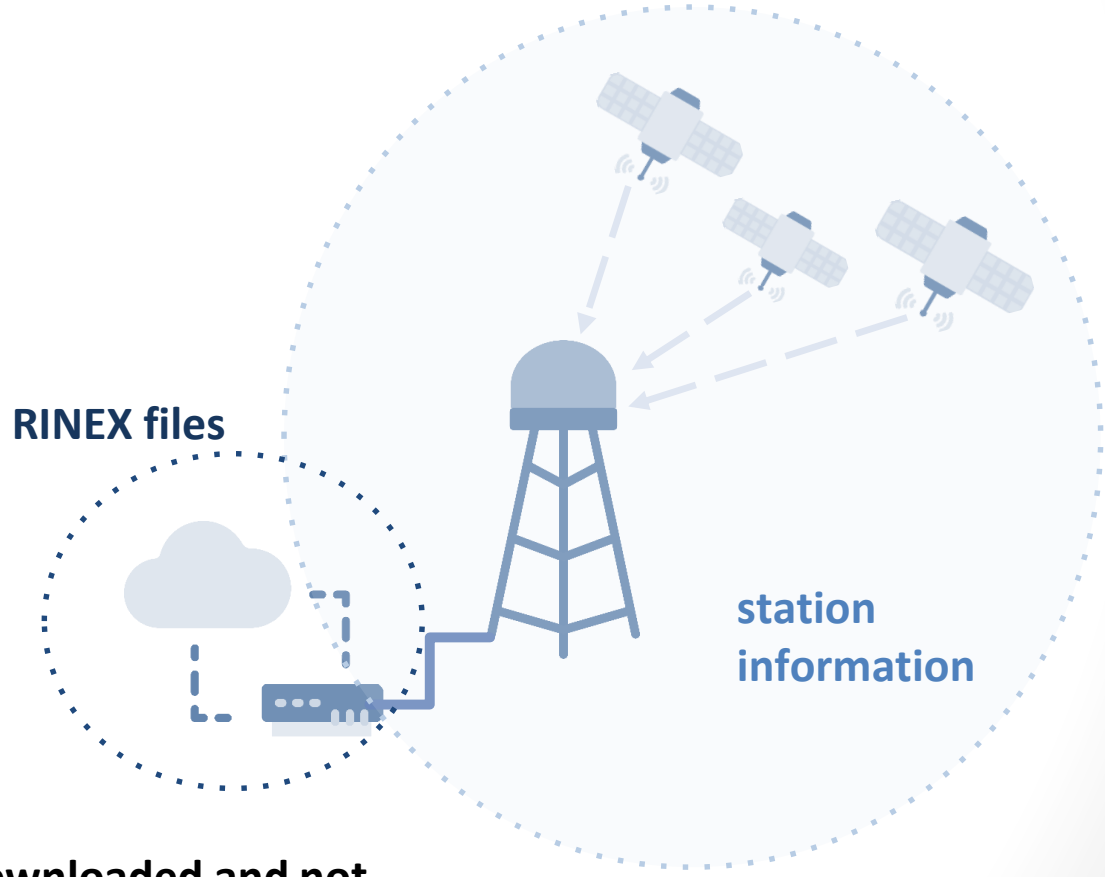
Historical archive of daily RINEX data of EPN stations

For each station

- Site log
- Data license (if provided)
- Networks to which station belongs
- Precise coordinates

For each RINEX file

- File size
- Owner of data
- Information about included observations
- Data quality metrics
- ….



All info is there, but it's not used for selecting files to be downloaded and not provided with the downloaded RINEX files.

Solution?

Evolve towards FAIR GNSS data

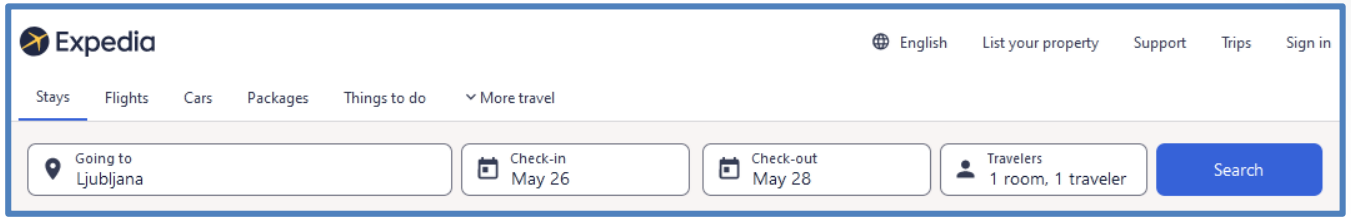
Findable
Accessible
Interoperable
Reusable

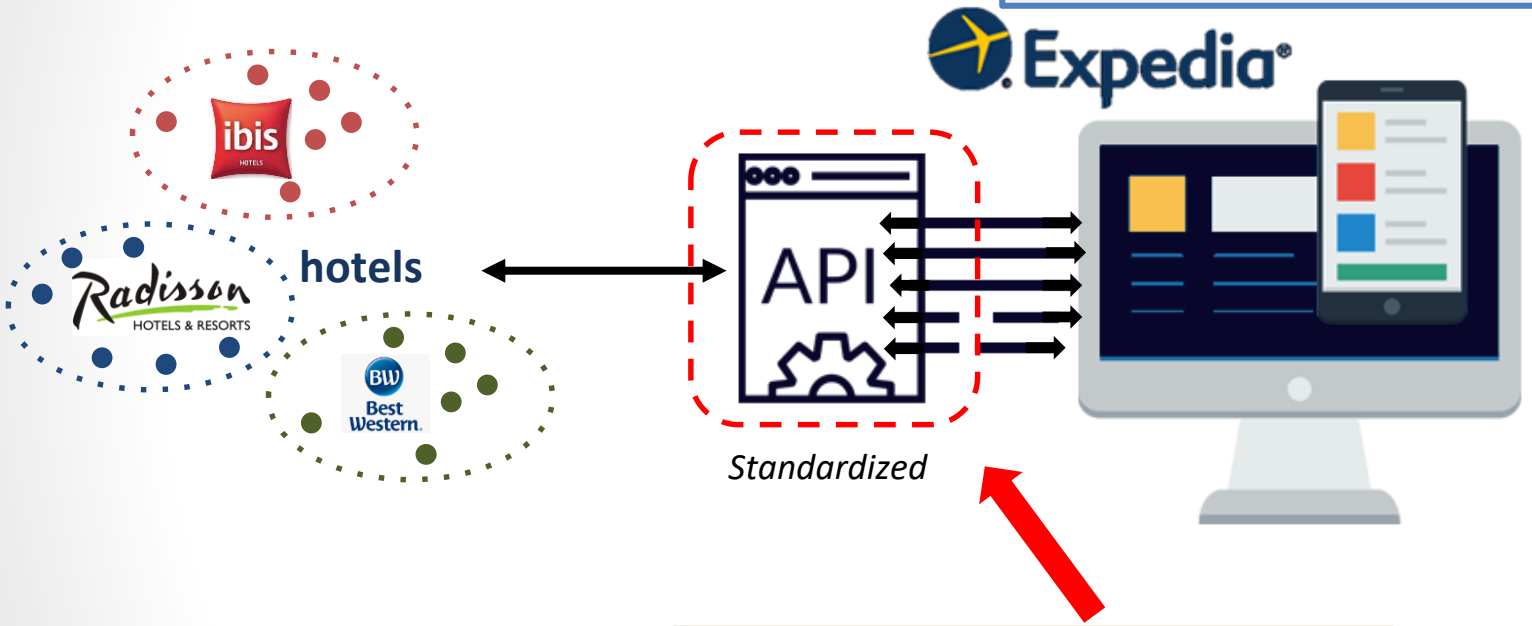
Provide API (Application Program Interface) to access the data
✓ allow to use search criteria
✓ provide all necessary information (=metadata) about the data/station

