Draft guidance on technical specifications of WIS2

World Meteorological Organization

Date: 2023-05-22

Version: 2.0.0alpha1

Document location: https://community.wmo.int/wis2-guide

Standing Committee on Information Management and Technology (SC-IMT)[1]

Commission for Observation, Infrastructure and Information Systems (INFCOM)[2]

Date: 2023-05-22

Version: 2.0.0alpha1

Table of Contents

Chapter 1. Preamble

1.1. Purpose

The purpose of this document is to provide technical guidance and additional information for participants in the WIS2 pilot phase and other early adopters of WIS2 to assist them in implementing the practices, procedures, and specifications defined in the Manual on WIS (WMO No. 1060), Vol II.

1.2. Audience

This draft guidance document is provided primarily for technical teams of organisations participating in the WIS2 pilot phase.

1.3. Status of this document

This draft guidance is as an informal document. The content will be updated, modified and improved throughout the WIS2 pilot phase as detailed specifications and procedures are validated through testing. The information provided herein will be consolidated into a new version of the Guide to WIS (WMO No. 1061) upon completion of the pilot phase.

[1] https://community.wmo.int/gvernance/commission-membership/commission-observation-infrastructures-and-information-systems-infcom/commission-infrastructure-officers/infcom-management-group/standing-committee-information-management-and-technology-sc-imt

[2] https://community.wmo.int/governance/commission-membership/infcom

Chapter 2. Introduction of WIS2

2.1. Why WIS2?

The WMO Information System (WIS) provides high-performance and reliable data sharing for all WMO Programmes. WIS2 evolves the original solution that became operational in 2012 in response to technical advances, increases in data volumes, and increased diversity in the people and systems that seek to consume data generated, curated, and shared by WMO Programmes.

From a business requirements perspective, WIS2 supports the following initiatives:

1. Response to emerging data issues.

2. Support for implementation of the Global Basic Observing Network (GBON).

3. Support for implementation of the WMO Unified Data Policy (WMO Res. 1).

4. Improved support for WMO Programmes beyond the World Weather Watch.

5. Enabling low barrier and open access to weather, water, and climate data.

These business requirements translate into the following functional requirements:

1. Use of open standards.

2. Simple Dataset discovery.

3. Integration with the World Wide Web.

4. Provision of cloud-ready solutions.

5. Enablement of big data sharing.

6. Replacement of the Global Telecommunications System (GTS).

Four Technical Foundations are central to how WIS evolved to meet these business and functional requirements:

1. Scalable Web-services architecture.

2. Dataset discovery using the Open Geospatial Consortium (OGC) API Records standard.

3. Real-time notifications about data availability using the OASIS MQTT messaging protocol.

4. Unified monitoring using the OpenMetrics cloud-native, highly scalable metrics protocol (IETF draft)

Underpinning these WIS2 Technical Foundations are the Principles of WIS2. These are described in Appendix A to the Manual on WIS (WMO No. 1060), Vol II. WIS2.

2.2. What is WIS2?

WIS2 is a federated "system of systems". All WIS Centres will comply with the Technical Regulations defined in the Manual on WIS (WMO No. 1060) with regards to how they share data in support of WMO Programme objectives.

In addition, Global Information System Centres (GISC) coordinate the operation of WIS within their Area of Responsibility (AoR) and collectively provide Global Services that ensure highly available data sharing within the WIS2 ecosystem.
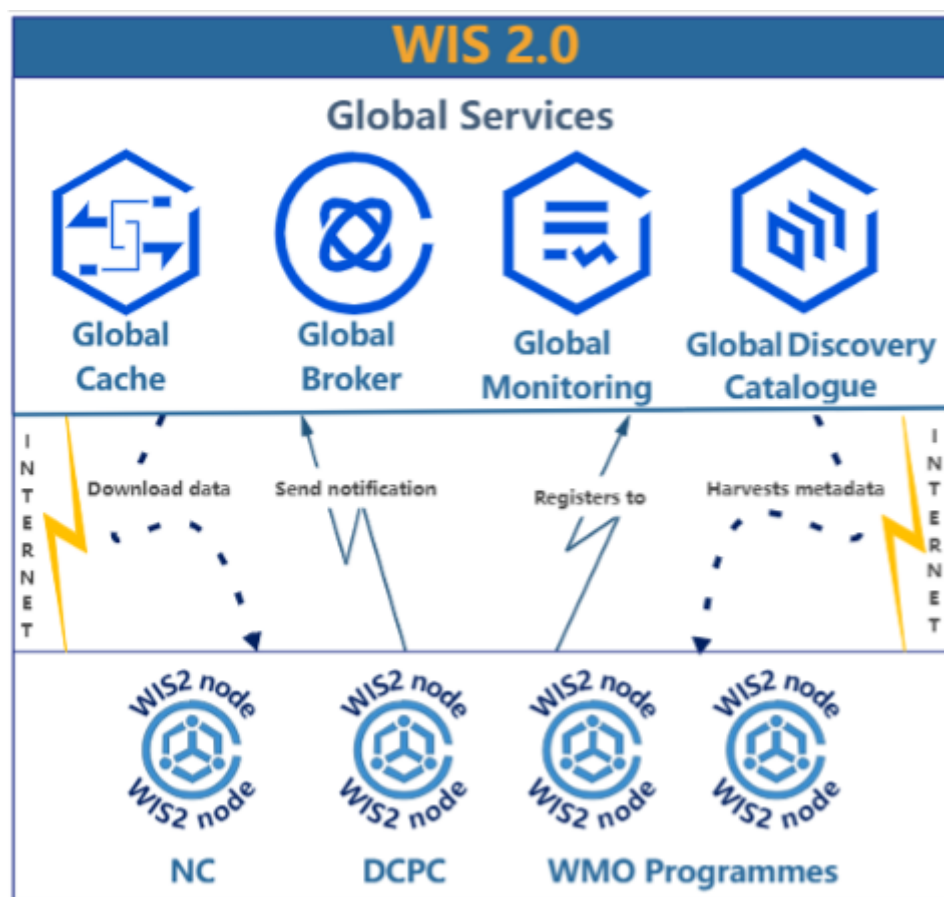
Figure 1. Schematic overview of WIS2

The list of NCs, DCPCs, and GISCs is available in Appendix D to the Manual on WIS (WMO No. 1060), Vol II. WIS2: Approved WIS Centres [3].

When describing the functions of WIS2 there are 3 roles to consider:

1. Data Publisher:

   ◦ This role is fulfilled by NC and DCPC.

   ◦ Data Publishers operate a "WIS node" to share their data within the WIS2 ecosystem.

   ◦ Data Publishers manage, curate, and provide access to one or more "Datasets".

   ◦ For each Dataset, a Data Publisher provides:

     1. "Discovery metadata" to describe the Dataset, provide details on how it can be accessed, and under what conditions.

     2. An API or Web-service to access (or interact with) the Dataset.

     3. Notification messages advertising the availability of new data and metadata.

2. Global Coordinator:

   ◦ This role is exclusive to GISCs.

   ◦ All GISCs supporting WMO Members in their AoR fulfil their data sharing obligations via

WIS2.

- Sufficient to provide highly available global coverage, a GISC may operate one or more Global Services –

    1. Global Discovery Catalogue: enables users (humans and software agents) to search all Datasets provided by Data Publishers and discover where and how to interact with those Datasets (e.g., subscribe to updates, access/download/visualize data, or access more detailed information about the Dataset).

