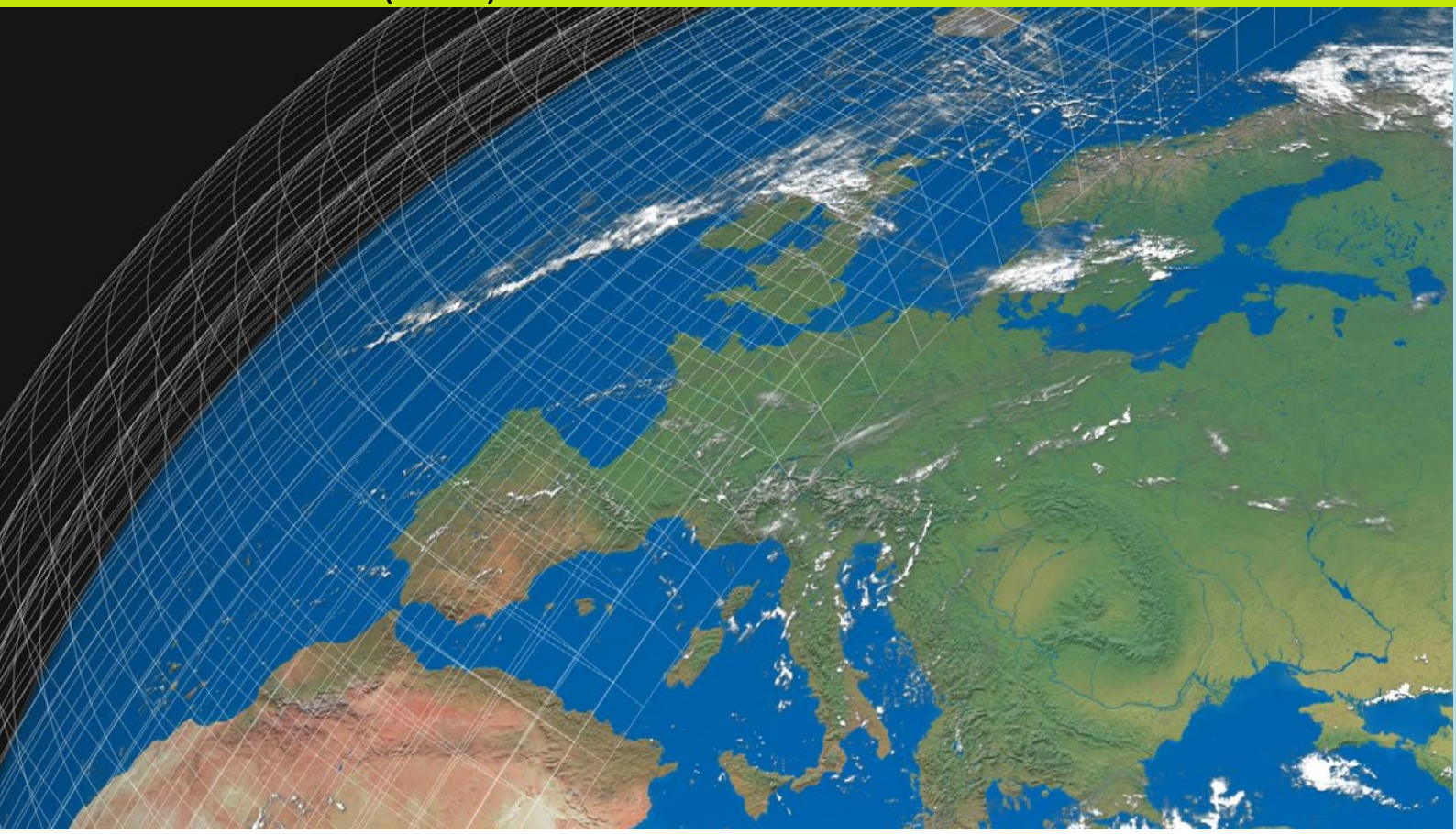
Met Office

WMO Information System 2 (WIS2) Global Cache on Amazon Web Services (AWS)



INTRODUCTION

The World Meteorological Organisation (WMO) is the United Nations system's authoritative voice on the state and behaviour of the Earth's atmosphere, its interaction with the land and oceans, the weather and climate it produces and the resulting distribution of water resources.

WMO defines the authoritative exchange of Earth system data between its Members for the benefit of public, private, academic and other sectors.

The WMO Unified Data Policy, Resolution 1 (Cg-Ext(2021)) defines which Earth system data are necessary for efforts to monitor, understand and predict the weather and climate - including the hydrological cycle, the atmospheric environment and space weather. Data considered essential for provision of services for the protection of life and property and for the well-being of all nations are termed "Core data". Core data is provided on a free and unrestricted basis, without charge and with no conditions on use. "Recommended data" is exchanged in support of Earth system monitoring and prediction efforts. Recommended data may be provided with conditions on use and/or subject to a license. For more information see https://library.wmo.int/idurl/4/58009.

The second generation of the WMO Information System (WIS2) provides the foundation for how this Earth system data is exchanged internationally.

WIS2 leverages the Web and open standards to maximise participation. It is federated system comprised of WIS2 Nodes that publish data and Global Services that enable fault tolerant, highly available, low latency data distribution.

WIS2 provides access to data via a globally distributed set of data servers using the Hyper Text Transfer Protocol (HTTP). Where data is needed for real-time or near-real-time applications, notification messages are published using the MQTT protocol that advertise the availability of new data from those HTTP data servers.

WIS2 is currently pre-operational. It will become operational from January 2025.