Has to be done in standardized way to ensure machine-readability and interoperability with other datasets

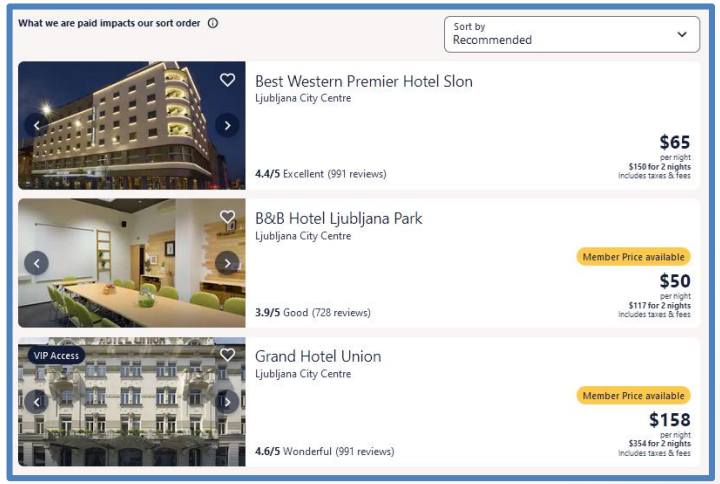
What are APIs?

… an everyday life example of an API





Application Programming Interface (API) :
a software that allows two applications to communicate with each other



https://www.expedia.com/static/mobile/APIConsole/hotel.html

FAIR applied to GNSS

ROYAL OBSERVATORY OF BELGIUM

FAIR GNSS data
EUREF Symposium
May 27, 2021

GHENT UNIVERSITY

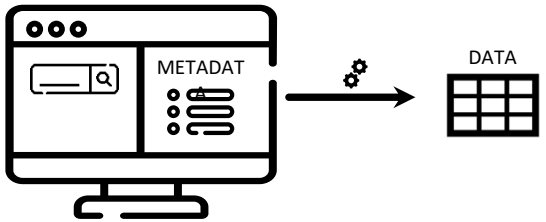
STEP 1

Assign a persistent identifier (**PID**)

Globally unique and **persistent identifiers**

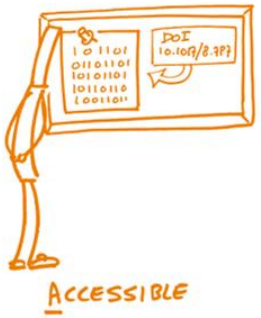
STEP 2

Data are described with **rich metadata**

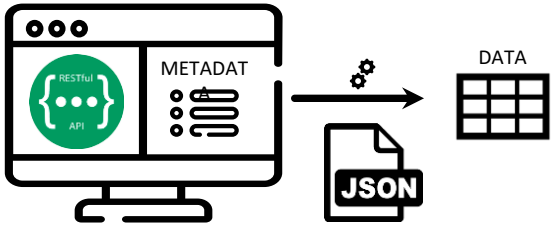


Metadata, machine-readable and structured documentation

STEP 3



Standard communications protocol to retrieve (meta)data



APIs (Application Programming Interface)

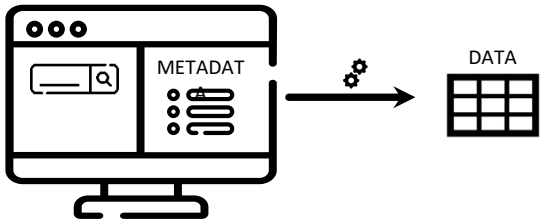
STEP 1

Assign a persistent identifier (**PID**)

Globally unique and **persistent identifiers**

STEP 2

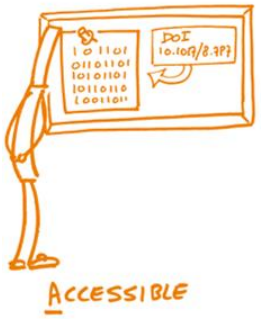
Data are described with **rich metadata**

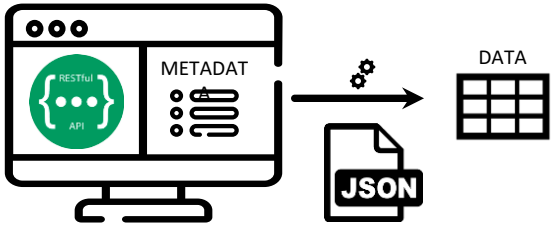


Metadata, machine-readable and structured documentation

STEP 3

Standard communications protocol to retrieve (meta)data





APIs (Application Programming Interface)

ROYAL OBSERVATORY OF BELGIUM

FAIR GNSS data

EUREF Symposium

May 27, 2021

GHENT UNIVERSITY

Persistent identifiers (PIDs) for GNSS data

Most obvious choice: Digital Object Identifiers provided by DataCite

- ✓ *Seismological community*
- ✓ *UNAVCO, GFZ, CDDIS, EPOS, ROB, …*
- ✓ *GGOS working group on DOI ! Common approach to assign DOI to geodetic data and products*

Digital Object Identifiers (DOIs) for data

DOI is a character string (standardized by ISO) used to uniquely identify an object such as journal articles, research reports and data sets, e.g

https://doi.org/10.24414/FST8-P256 → landing page

Prefix

Suffix

Examples of DOIs

GFZ

https://doi.org/10.5880/GFZ.1.1.2020.001 → landing page

GNSS data of the global GFZ tracking network

Data Citation: Ramatschi, Markus; Bradke, Markus; Nischan, Thomas; Männel, Benjamin (2019): GNSS data of the global GFZ tracking network. V. 1. GFZ Data Services. https://doi.org/10.5880/GFZ.1.1.2020.001