    2. Global Broker: provides highly available messaging services where users may subscribe to notifications about all Datasets provided by Data Publishers.

    3. Global Cache: provides highly available download service for cached copies of core data downloaded from Data Publishers' Web-services.

    4. Global Monitor: gathers and displays system performance, data availability, and other metrics from all WIS nodes and Global Services.

3. Data Consumer:

    - This role represents anyone wanting to find, access, and use data from WIS2 – examples include (but are not limited to): NMHS, government agency, research institution, private sector organisation, etc.

    - Searches or browses the Global Discovery Catalogue (or other search engine) to discover the Dataset(s) that meet their needs (i.e., "Datasets of interest").

    - Subscribes via the Global Broker to receive notification messages about the availability of data or metadata associated with Datasets of interest.

    - Determines whether the data or metadata referenced in notification messages is required.

    - Downloads data from Global Cache or WIS node.

"Quickstart" guidance is provided in Quickstarts TODO: ref for each of these roles.

More details on these aspects are provided in Manual on WIS (WMO No. 1060), Vol II. Part III. Functions of WIS. Section 3.2 (Functional Architecture of WIS) provides a simple overview of how WIS2 functions

2.3. Why are datasets so important?

The purpose of WIS is to enable sharing of data. "Datasets" provide the mechanism for us to describe what data we publish, or what data we're interested in.

Data comes in all shapes and sizes. Often, our data is continually updated (e.g., new observations from an automatic weather station). Sometimes it is very, very big (e.g., a set of climate predictions or a multi-decadal reanalysis). WIS2 aims to make it simple to talk about collections of data that share similar characteristics.

We call these collections of data "Datasets".

A Dataset may comprise of thousands of individual files – like scene from an imaging satellite, or fields from a numerical weather prediction model. Or it may comprise huge numbers of records in

a database – like a time-series of physical property values observed at numerous locations.

However, the data is stored, we describe the collection of files, records, whatever (!) as a Dataset. But there are some rules. The data in a Dataset must be consistent – it must have the same type (e.g., time-series at point locations), subject (e.g., observations of physical properties), and any spatiotemporal referencing so that people can structure sensible queries on the Dataset. A Dataset must also have a single usage or access policy, which implies a Dataset will be subject to a single governance regime and will likely have a single owner. Most importantly, the contents of a Dataset must be available from a single access point so that we can tell Data Consumers where they can download data from. We also need to use a single topic for publishing notifications about the availability of new data within a Dataset.

It is up to Data Publishers to decide how they group data into Datasets. That said, we recommend that, subject to the consistency rules, Data Publishers should organise their data into as few Datasets as possible. For a Data Publisher this means fewer discovery metadata records to maintain. For a Data Consumer this means fewer topics to subscribe to, and lower barrier APIs and Web services to access.

A real-world example for an NMHS to package all synoptic observations for the entire country or territory into a single Dataset [4] [5]. Described with a single discovery metadata record. Accessible via a single Web-service.

Finally, note that the notification messages used to advertise the availability of new data themselves contain metadata (time, location, etc.). This allows for Data Consumers to decide whether or not they want to download a specific file or record.

[3] As of October 2022 the list of approved WIS centres is empty because none have been formally designated as WIS2 centres, having completing the migration from WIS/GTS to WIS2. For designated "WIS 1" centres please refer to Appendix B of the Manual on WIS (WMO No. 1060), Vol I.

[4] Probably keep archived synoptic observations separate from those being collected in (near) real-time because they're likely accessed through different Web-services.

[5] This is contrary to earlier practices in WIS where each GTS bulletin was considered a Dataset. This meant that the WIS Catalogue became overburdened with hundreds of thousands of records, each describing more-or-less the same thing. The result was a very poor search experience – always too many 'hits'!

Chapter 3. Quickstarts

This section lists the sections in this technical guidance document that are essential for Data Publishers operating a WIS node (e.g., an NC or a DCPC) and Data Consumers to read before getting started. Operators of GlSCs should have an in-depth understanding of how WIS2 functions and operates – not least so that they are able to effectively support the WIS Centres, Data Publishers, and Data Consumers in their Area of Responsibility (AoR). As such, all contents of this guidance document are relevant to GISCs.

3.1. Data publisher

As a Data Publisher planning to operate a WIS node, you should read the following sections: TODO: fix crossrefs * 1.2: What is WIS2? * 1.3 Why are Datasets so important? * PART III Data Metadata flows * PART X Information Management

3.1.1. How to publish Dataset discovery metadata to the Global Discovery Catalogue

A Data Producer publishes discovery metadata to the Global Discovery Catalogue by generating a discovery metadata record (see PART V Discovery Metadata) and publishing it via a WIS Node broker in accordance to the standard topic hierarchy. Global Services subscribed to WIS Nodes then validate, ingest and publish the Dataset discovery metadata record to the Global Discovery Catalogue.

3.1.2. How to publish data notifications

A Data Producer publishes data notifications by generating a notification message (see 7.2.2 Notification_messages) via a WIS2 Node broker in accordance with the standard topic hierarchy. Global Brokers subscribe to WIS2Nodes and then re-publish the data notification. If applicable, Global Cache ingests the associated file or data for re-publication to the Global Broker.

3.1.3. How to publish data to the Global Broker

Data publishing is realized as a function of publishing data notifications (see 2.1.2 How to publish data notifications).

TODO: To be completed

3.2. Data consumer

As a Data Consumer wanting to use data published via WIS2 you should read the guidance presented here. In addition, a list of references to useful sections in this guidance document is provided where Data Consumers can gain a better understanding of how WIS2 works.

3.2.1. How to search the Global Discovery Catalogue to find Datasets

A Data Consumer may browse discovery metadata provided by the Global Discovery Catalogue to

locate a Dataset of interest. Discovery metadata follows a standard scheme (see PART V Discovery Metadata). A Data Consumer may discover a Dataset using keywords, topics, geographic area of interest, or date(s) associated with the Dataset. Matching search results from the Global Discovery Catalogue provide high-level information (title, description, keywords, spatiotemporal extents, data policy, licensing), from which a Data Consumer can assess and evaluate their interest in accessing/downloading the Dataset's data.

A key component of Dataset records in the Global Discovery Catalogue is that of "actionable" links. A Dataset record provides one to many links that clearly identify the nature of the link (informational, direct download, API or Web service) so that the Data Consumer can interact with the data accordingly. For example, a Dataset record may include a link to subscribe to notifications (see 2.2.2 how to subscribe to notification about availability of new data) about the data, or a API or Web service, or an offline archive retrieval service.

TODO: To be completed

3.2.2. How to subscribe to notifications about availability of new data

A Data Consumer may browse the topics provided by a Global Broker to locate a Dataset of interest to them. Topic structures follow a standard scheme (see WIS2 messages Standard topic hierarchy). A Data Consumer may discover more about the Dataset associated with a particular topic by using the topic-name as a key to search the Global Discovery Catalogue.

Data Consumers should subscribe to receive "data availability" messages from Global Brokers. Exceptionally, a Data Consumer may choose to subscribe directly to the local message broker at the originating NC/DCPC – but note that the NC/DCPC may not provide public access to its local message broker. Data Consumers should not subscribe to the local message broker at Global Cache instances.

Real-time and near real-time data designated as "Core" in the WMO Unified Data Policy (Resolution 1) will be available from the Global Cache. As per clause 3.2.13 from the Manual on WIS, Vol II. Data Consumers should access data from the Global Cache.