For more information, please refer to:

- Manual on the WMO Information System (WMO-No. 1060), Volume II – WMO Information System 2.0 – available from the WMO library in English (https://library.wmo.int/idurl/4/68731), Chinese, Russian, Spanish, Arabic, and French.
- Guide to the WMO Information System (WMO-No. 1061), Volume II – WMO Information System 2.0 – approved by the WMO Commission on Infrastructure and pending publication, currently available as a draft in English (https://wmo-im.github.io/wis2-guide/guide/wis2-guide-DRAFT.html).

As a potential user of WMO Earth system data, or "data consumer", you should read section 1.2 of the Guide to WIS, Volume II: https://wmo-im.github.io/wis2-guide/guide/wis2-guide-DRAFT.html#_data_consumer.

WHAT IS A WIS2 GLOBAL CACHE?

WIS2 uses several Global Caches to distribute Core data globally. Each Global Cache re-publishes Core data on its own highly-available data server and publishes a new notification message advertising the availability of that data from that Global Cache's data server.

There are four Global Caches – including `data-metoffice-noaa-global-cache` that is co-sponsored by the Met Office (https://www.metoffice.gov.uk/) and the National Oceanic and Atmospheric Administration (NOAA) (https://www.noaa.gov/).

`data-metoffice-noaa-global-cache` is hosted on AWS. The contents of this Global Cache are published as open data via the Amazon Sustainability Data Initiative (ASDI).

WHAT IS THIS DATASET?

This dataset contains Earth system data deemed by WMO as essential for provision of services for the protection of life and property and for the well-being of all nations – so called "Core data". The data covers the following Earth system disciplines:

- Weather
- Climate
- Hydrology
- Atmospheric Composition
- Cryosphere
- Oceans
- Space Weather

The initial Catalogue of Core Data is available here: http://library.wmo.int/doc_num.php?explnum_id=11001#page=139. This includes in-situ observations (e.g., from automatic weather stations), remotely sensed observations (e.g., from meteorological satellites) and weather predictions (e.g., from numerical models).

The data is sourced from all 193 countries and territories that are Members of the WMO and is retained for a 24-hour period before being discarded.

HOW IS THE DATA ORGANISED?

All the data is available from an S3 bucket in the 'us-east-1' region: https://wis2globalcache.s3.us-east-1.amazonaws.com/.

Each data object is given a key derived from the unique data identifier in the WIS2 Notification Message ('properties. data_id'). Generally, this is a concatenation of data topic plus an instance identifier.

For example: 'ca-eccc-msc/data/core/weather/surface-based-observations/synop/ISAB05_ CWAO_091200___19759'.

The WIS2 Topic Hierarchy is specified here:
https://github.com/wmo-im/wis2-topic-hierarchy/.

At time of writing, the dataset contains approximately 900,000 objects, with an average size around 80kB. Consequently, browsing the S3 bucket contents is not an effective way to access the data.

Instead, data consumers should subscribe to notifications about the availability of new objects in the dataset – see below.

HOW TO RECEIVE NOTIFICATIONS ABOUT NEWLY AVAILABLE DATA

The primary mechanism to access data from the Global Cache is to subscribe to WIS2 Notification Messages via one of the WIS2 Global Broker services using the MQTT protocol.

For more information on MQTT please see https://www.hivemq.com/blog/how-to-get-started-with-mqtt/.

The WIS2 Notification Message itself is encoded as GeoJSON
(RFC 7946 https://datatracker.ietf.org/doc/html/rfc7946).
Technical Specifications are provided here https://github.com/wmo-im/wis2-notification-message.

Connections:

- There are four Global Brokers to choose from. Connection details for the Meteo France Global Broker are as follows –
  - Hostname: globalbroker.meteo.fr
  - Port: 8883 (MQTT) or 443 (MQTT over Websocket)
  - MQTT version: 5.0
  - TLS: yes
  - Username: everyone
  - Password: everyone

SUBSCRIPTIONS:

- You may subscribe to all messages using the topic 'cache/a/wis2/#'
- Alternatively, you may subscribe to subsets of messages. For example, the topic 'cache/a/wis2/ca-eccc-msc/#' will provide messages related data from the Meteorological Service of Canada (Environment and Climate Change Canada), where as 'cache/a/wis2/+/data/core/weather/surface-based-observations/#' will provide messages relating to surface-based weather observations from all WMO Members. For more information on MQTT topics, please see https://www.hivemq.com/blog/mqtt-essentials-part-5-mqtt-topics-best-practices/. For details of the WIS2 Topic Hierarchy, see https://github.com/wmo-im/wis2-topic-hierarchy/.

Please note that the Global Broker will provide WIS2 Notification Messages from all Global Caches. Client-side filtering is needed to identify messages published by 'data-metoffice-noaa-global-cache': within a 'links' object where 'links.rel' equals 'canonical', 'links.href' must begin with 'https://wis2-global-cache.s3.eu-west-2.amazonaws. com/'.

FUTURE PLANS

Met Office and NOAA have plans to make this dataset easier to use within the AWS ecosystem. Planned enhancements include:

- Providing Spatio Temporal Asset Catalogue (STAC) metadata for the Global Cache content. Using ASDI's STAC service, users will be able to search and browse the Global Cache to find the content they need. For more information on STAC see https://stacspec.org/.
- Providing AWS-native notifications via the AWS Simple Notification Service (SNS). Users will be able to subscribe to notifications for data available from the Global Cache service on ASDI.

Produced by the Met Office. Met Office and the Met Office logo are registered trademarks. © Crown Copyright 2024, Met Office 02534