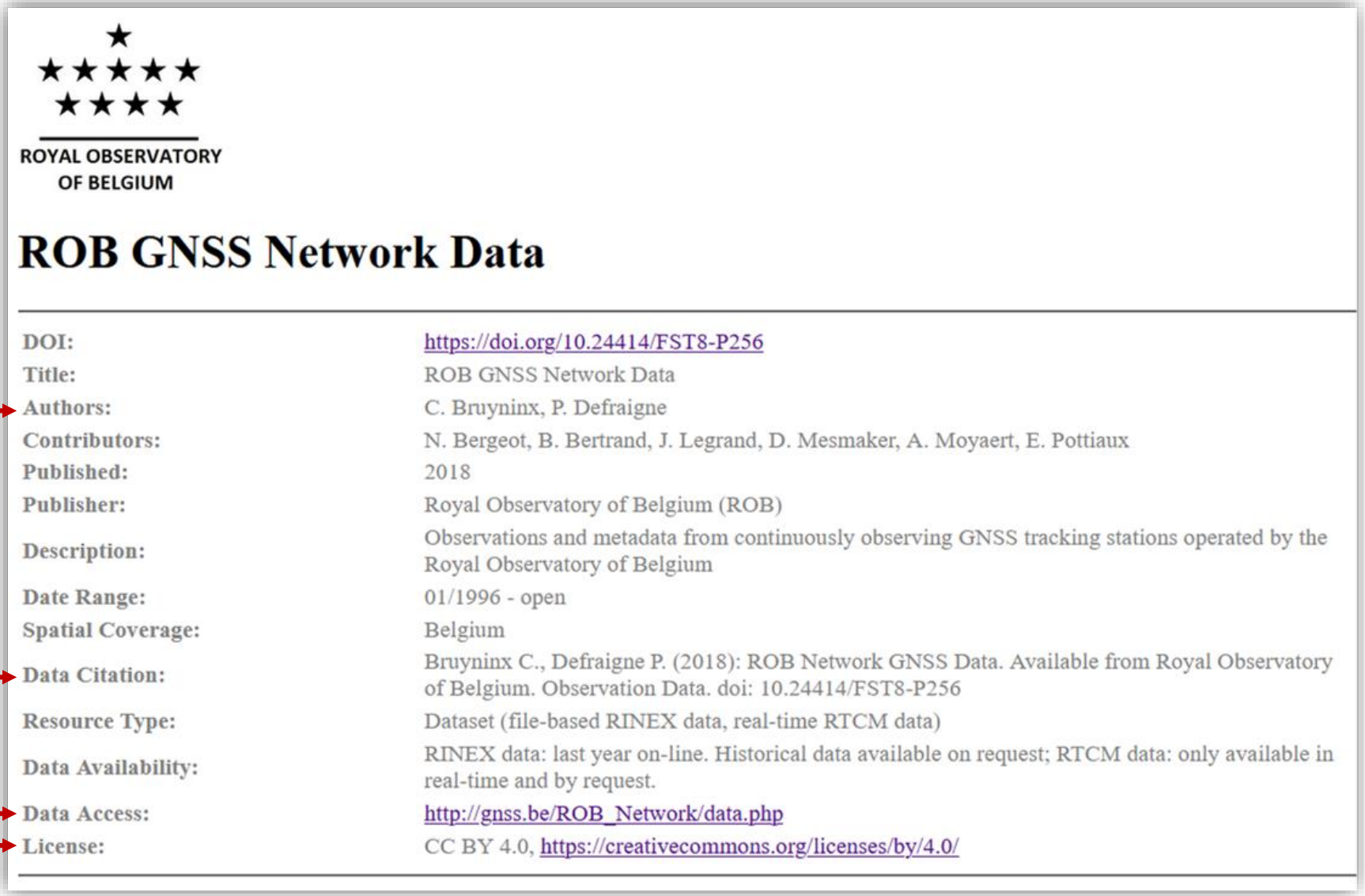
UNAVCO

https://doi.org/10.7283/ZR9Z-K767 → landing page

CVO GPS Network - YOCR-Yocum Ridge P.S. - GPS/GNSS Observations Dataset

Data Citation: Kramer, Rebecca, Pauk, Benjamin, Montgomery-Brown, Emily, 2021, CVO GPS Network - YOCR-Yocum Ridge P.S., The GAGE Facility operated by UNAVCO, Inc., GPS/GNSS Observations Dataset at https://doi.org/10.7283/ZR9Z-K767.

Example of DOI landing page



ROYAL OBSERVATORY OF BELGIUM

ROB GNSS Network Data

DOI:	https://doi.org/10.24414/FST8-P256
Title:	ROB GNSS Network Data
Authors:	C. Bruyninx, P. Defraigne
Contributors:	N. Bergeot, B. Bertrand, J. Legrand, D. Mesmaker, A. Moyaert, E. Pottiaux
Published:	2018
Publisher:	Royal Observatory of Belgium (ROB)
Description:	Observations and metadata from continuously observing GNSS tracking stations operated by the Royal Observatory of Belgium
Date Range:	01/1996 - open
Spatial Coverage:	Belgium
Data Citation:	Bruyninx C., Defraigne P. (2018): ROB Network GNSS Data. Available from Royal Observatory of Belgium. Observation Data. doi: 10.24414/FST8-P256
Resource Type:	Dataset (file-based RINEX data, real-time RTCM data)
Data Availability:	RINEX data: last year on-line. Historical data available on request; RTCM data: only available in real-time and by request.
Data Access:	http://gnss.be/ROB_Network/data.php
License:	CC BY 4.0, https://creativecommons.org/licenses/by/4.0/

Advantages of DOIs

✓ A stable link to the data

✓ Facilitating access to, sharing and reuse of the data
- track data provenance (data owner)
- access specific versions of datasets
- provide the user with information on data access restrictions

✓ Reliable long-term citation of the data
- acknowledge merits of data provider
- provide statistics on data usage for funders

✓ Cross-linking through various objects
- publications and underlying data
- network-DOIs (e.g. for EPN data) and station-specific DOI

How to get a DOI?

https://datacite.org/members.html

1) Agency that is member of DataCite (e.g. TU Delft, GFZ, …)

2) Agency which has a contract with DataCite member agency (e.g. ROB) = DOI minting agency