When a Dataset is available from the Global Cache, its discovery metadata will include a link to subscribe to "data availability" messages from the Global Cache instances. Data Consumers will subscribe to these "Global Cache" topics and not to the topic that contains messages from the originating NC/DCPC.

TODO: To be completed

3.2.3. How to use a notification message to decide whether to download data

Each Global Cache will publish a "data availability" message as the associated data object becomes available from that Global Cache instance. Each "data availability" message will refer to the same logical data object but hosted at a different location within WIS. Data Consumers need to consider their strategy for managing these duplicate messages. From a data perspective, it does not matter which Global Cache instance is used – they will all provide an identical copy of the data object published by the originating NC/DCPC. The simplest strategy is to accept the first "data availability" message and download from the Global Cache instance that the message refers to (i.e., using a URL for the data object at that Global Cache instance). Alternatively, a Data Consumer may have a

preferred Global Cache instance, for example, that is located in the Region. Whichever Global Cache instance is chosen, Data Consumers will need to implement logic to discard duplicate "data availability" messages.

TODO: To be completed

3.2.4. How to download data from the Global Cache and WIS nodes

Dataset links are made available through Dataset discovery metadata (via the Global Discovery Catalogue) as well as data notification messages (via Global Brokers). Links can be used to download the data (according to the network protocol and content description provided in the link) using a mechanism appropriate to the workflow of the Data Consumer. This could include web and/or desktop applications, custom tooling, or other approaches.

TODO: To be completed

3.2.5. Additional sections of interest to a Data Consumer (TODO: fix crossref)

- 1.2 What is WIS2?
- 1.3 Why are Datasets so important?

Chapter 4. Data and metadata flows in WIS2

This section provides a visual overview of how data and discovery metadata are shared within WIS2.

Figure 2 introduces the symbology used throughout this section. TODO: crossref
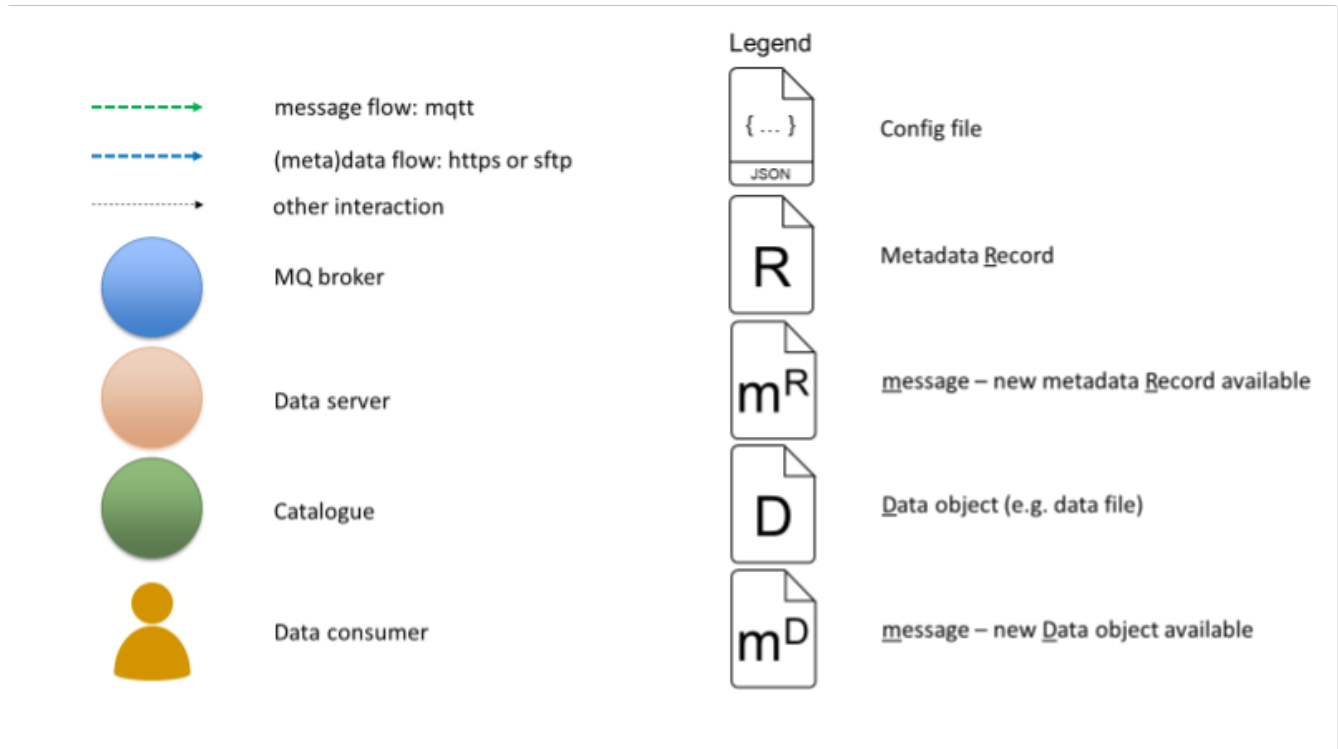
Figure 2. Symbology used to describe data and metadata flows

Introducing the key components:

- **MQ Broker**: Used to publish messages advertising the availability of new/updated/deleted discovery metadata and new Data Objects (e.g. data files that comprise the dataset). Implements MQTT 3.1.1 and/or MQTT 5.0.

- **Data server**: Provides mechanism for Data Consumers to download datasets or part thereof. May be implemented by publishing Data Objects (files) to a Web server or FTP server. Or as an interactive Web service that exposes an API. Supported protocols are HTTPS and SFTP.

- **Catalogue**: Contains discovery metadata records describing the datasets published by the NC/DCPC. May be implemented as a "static catalogue" with JSON files published on a Web server enabling a user to browse the catalogue, or as a "dynamic catalogue" via an application that offers a search API.

4.1. Simple NC/DCPC data sharing

Here we introduce the key concepts of publishing discovery metadata, notification messages, and data. For simplicity, the WIS2 Global Services are not shown here.

Figure 3 illustrates metadata publication and discovery by a Data Consumer.
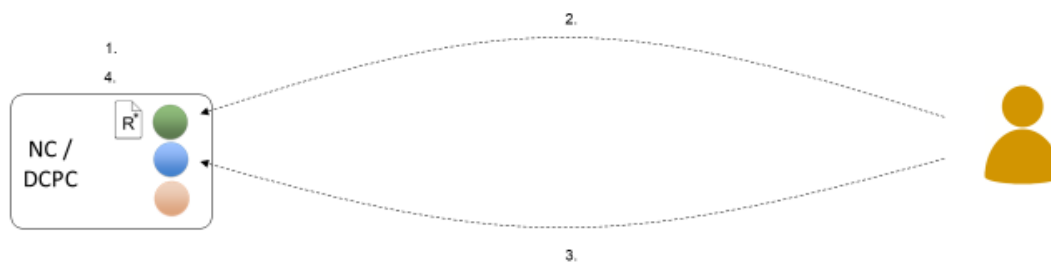
Figure 3. Simple NC/DCPC data sharing - metadata and discovery

1. NC/DCPC shares a new dataset. They publish a discovery metadata record (**R**) to their local catalogue describing this dataset.

2. Data Consumer searches/browses the NC/DCPC's local catalogue and discover a dataset that they need. They use the links to identify how & where they can download the dataset and subscribe to updates (e.g., when new data is added to the dataset).