(Mostly) Data centre responsible for long-term storage of the data

Will assign the DOI to the data

Procedure

- DataCite member agency provides prefix
- DOI minting agency chooses suffix, provides DOI metadata, creates DOI, and maintains DOI landing page

https://doi.org/10.24414/FST8-P256

Step 1: In practice

Through M³G:

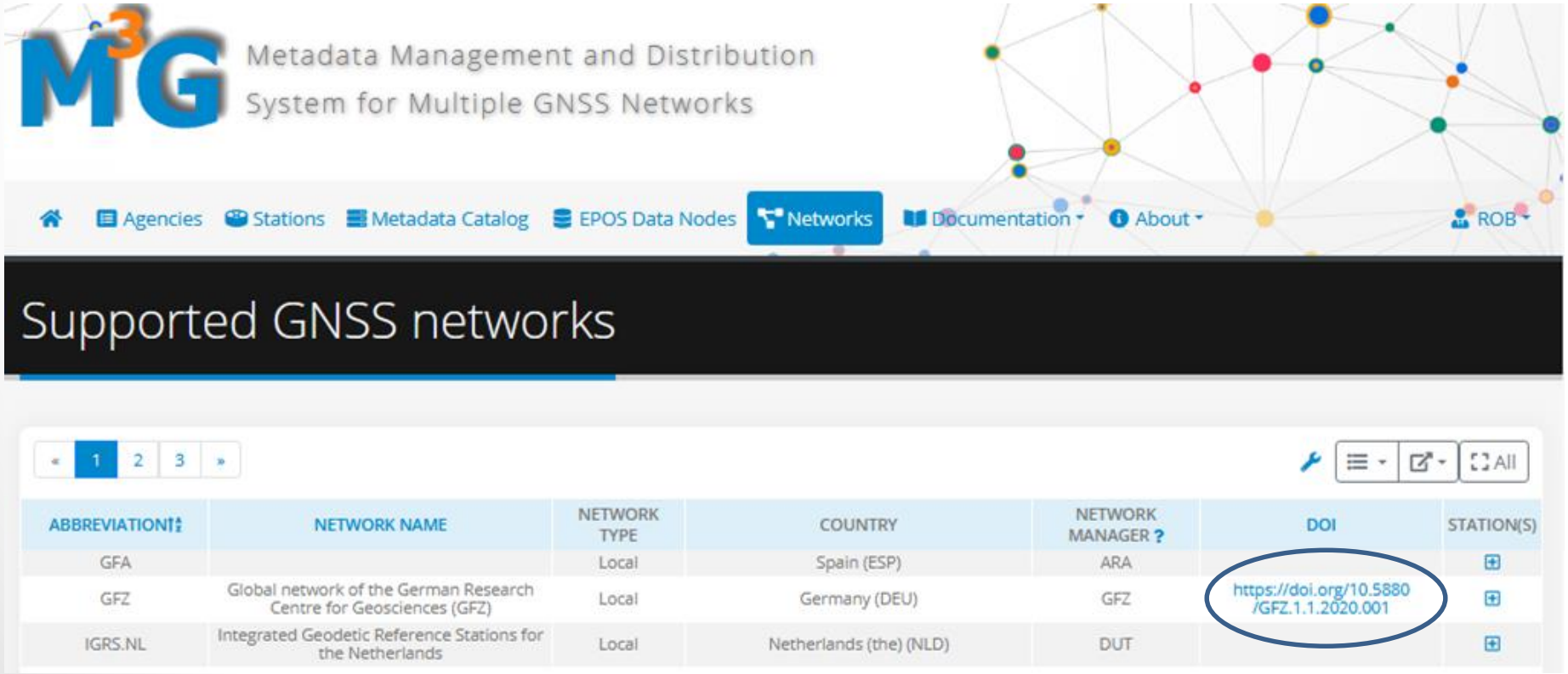
- Collection of existing DOIs for GNSS networks and GNSS stations

STEP 1

Assign a persistent identifier (**PID**)

doi DataCite

Globally unique and **persistent identifiers**

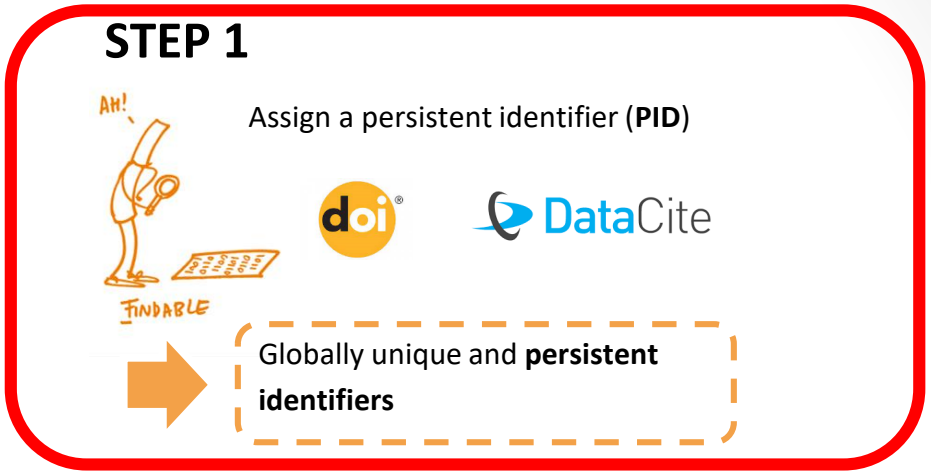


ABBREVIATION	NETWORK NAME	NETWORK TYPE	COUNTRY	NETWORK MANAGER ?	DOI	STATION(S)
GFA		Local	Spain (ESP)	ARA		
GFZ	Global network of the German Research Centre for Geosciences (GFZ)	Local	Germany (DEU)	GFZ	https://doi.org/10.5880/GFZ.1.1.2020.001	
IGRS.NL	Integrated Geodetic Reference Stations for the Netherlands	Local	Netherlands (the) (NLD)	DUT		

Step 1: In practice

Through M^{3}G:

- Collection of existing DOIs for GNSS networks and GNSS stations



EPN historical data archive at ROB:

- Provide service to EPN station managers to assign DOI to their EPN data (only if data does not yet have a DOI!)

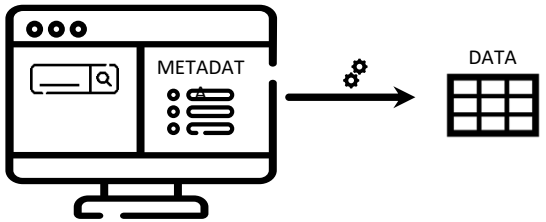
STEP 1

Assign a persistent identifier (**PID**)

Globally unique and **persistent identifiers**

STEP 2

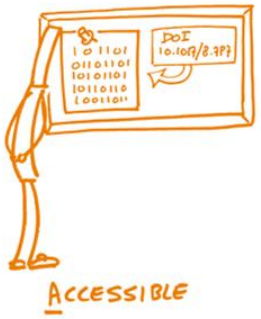
Data are described with **rich metadata**

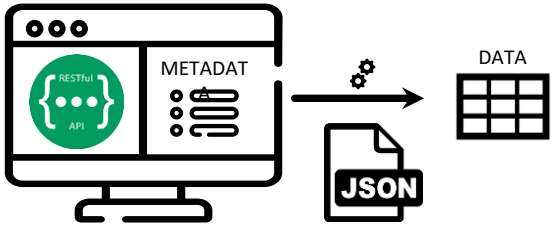


Metadata, machine-readable and structured documentation

STEP 3

Standard communications protocol to retrieve (meta)data





APIs (Application Programming Interface)

Download daily RINEX data

TODAY

BRUX00BEL
02/2021-012/2021

RINEX data files

WTZR00DEU
001/2021-010/2021

RINEX data files

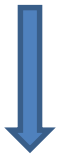
ACOR00ESP
005/2021-010/2021

RINEX data files

Download daily RINEX data

FUTURE

BRUX00BEL
02/2021-012/2021



RINEX data files
Metadata files

WTZR00DEU
001/2021-010/2021

RINEX data files
Metadata files

ACOR00ESP
005/2021-010/2021

RINEX data files
Metadata files

Metadata should provide all info a user needs to know about downloaded data
→ **RICH metadata**

Rich metadata

Station-specific metadata:

Site log + DOI (including data license)

Optionally: precise ETRS89 coordinates, GNSS network to which the station contributes, indiv. antenna calib, …

File-specific metadata (i.e. for each data file and created on-the-fly upon data download):

Data identifier, creation & publication date, type of RINEX (daily or hourly), summary and keywords, data license, observed constellations

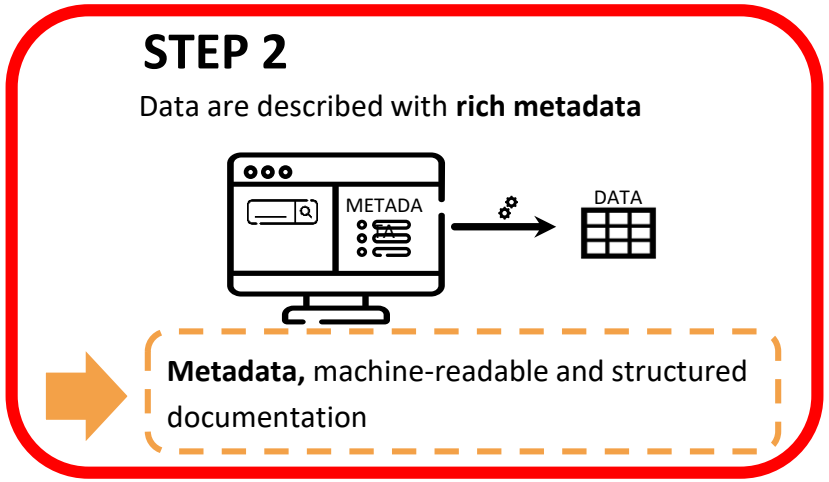
Optionally: quality check results, MD5, file size,…

Requirements

- Allow easy discovery and harvesting of specific data requests by machines (**standardized vocabulary and scheme**) based on PID
- Allow **interoperability** with other data sets (e.g. using Json-LD)

Allow metadata to be exported in various formats (user can select the one most suitable for him)

Step 2: In practice



- Review existing **metadata standards** to reach community-agreement (standardization!)
  - Station metadata:
    - ✓ GeodesyML, working on extension with IGS infrastructure committee (DOI)
  - File metadata:
    - Discuss with organizations providing file-specific metadata (e.g. UNAVCO)
      - ✓ Extended GeodesyML?
      - ✓ (already existing) standardized metadata scheme?
      - ✓ New?

- Review **vocabularies** used
  - Global Change Master Directory (GCMD) Keywords from NASA

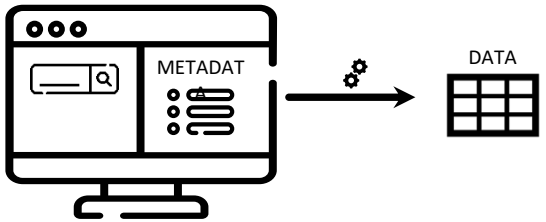
STEP 1

Assign a persistent identifier (**PID**)

Globally unique and **persistent identifiers**

STEP 2

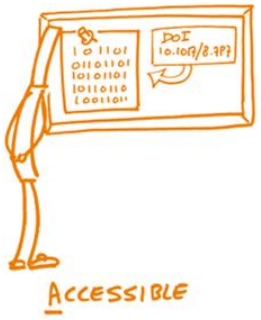
Data are described with **rich metadata**

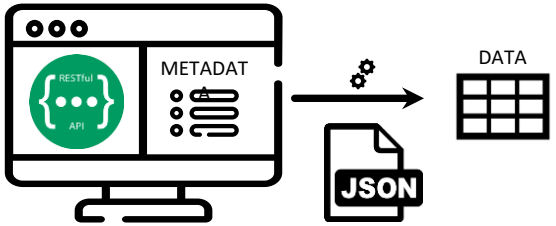


Metadata, machine-readable and structured documentation

STEP 3

Standard communications protocol to retrieve (meta)data





APIs (Application Programming Interface)

ROYAL OBSERVATORY OF BELGIUM

FAIR GNSS data

EUREF Symposium

May 27, 2021

GHENT UNIVERSITY

Web services APIs for GNSS data

Several APIs available to access station descriptions (e.g. site log):

M³G, GFZ, Geoscience Australia

Few APIs to search/access GNSS data:

- BEV: working on it!
- UNAVCO & Geoscience Australia: Beta APIs
- EPOS: under testing