3. This is a near-real-time dataset that will be routinely updated. The Data Consumer decides to subscribe to the topic for this dataset on the NC/DCPC's local MQ broker.

4. Later, NC/DCPC updates the scope of the dataset (e.g. adding data from another observing station). They publish an updated discovery metadata record (**R***) to their local catalogue

Note: Data Consumer does not need to update their subscription to receive messages about data availability from this new observing station. They have subscribed to updates about the **dataset** - which now includes data from the additional observing station.

Figure 4 illustrates how a Data Consumer subscribes to receive notification messages and responds to receipt of a message by downloading data.
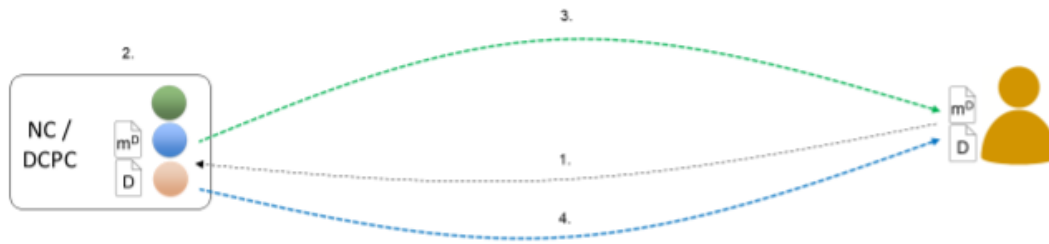
Figure 4. Figure 4 Simple NC/DCPC data sharing - data distribution

1. [Precondition] Data Consumer subscribes to the MQ topic at the NC/DCPC.

2. As the dataset is produced Data Objects (**D**), e.g., observation data BUFR-files, are made available for download, and messages (**m^D**) are published to the topic for this dataset on the local MQ broker.

3. Message (**m^D**) indicating availability of new data are pushed to the subscriber.

4. Data Consumer parses the messages to extract the URL for the new Data Object (**D**) and downloads the file.

Note: In this example, the dataset is published as files that can be downloaded – but the same pattern can be used to advertise new data that is made available via a Web service API.

4.2. Message distribution via Global Broker

Here we add the Global Broker service.

Note that the Global Broker operates two "channels": one for messages originating from NC or DCPCs, the second for messages from Global Cache instances. These are designated "O" and "GC" accordingly. Figure 5 illustrates how messages are published from the NC/DCPC to the Global Broker.
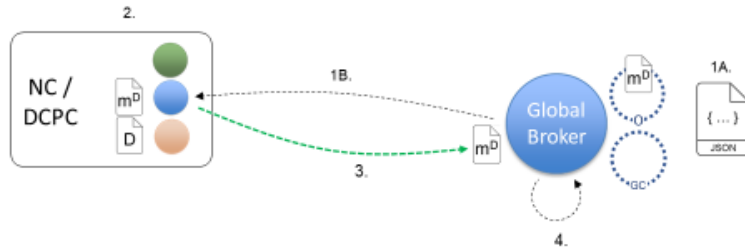
Figure 5. Message distribution via Global Broker - publishing messages from NC/DCPC

1. A[Pre-condition] Bootstrap the Global Broker with a config file that indicates (i) NC/DCPCs and Global Cache instances from which to re-publish messages, and (ii) connection details for at least two other Global Broker instances.

2. B[Precondition] Global Broker determines the MQ broker topics to subscribe to and subscribes to them.

3. As the dataset is produced Data Objects (**D**), e.g., observation data BUFR-files, are made available for download, and messages ($\mathbf{m^D}$) are published to the topic for this dataset on the local MQ broker.

4. Message ($\mathbf{m^D}$) indicating availability of new data are pulled by the subscriber.

5. Message ($\mathbf{m^D}$) is assessed using "anti-loop" logic to avoid resending duplicate messages. If approved, it is added to the appropriate MQ topic of the "originating centre" channel of Global Broker.

6. Message ($\mathbf{m^D}$) is validated against the agreed message format (insert reference). If the message is not compliant, it is discarded.

Figure 6 illustrates how messages are propagated to other Global Broker instances.
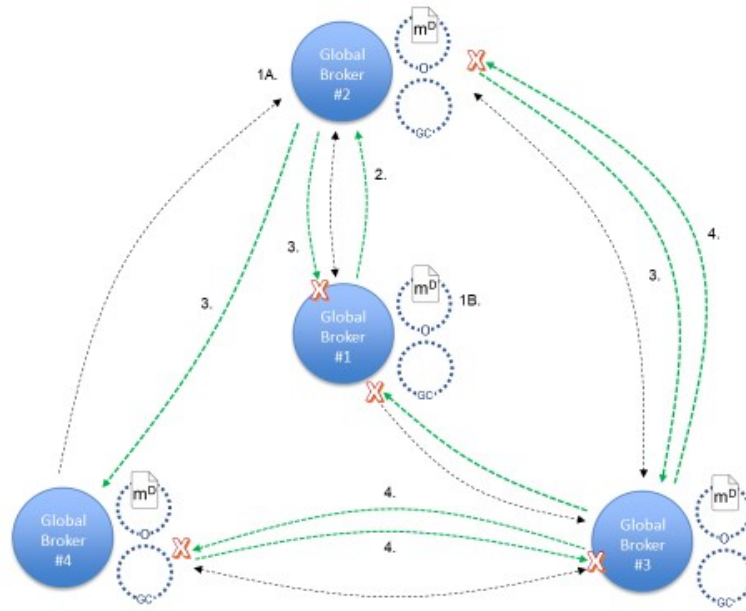
Figure 6. Message distribution via Global Broker – publishing messages from other Global Brokers

1. A[Precondition] Global Broker instances subscribe to MQ topics from at least 2 other Global Broker instances. These are not necessarily bilateral relationships.

2. B[Precondition] Message ($\mathbf{m}^D$) already (re-)published by Global Broker #1.

3. Message ($\mathbf{m}^D$) is pulled by subscriber: Global Broker #2. This is not a duplicate message, so GB#2 republishes the message.

4. Message ($\mathbf{m}^D$) is pulled by subscribers: Global Broker instances #1, #3 and #4. GB#3 and GB#4 have not seen this message before, so they republish the message. However, GB#1 identifies the message as a duplicate and discards it.

5. Message ($\mathbf{m}^D$) is made available to subscribers from Global Broker instances #3 and #4. All Global Broker instances already have a copy of the message, so the new copy of the message is discarded thereby avoiding publication of duplicate messages.

4.3. Data distribution via Global cache

Here we add the Global Cache service.

Figure 7 illustrates how "data availability" notification messages are published by a Global Cache instance once it has downloaded a copy of a data object from an NC or DCPC.
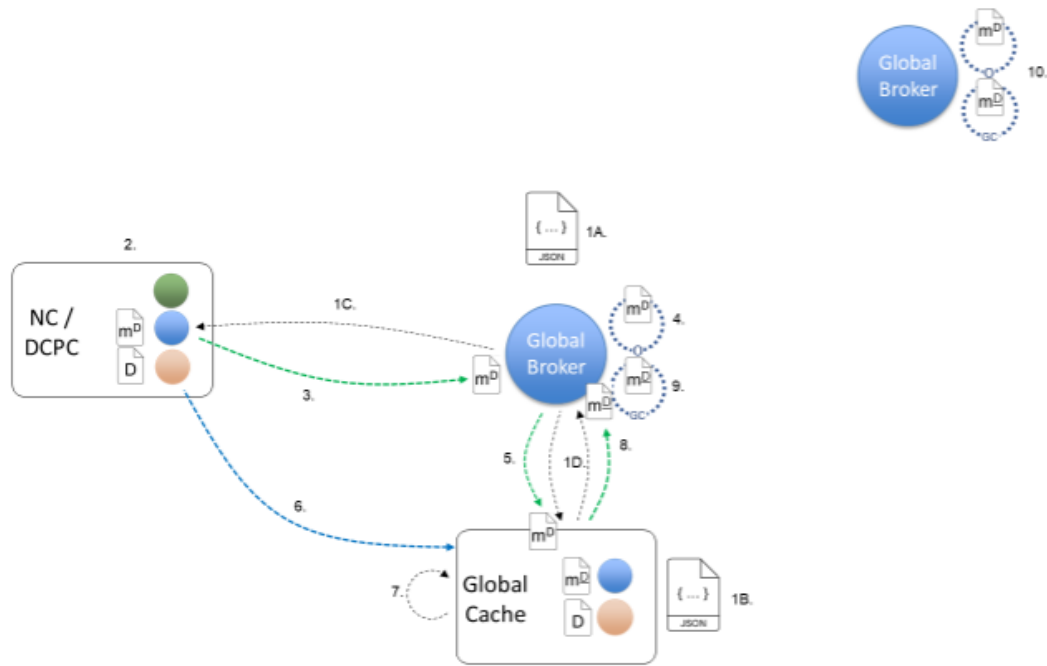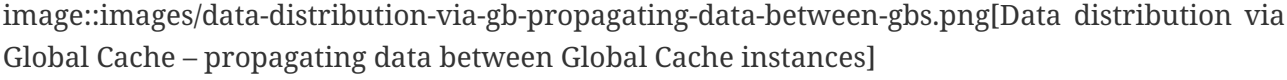
Figure 7. Data distribution via Global Cache – publishing "data availability" messages

1. A[Pre-condition] Bootstrap the Global Broker with a configuration that indicates (i) NC/DCPCs and Global Cache instances from which to re-publish messages, and (ii) connection details for at least two other Global Broker instances.

2. B[Pre-condition] Bootstrap the Global Cache with a configuration that indicates connection details for at least one Global Broker instance, the list of datasets that need to be cached, and the list of NC/DCPC for which this instance serves as a primary Global Cache.

3. C[Pre-condition] Global Broker determines the MQ broker topics to subscribe to and subscribes to them.

4. D[Pre-condition] Global Cache determines the MQ broker topics to subscribe to and subscribes to them.

5. As the dataset is produced Data Objects (**D**) are made available for download, and messages (**m^D**) are published to the topic for this dataset on the local MQ broker.

6. Message (**m^D**) indicating availability of new data are pushed to the subscriber.

7. Message (**m^D**) is not a duplicate, so Global Broker republishes it.

8. Message (**m^D**) indicating availability of new data are pushed to the subscriber.

9. Global Cache parses the message and downloads the Data Object (**D**).

10. Global Cache adds Data Object (**D**) to its data server and publishes a "data availability" message (**m^D**) that advertises data availability from the Global Cache instance.

11. Message (**m^D**) indicating availability of new data are pushed to the subscriber.

12. Message (**m^D**) is not a duplicate, so Global Broker republishes it – this time on the Global Cache channel.

13. Messages (mD, mD) are pushed to other Global Broker instances and re-published there.

Figure 8 illustrates how data propagates between Global Cache instances.

1. Data distribution via Global Cache – propagating data between Global Cache instances image::images/data-distribution-via-gb-propagating-data-between-gbs.png[Data distribution via Global Cache – propagating data between Global Cache instances]

2. A[Pre-condition] Bootstrap the Global Caches with a configuration that indicates connection details for at least one Global Broker instance and the list of datasets that need to be cached.

3. B[Pre-condition] Global Cache determines the MQ broker topics to subscribe to and subscribes to them.

4. C[Pre-condition] NC/DCPC has published Data Object (**D**) and messages (**m^D**) advertising availability of that Data Object.

5. D[Pre-condition] Global Cache #1 has downloaded Data Object (**D**) and published "data availability" message (**m^D**).

6. E[Pre-condition] Global Broker has re-published "data availability" messages from NC/DCPC and Global Cache instance #1 (mD, mD).

7. Global Cache #2 subscribes to both channels on the Global Broker. Both messages advertising Data Object (mD, mD) are pushed to GC#2. Message (**m^D**) arrives first, so message (**m^D**) is discarded as a duplicate.

8. Global Cache #2 parses the message and downloads the Data Object (**D**) from GC#1.

9. GC#2 adds Data Object (**D**) to its data server and publishes a "data availability" message (m2) that advertises data availability from GC#2.

10. Global Broker also subscribes to GC#2 (details not shown for brevity). New "data availability" message (m2) is pushed to the Global Broker and added to the "Global Cache" channel. It is not a duplicate of message (**m^D**) because it refers to a different download location – GC#2 not GC#1. Both messages (mD, m2) are published to the same topic on the Global Cache channel, so subscribers to that topic will receive multiple messages.

11. Messages (mD, mD, m2) are pushed to other Global Broker instances and re-published there.

4.4. Populating the Global Catalogue

Here we add the Global Discovery Catalogue service.

Figure 9 illustrates how discovery metadata records propagate from the NC or DCPC via a Global Cache instance to the Global Discovery Catalogue.
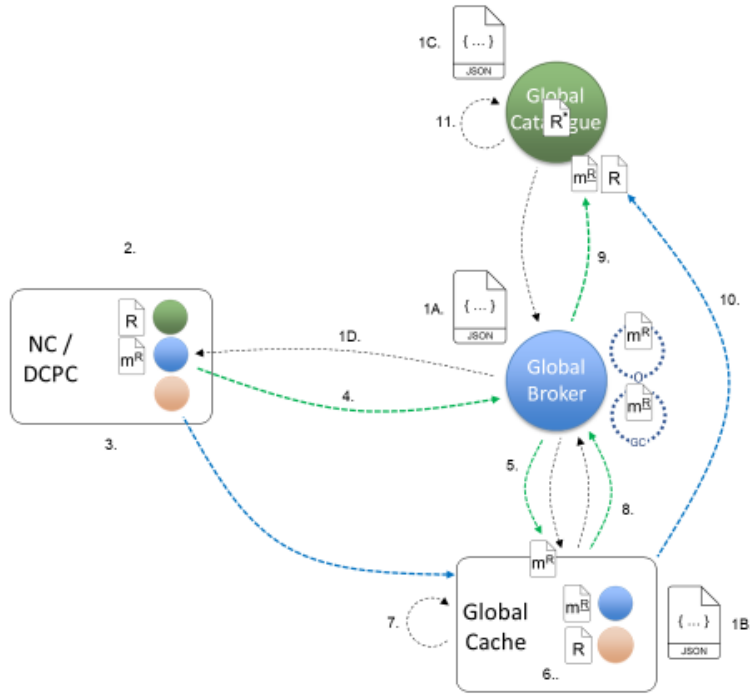
Figure 8. Populating the Global Catalogue with discovery metadata records

1. A[Pre-condition] Bootstrap the Global Broker with a configuration that indicates (i) NC/DCPCs and Global Cache instances from which to re-publish messages, and (ii) connection details for at least two other Global Broker instances.