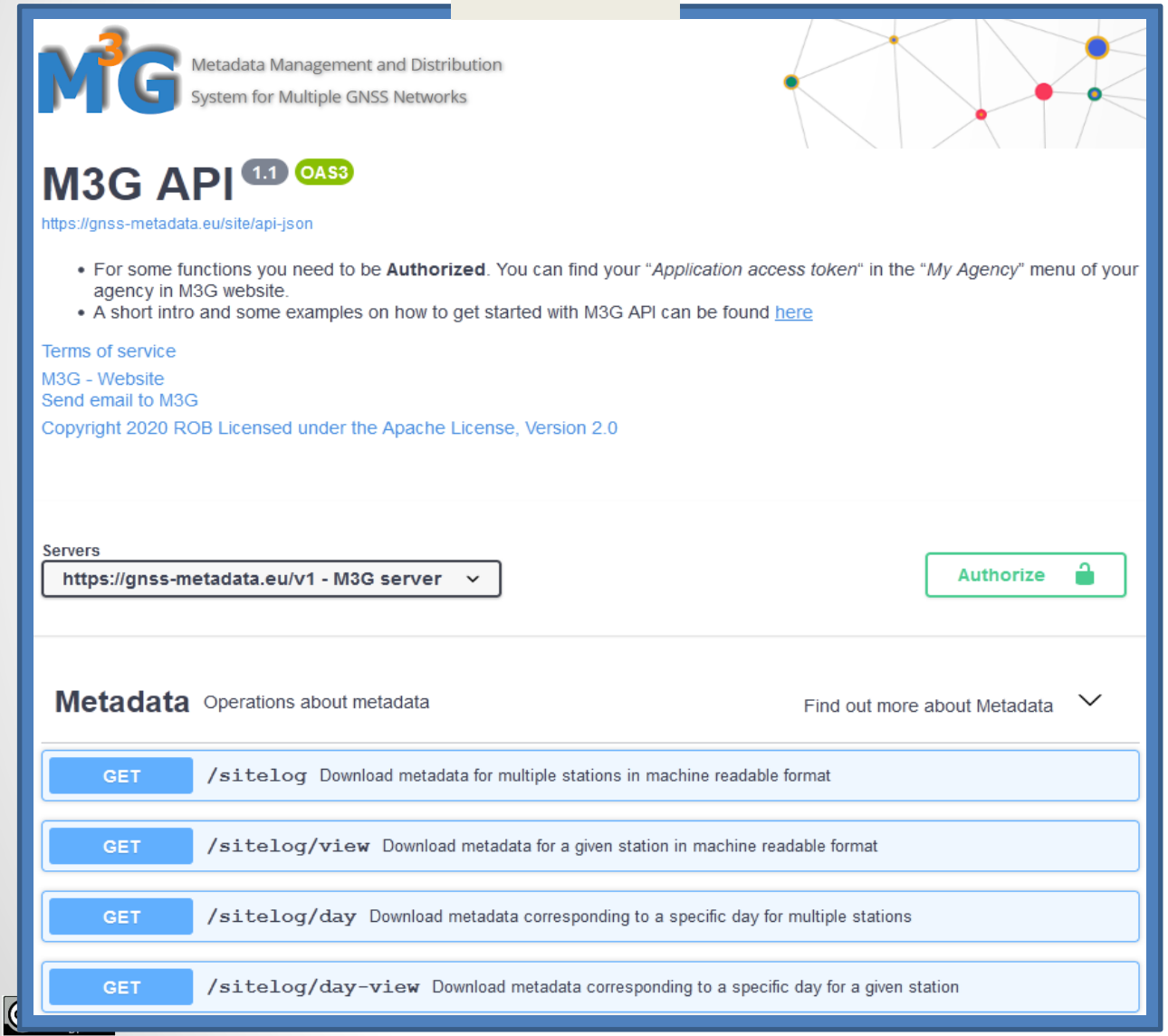
…or private companies:

- SmartNet North America

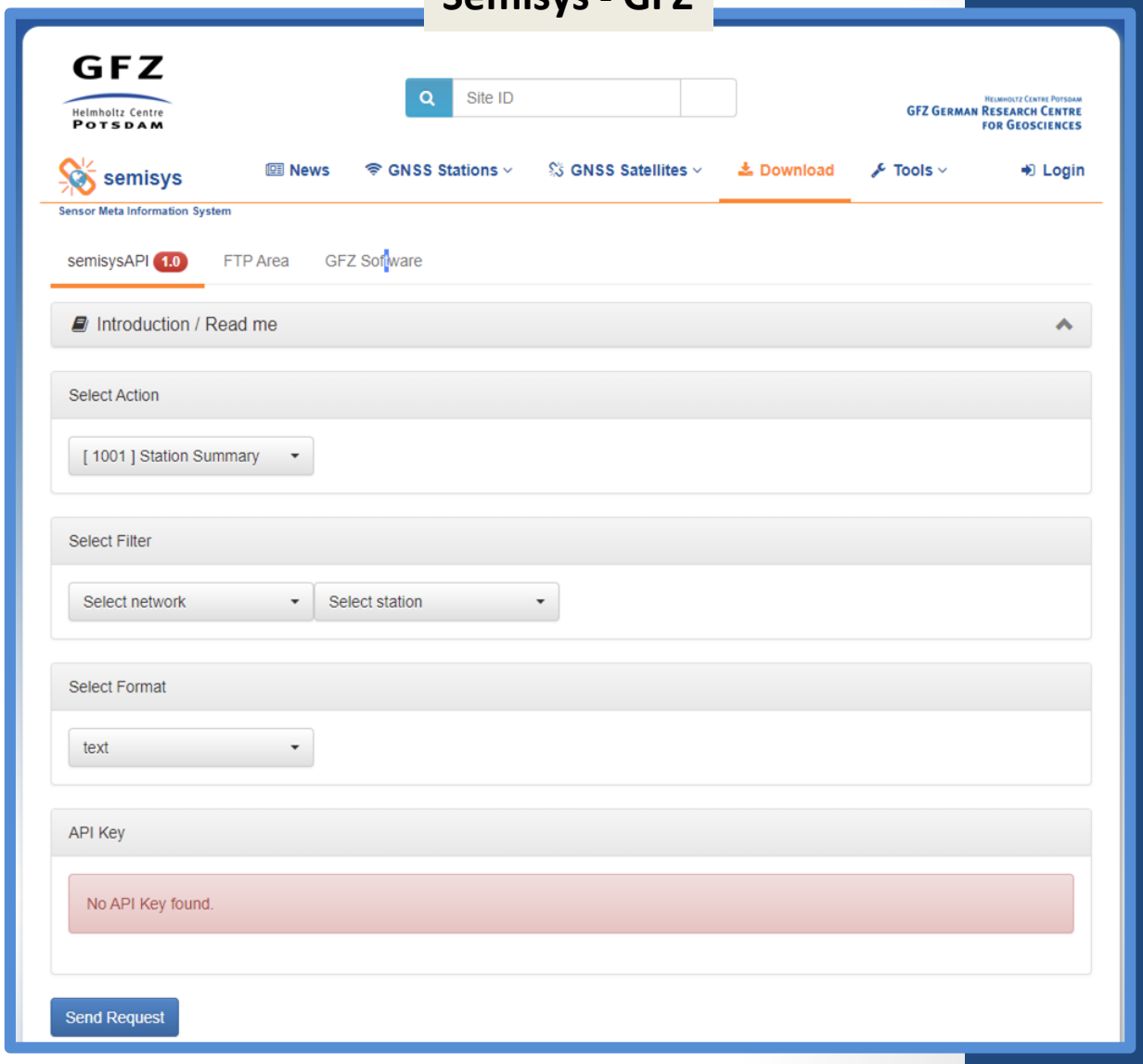
APIs to download GNSS station description