2. B[Pre-condition] Bootstrap the Global Cache with a configuration that indicates connection details for at least one Global Broker instance, the list of datasets that need to be cached, and the list of NC/DCPC for which this instance serves as a primary Global Cache.

3. C[Pre-condition] Bootstrap the Global Catalogue with a configuration that indicates connection details for at least one Global Broker instance.

4. D[Pre-condition] Global Broker, Global Cache, and Global Catalogue determine MQ broker topics to subscribe to and subscribes to them. Global Catalogue subscribes only to the "Global Cache" channel.

5. NC/DCPC publishes a discovery metadata record ® to their local catalogue describing the dataset they are sharing.

6. NC/DCPC advertises the availability of this metadata record by publishing a message ($\mathbf{m^R}$) to the 'metadata' topic on the local MQ broker.

7. Message ($\mathbf{m^R}$) indicating the availability of new metadata is pushed to the Global Broker. The Global Broker republishes the message on the metadata topic of the "Originating Centre" channel.

8. Message ($\mathbf{m^R}$) indicating the availability of new metadata is pushed to the Global Cache.

9. Global Cache parses the message and downloads the discovery metadata record ®.

10. Global Cache adds discovery metadata record ® to its data server and publishes a message ($\mathbf{m^R}$) that advertises availability of the record from the Global Cache instance.

11. Message ($\mathbf{m^R}$) indicating availability of new metadata at the Global Cache is pushed to the Global Broker. The Global Broker republishes the message on the metadata topic of the "Global Cache"

channel.

12. Message ($\mathbf{m}^R$) indicating the availability of new metadata at the Global Cache is pushed to the Global Catalogue.

13. Global Catalogue parses the message ($\mathbf{m}^R$) indicating the availability of new metadata, asserts that this is not a duplicate message, and downloads the discovery metadata record ® from the Global Cache. It does not matter which Global Cache instance the record is downloaded from – a Global Cache will only publish a message ($\mathbf{m}^R$) when they hold a copy of record ®.

14. Global Catalogue creates a copy of the original record (new record is R*) and adds links indicating where Data Consumers can subscribe for updates to this dataset at each Global Broker instance. Global Catalogue adds record R* to the catalogue.

4.5. Data distribution including the Global Services

Here we bring all the components together to show how data is distributed from NCs or DCPCs to Data Consumers in WIS2. See Figure 10.
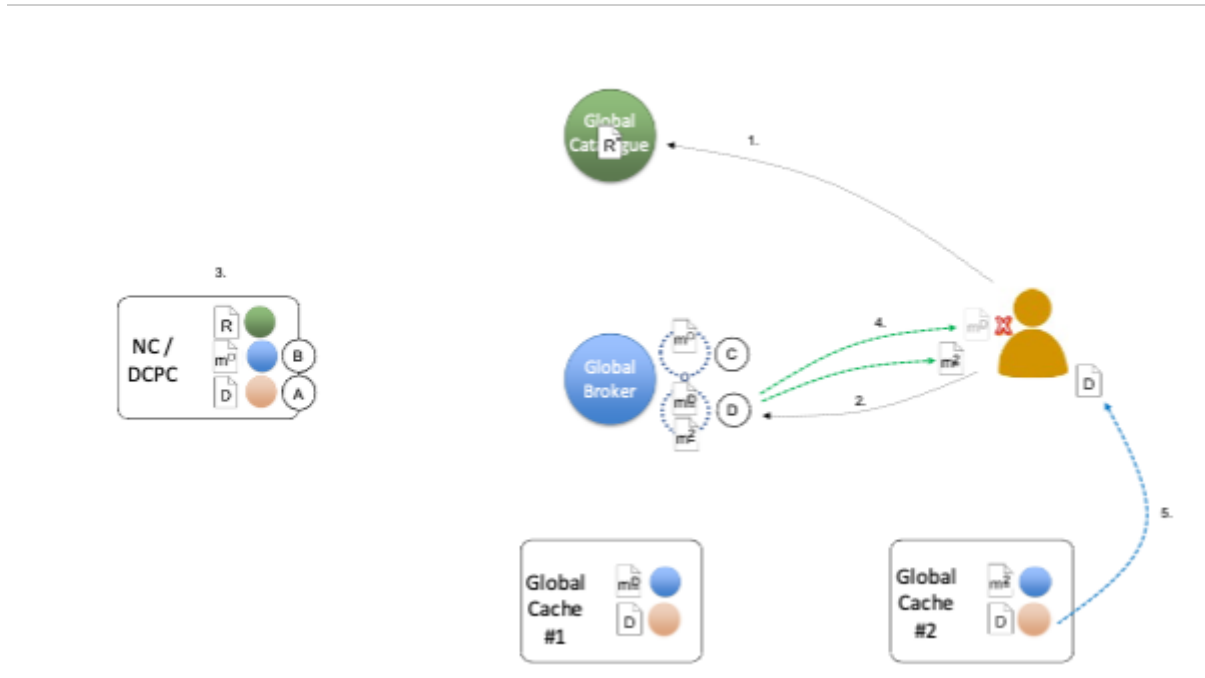
Figure 9. Data distribution with Global Services

Pre-conditions:

- WIS2 Global Service Components are correctly configured and operating.

- NC/DCPC has published a discovery metadata record ® to their local catalogue describing the dataset they are sharing. This record has propagated to the Global Catalogue, where an amended metadata record (R*) is published. Record (R*) includes additional actionable links to Global Broker instances where Data Consumers can subscribe for updates to the dataset.

  1. Data Consumer searches/browses the Global Catalogue and discovers a dataset that they need. Discovery metadata record (R*) provides links indicating how and where to download the data set and/or subscribe to updates.

     1. Download from NC/DCPC.

2. Subscribe to updates from NC/DCPC local MQ broker.

3. Subscribe to updates from Global Broker's "Originating Centre" channel.

4. Subscribe to updates from Global Broker's "Global Cache" channel.

2. Data Consumer wants to download Data Objects from the Global Cache, so decides to subscribe to dataset updates from the Global Broker's "Global Cache" channel.

3. [Data Object publication and propagation to Global Cache instances]

- NC/DCPC has published Data Object (**D**) as part of their dataset. Data Object (**D**) is propagated to Global Cache #1 and Global Cache #2.

- NC/DCPC has published a message ($\mathbf{m}^{\mathbf{D}}$) advertising availability of Data Object (**D**).

- GC#1 has published a message ($\mathbf{m}^{\mathbf{D}}$) advertising availability of Data Object (**D**).

- #2 has published a message (m2) advertising availability of Data Object (**D**).

- Global Broker has republished message ($\mathbf{m}^{\mathbf{D}}$) on its "Originating Centre" channel, and messages (mD, m2) on its "Global Cache" channel.

1. Messages (mD, m2) are pushed to the subscriber. Message ($\mathbf{m}^{\mathbf{D}}$) arrives second, so it is discarded.

2. Data Consumer parses the message (m2) to extract the URL for the new Data Object (**D**) hosted on GC#2 and downloads the Data Object.

Chapter 5. Real-time data exchange

5.1. What's wrong with the GTS?

The Global Telecommunication System (GTS) has been the mainstay of real-time data exchange within the WMO community since the 1970s. It has evolved over during that time, but its architecture impedes the free flow of information.

Being "on the GTS" requires operation of a message switching system (MSS) and update of the Routing Tables that control the flow of data between MSS. Once data is published on the GTS, anyone with a MSS can use the data they receive – or forward that data to other MSS. Each hop from MSS to MSS introduces delays and introduces a risk of data loss. Data publishers must trust recipients to respect terms and conditions in their data licenses. More fundamentally, updates to Routing Tables and MSS requires manual intervention – a model that simply will not scale to support the ever-growing community of consumers that need weather, water, and climate data.

5.2. Real-time data exchange with WIS2

WIS2 addresses the problems with the GTS as follows:

1. Controlled access to data: Web services are used to provide access to data in WIS2. A data publisher may choose to implement controlled access to their data by securing their Web-service [6].

2. Point-to-point data distribution: WIS2 uses the Internet to transport data directly between data suppliers and data consumers [7]. The Internet provides efficient and robust [8] transport of data throughout the globe – with routing managed at the network infrastructure level. WIS2 no longer needs MSS to manually route data via intermediate "hops".

3. Scaling to support many data consumers: WIS2 uses notification messages to alert consumers about the availability of new data. "Publish-subscribe" (Pub/Sub) messaging makes it simple to distribute these notifications in real-time to many consumers. A message broker allows an authorized agent (person, software system etc.) to create and publish messages to "topics". The message broker then manages requests from consumers to subscribe to those topics – without human intervention. All messages published to a topic are immediately sent to all subscribers. The publisher only needs to know about the topic – they don't need know who the message broker forwards their message to. Upon receipt of a notification message a data consumer will decide whether to download the data from the Web-service designated in the message. Consequently, WIS2 no longer needs Routing Tables.

WIS2 uses the OASIS MQTT messaging protocol, versions 3.1.1 and 5.0.

Several open-source message brokers are available that support MQTT. These include Mosquitto and RabbitMQ.

5.3. Notification message format and structure

WIS2 notification messages must be understood by all participants of WIS. Consequently, these

messages must comply with a standard format and structure.

WIS2 notification messages extend the GeoJSON format (RFC 7946), requiring the following elements to be included in each message:

1. A globally unique identifier (GUID) for the notification message
2. Geometry (point or polygon) of the associated data object (may be null)
3. Publication time of the notification message
4. Date-time of the associated data object
5. Topic hierarchy and unique identifier for the associated data object
6. Actionable links that can be followed to access the associated data object

Details of the WIS2 notification message format are available in WIS2 messages

5.4. Standard topic hierarchy

A 'Pub/Sub' message broker uses "topics" to organise messages. Messages are published to a specific topic, and consumers subscribe to the topics they are interested in so that the message broker sends them all relevant messages.

To simplify things for all Data Publishers and Data Consumers, a standard topic structure is used throughout WIS2: the "WIS2 Topic Hierarchy".

The standard topic structure enables notification messages to be grouped according to Member (e.g., country or territory), originating WIS Centre, and data type. It embeds the classification scheme of Earth System domains used Annex a of the WMO Unified Data Policy (Resolution 1, Cg-Ext (2021)).

The top-level of the standard topic structure "channel" is used to distinguish between notification messages from WIS Centres (channel = "origin") and notification messages from Global Cache services that host cached copies of data from the originating WIS Centre (channel = "cache"). This makes it simple for Data Consumers to comply with clause 3.2.13 from the Manual on WIS, Vol II:

```
3.2.13 Data consumers may access data from Global Caches or directly from a WIS node.
Data consumers SHOULD access real-time and near real-time core data via Global Caches.
```

There is a one-to-one mapping [9] between the theme (i.e., keyword) used to describe a Dataset in discovery metadata and the topic used to publish notifications about that Dataset.

Details of the WIS2 Topic Hierarchy are available in WIS2 topic hierarchy.

Note: the WIS2 Topic Hierarchy, particularly Earth System-Domain-Subcategory and below, is expected to evolve during the WIS2 pilot phase.

[6] Data accessible via a WIS2 Global Cache will be "Core data" (as defined in the WMO Unified Data Policy) – consequently it must be provided without restriction on use.

[7] A WIS2 Global Cache may be used ensure efficient and highly available access to data. See 8.4 Global Cache for more information.

[8] Even if generally robust, the Internet, as opposed to private networks, such as dedicated lines or the RMDCN, has no end-to-end Service Level Agreement (SLA). Robustness is based factual experience and not contractual SLA.

[9] Excluding the top two levels ("channel" and "version") – these are not considered part of the data classification scheme in the WMO Core Metadata Profile because they relate to data distribution mechanisms.

Chapter 6. Discovery metadata

Discovery metadata provides a description of a dataset sufficient for someone to determine whether that dataset meets their needs and includes links to where one may access more detailed information about the dataset or access the dataset itself.

Discovery metadata is an essential component of WIS2.

Data Publishers must provide a discovery metadata record for each dataset they make available within WIS. These discovery metadata records are aggregated within the Global Discovery Catalogue providing Data Consumers the means to search and browse every dataset within WIS.

Discovery metadata provided in WIS2 must conform to the WMO Core Metadata Profile version 2 (WCMP2), a profile of the OGC API - Records standard.

The WCMP2 specification, including usage examples, is published WCMP2.

TODO: Add link to the WCMP2 specification

Chapter 7. WIS metrics

The availability of data and performance of system components within WIS2 are actively monitored by GISCs and the Global Monitor service to ensure proactive response to incidents and effective capacity planning for future operations.

WIS2 requires that metrics are provided using OpenMetrics – the de-facto standard [10] for transmitting cloud-native metrics at scale. Widely adopted, many commercial and open-source software components already come preconfigured to provide performance metrics using the OpenMetrics standard. Tools such as Prometheus and Grafana provide aggregation and visualisation of metrics provided in this form, making it simple to generate performance insights. The OpenMetrics standard can be found at openmetrics.io [11].

The roles of centres with respect to the monitoring as well as the content and structure of metrics specific to WIS2 (e.g., data availability) are described below.

7.1. Roles

7.1.1. Global Monitoring (GM)

The Global Monitoring (Centres) are the entry points for users and provide the monitoring results. The main task of the Global Monitoring is to regularly query the provided metrics from the relevant WIS2 entities, aggregate and process the data and then provide the results to the end user in a suitable presentation.

7.1.2. Global Services

The WIS2 Global Services, namely the Global Broker, Global Cache, Global Discovery Catalogue and so on expose monitoring metrics on their respective service to the Global Monitoring. See [wis2-metric-hierarchy] for further details.

7.1.3. WIS2 Nodes

There is no requirement on WIS2 Nodes to provide monitoring metrics. However their WIS2 interfaces may be queried remotely by Global Services, which in turn can provide metrics on the availability of WIS2 Nodes.

7.1.4. Sensor Centres and WMO programs

WMO Programs can define monitoring requirements in order to be able to verify if the program's requirements are met. To do so, they need to define the relevant metrics. Metrics not directly related to the core services of WIS2 can be exported by so called Sensor Centres. Sensor Centres can read data or query services defined by the relevant WMO Program and provide the results in OpenMetrics format.

7.2. WIS2 Metric Hierarchy

7.2.1. Conventions for metric naming

Metrics for the WIS2 monitoring should follow the naming convention:

```
wmo_<program>_<name>
```

where program is the name of the responsible WMO Program and name is the name of the metric. Examples for WIS2 metrics can look like

```
wmo_wis2_gc_received
```

```
wmo_wis2_gb_messages_invalid
```

The full set of the WIS2 monitoring metrics is given in WMO: WIS2 Metric Hierarchy [12]

7.2.2. Metrics for Global Brokers

https://raw.githubusercontent.com/wmo-im/wis2-metric-hierarchy/main/metric-hierarchy/gb.csv

See [12] for further details.