ROYAL OBSERVATORY OF BELGIUM

M³G - ROB



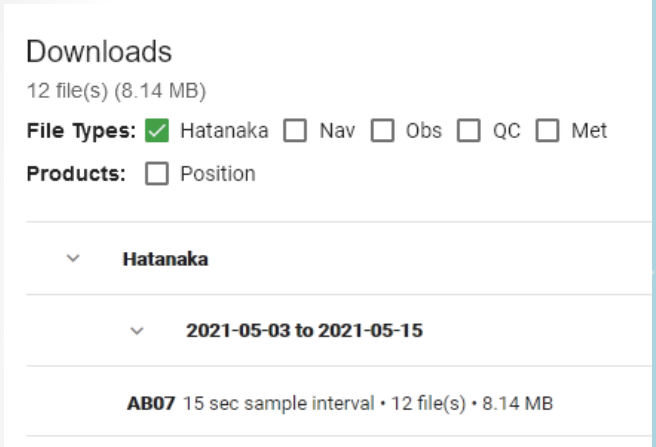
Semisys - GFZ

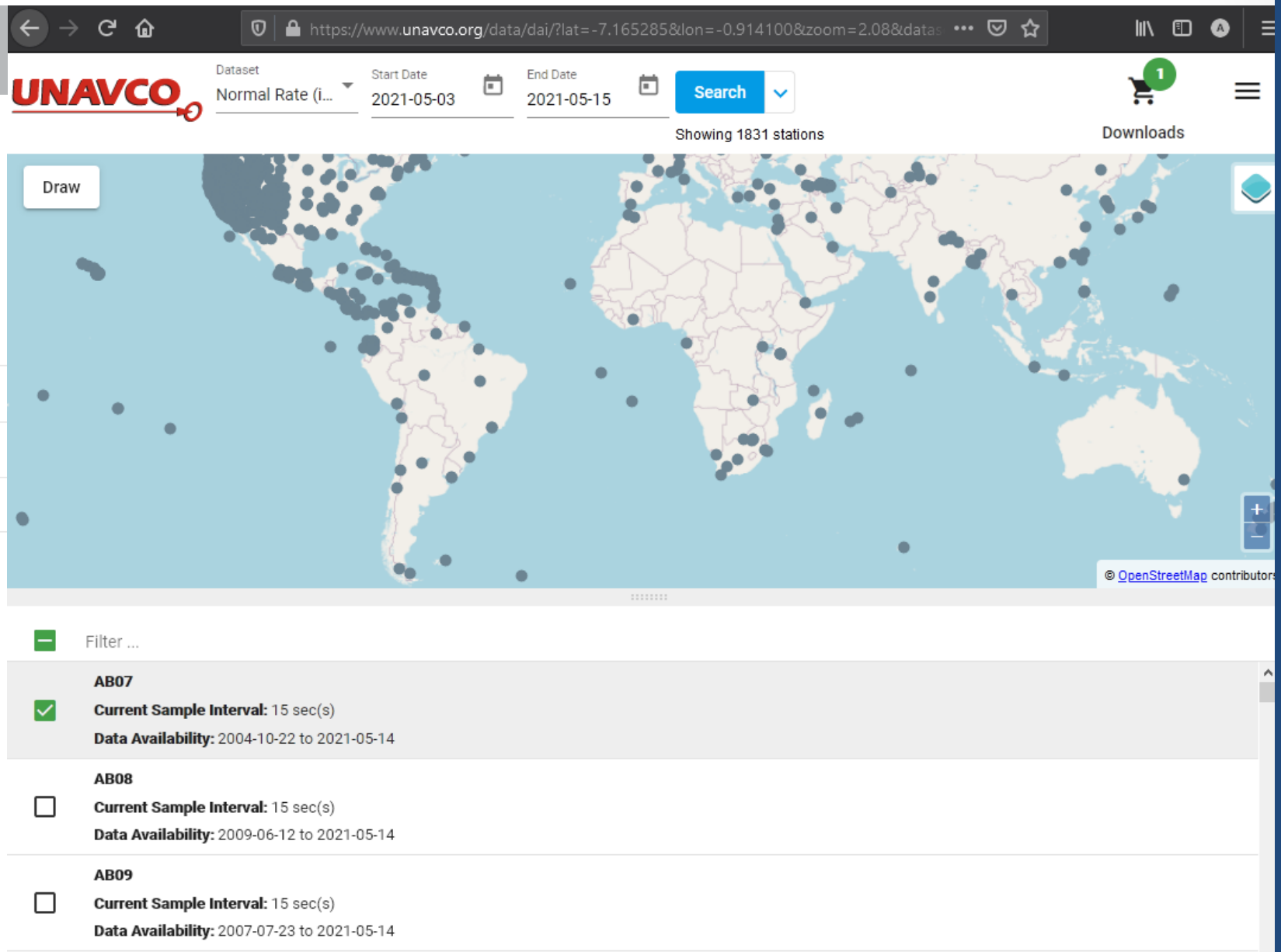


APIs to search/download GNSS data

GAGE Operated by UNAVCO | UNAVCO

https://www.unavco.org/data/dai/





ROYAL OBSERVATORY OF BELGIUM

FAIR GNSS data

EUREF Symposium

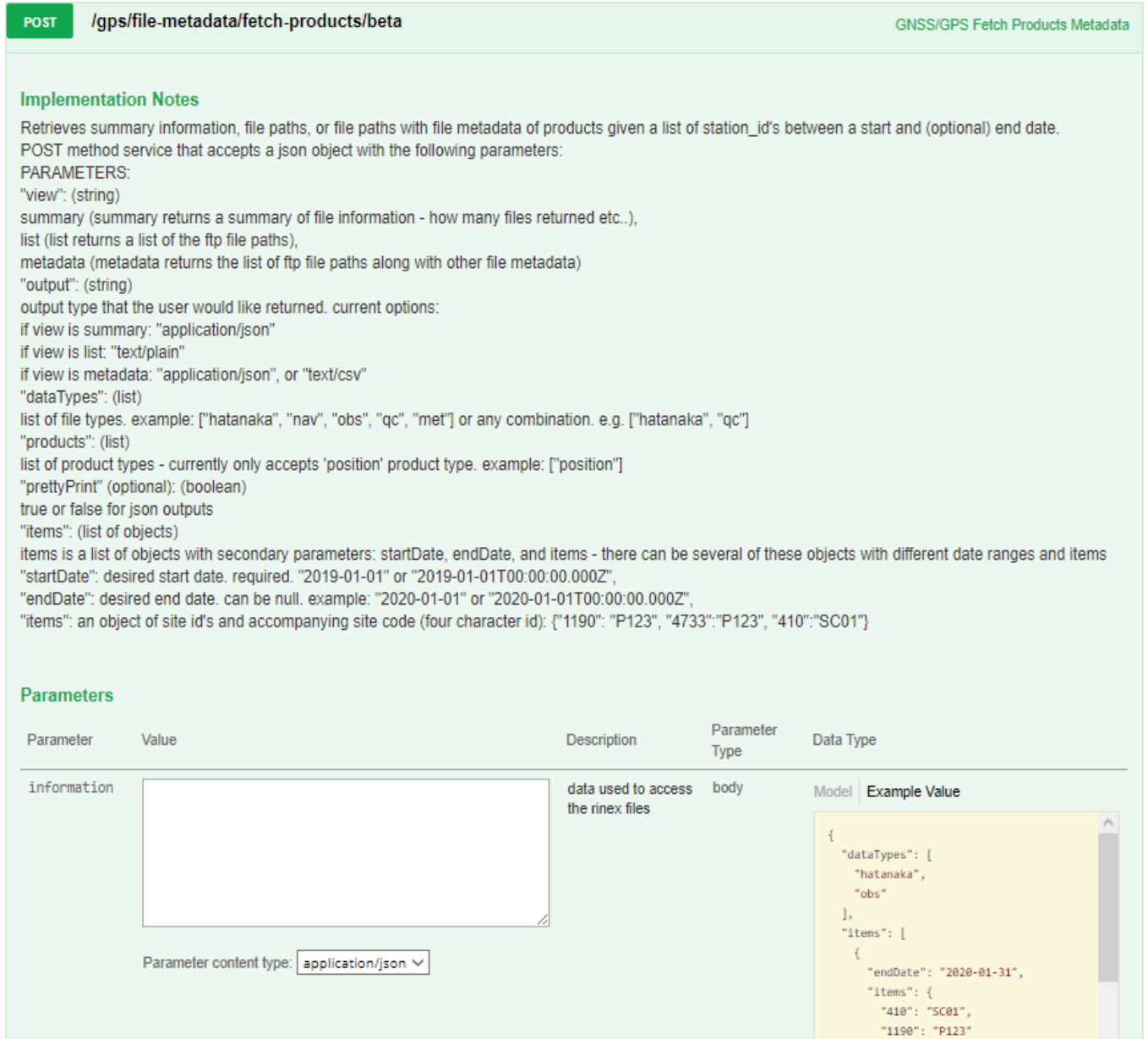
May 27, 2021

GHENT UNIVERSITY

APIs to search/download GNSS data

UNAVCO Unified Web Services

https://www.unavco.org/data/web-services/documentation/documentation.html#!/GNSS47GPS/post_gps_file_metadata_fetch_products_beta



POST /gps/file-metadata/fetch-products/beta — GNSS/GPS Fetch Products Metadata

Implementation Notes

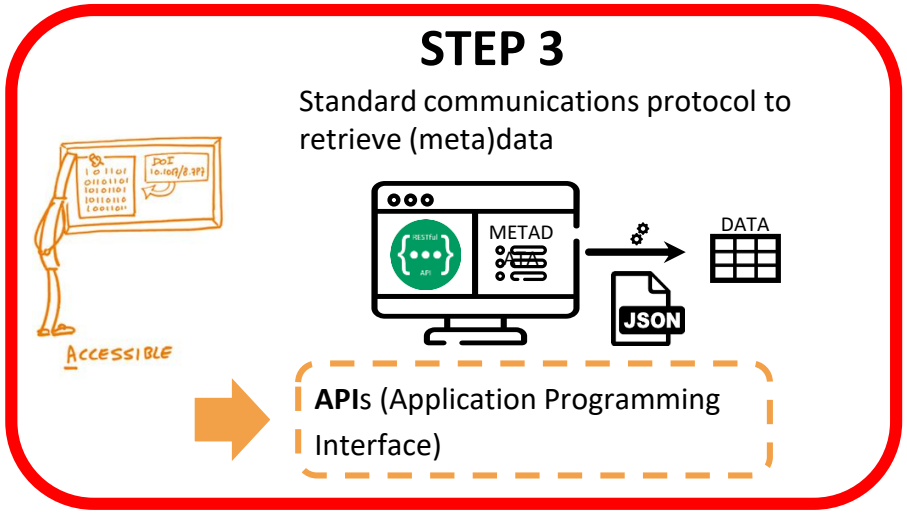
Retrieves summary information, file paths, or file paths with file metadata of products given a list of station_id's between a start and (optional) end date.
POST method service that accepts a json object with the following parameters:
PARAMETERS:
"view": (string)
summary (summary returns a summary of file information - how many files returned etc..),
list (list returns a list of the ftp file paths),
metadata (metadata returns the list of ftp file paths along with other file metadata)
"output": (string)
output type that the user would like returned. current options:
if view is summary: "application/json"
if view is list: "text/plain"
if view is metadata: "application/json", or "text/csv"
"dataTypes": (list)
list of file types. example: ["hatanaka", "nav", "obs", "qc", "met"] or any combination. e.g. ["hatanaka", "qc"]
"products": (list)
list of product types - currently only accepts 'position' product type. example: ["position"]
"prettyPrint" (optional): (boolean)
true or false for json outputs
"items": (list of objects)
items is a list of objects with secondary parameters: startDate, endDate, and items - there can be several of these objects with different date ranges and items
"startDate": desired start date. required. "2019-01-01" or "2019-01-01T00:00:00.000Z",
"endDate": desired end date. can be null. example: "2020-01-01" or "2020-01-01T00:00:00.000Z",
"items": an object of site id's and accompanying site code (four character id): {"1190": "P123", "4733":"P123", "410":"SC01"}

Parameters

Parameter	Value	Description	Parameter Type	Data Type
information		data used to access the rinex files	body	Model / Example Value

Parameter content type: application/json

Step 3: In practice



- Need to have some community-agreed approach on API (type of queries, type of response, metadata scheme,…)
- Study existing APIs and discuss with organizations having/developing APIs
- Discuss with users which kind of search criteria they need
- Create APIs that are as much as possible in-line with what is already existing

Discussion

Questions?

Remarks?

DOI (2021)

Rich metadata (2021)

APIs to download (meta)data (2022)