7.2.3. Metrics for Global Caches

https://raw.githubusercontent.com/wmo-im/wis2-metric-hierarchy/main/metric-hierarchy/gc.csv

See [12] for further details.

7.2.4. Metrics for Global Discovery Catalogues

https://raw.githubusercontent.com/wmo-im/wis2-metric-hierarchy/main/metric-hierarchy/gdc.csv

See [12] for further details.

7.2.5. Other metrics

https://raw.githubusercontent.com/wmo-im/wis2-metric-hierarchy/main/metric-hierarchy/sensor.csv

See [12] for further details.

[10] OpenMetrics is proposed as a draft standard within IETF.

[11] https://openmetrics.io

[12] https://github.com/wmo-im/wis2-metric-hierarchy

Chapter 8. Implementation and operation of WIS2 Node

8.1. Practices and procedures

8.1.1. Registration and decommissioning of a WIS node

This section describes the process used to register or remove a WIS node within WIS2. During the initial part of the WIS2 pilot phase, a Member simply needs to notify the WMO Secretariat and primary GISC of the intent to register a new WIS node. The Secretariat and GISC will then assist in the registration. More formal procedures will be developed as the number of WIS nodes increases.

TODO: To be completed

8.1.2. Registration and removal of a Dataset

This section describes the process used to register a Dataset so that it may be discovered and shared within WIS2. In cases where a WIS Centre no longer wishes to share a Dataset via WIS, it must be removed as per the procedure described here.

TODO: to be completed

8.1.3. Connecting with Global Services

This section describes the process by which a WIS node is registered with one or more Global Broker and Global Cache components.

TODO: to be completed

8.2. Publishing data, discovery metadata, and notification messages

8.2.1. Discovery metadata

Discovery metadata shall be encoded according to the WMO Core Metadata Profile version 2 (WCMP2). See WCMP2

Discovery metadata may be published one of two ways. The simplest method is to encode the discovery metadata record as a file and publish it to an HTTP server. The URL of this file is included in the notification message advertising the availability of new metadata. Alternatively, a Data Publisher may choose to host a local catalogue themselves, enabling them to share discovery metadata records through an API (e.g., OGC API Records). In this case, the URL used in the notification message will refer to the API endpoint identifying the specific discovery metadata record (e.g., an item as part of their discovery metadata catalogue/collection).

These discovery metadata records are then propagated through the Global Service components into

to the Global Discovery Catalogue where Data Consumers can search and browse.

Discovery metadata must be published before data sharing through WIS begins.

TODO: to be completed

8.2.2. Notification messages

There is no requirement for an NC/DCPC to publish "data availability" notification messages relating to infrequently changing Datasets, such as a data archive, especially where the user community have no requirement to be instantly updated about changes to a Dataset (e.g., the addition of new records into a climate observation archive). Data Publishers should note that without providing notification messages their data will not be copied into the Global Cache. However, since the Global Cache only holds real-time (or near real-time) Datasets, this is not a concern for Data Publishers with infrequently changing Datasets.

TODO: to be completed

8.2.3. Data

WIS2 provides the "plumbing" for data sharing within the WMO community, but it defines neither which data to share, nor how that data should be encoded. WIS Centres need to evaluate WMO Programme requirements and the WMO Unified Data Policy to determine which Datasets should be made available through WIS.

WMO Technical Regulations may require that data is encoded in particular formats. For example: synoptic observations should be encoded in BUFR. The Manual on Codes (WMO No. 306) provides details of data formats formally approved for use in WMO.

However, Technical Regulations don't cover all data sharing requirements. In such cases, Data Publishers should select data formats that are widely adopted and understood in their target user community.

WIS2 does not require the use of specific file-naming conventions. Where communities commonly use file-naming conventions (e.g., with embedded metadata), Data Publishers should ensure that adequate documentation is provided to users. Data Publishers cannot assume that users will understand (or respect) their file-naming rules – many Data Consumers will simply treat the filename as an opaque string.

Data publishers also have choices about how they publish data.

As a minimum, Data Publishers may publish data objects (e.g., the atomic bits of data that comprise a Dataset) as files using a Web server (HTTP protocol) or FTP server (FTP protocol), using secure communications (e.g., HTTPS/SFTP). As each data object is published, a notification message should also be published to a topic in a message broker (see 4.3 Notification message format and structure, and WIS2 messages 4.4 Standard topic hierarchy).

A Dataset (for example, a collection of climate model runs) may comprise thousands or more files. A Data Publisher may choose to provide users with a mechanism to browse through the set of files, enabling them to identify those which are relevant to them. Examples of such mechanisms include:

- Web Accessible Folders (WAF) – a Web-based folder structure listing the data object files by name .

- Spatio-Temporal Asset Catalog (STAC) – a common language based on GeoJSON to describe geospatial data files so that it can be easily indexed, discovered, and accessed. Freely available, open-source tools present STAC records (one for each data object file) through a Web-based, browse-able user interface.

WAFs and STAC are provided to illustrate options. There is no requirement for a Data Publisher to provide any such browse-able user interface to their data.

Increasingly, interactive Web APIs are being used to provide access to datasets. Although requiring a little more sophistication to implement, a Web API provides significant advantages:

- Data Consumers can select and download only the parts of a dataset that they need – providing them will a smaller dataset subset to work with and reducing the burden on the Data Publisher's network infrastructure.

- Data Consumers are insulated from the complexities of how a Data Publisher chooses to persist their data. The Web API can provide access to Datasets in a way that is easy for users to understand.

- A Web API may allow Data Consumers to download data in their preferred file format or encoding.

WIS-TECHSPEC-2 states:

```
4.3.3 URLs provided for accessing core data, as defined in WMO Unified Data Policy
(Res. 1), and discovery metadata shall be directly resolvable, i.e., data or discovery
metadata can be downloaded simply by resolving the given URL without further action,
such as populating elements of an API, is required.
```

When using a Web API to publish "core" data, the URL included in the data availability notification message must be directly resolvable, i.e., the Data Consumer must not be required to complete any additional fields in the API request. This can be achieved by identifying the data object in the URL. A Data Consumer or a Global Cache instance can simply resolve the URL to download the data object regardless of the manner in which it is made available.

WIS2 isn't yet mature enough to prescribe the use of particular Web APIs. Instead, WIS2 seeks to leverage the experience of Data Publishers who have been using Web APIs to serve their communities.

First, interactive Web APIs should be self-describing. A Data Consumer should not need to know, a priori, how to make requests from a Web API. They should be able to discover this information from the Web API endpoint itself – even if this is just a link to a documentation page they need to read.

Second, we recommend that Web APIs are compliant with OpenAPI version 3 or later. OpenAPI provides a standardised mechanism to describe the API. Effectively, OpenAPI provides metadata that describes the Web API endpoint. Tooling(free, commercial, etc.) is widely available that can

read this metadata and automatically generate client applications to query the Web API.

Third, the Open Geospatial Consortium (OGC) have developed a suite of APIs (called "OGC APIs") that are designed specifically to provide APIs for geospatial data workflows (discovery, vizualisation, access, processing/exploitation) – all of which build on OpenAPI v3. Among these, OGC API – Environmental Data Retrieval (EDR), OGC API – Features, and OGC API - Coverages are considered particularly useful. Because these are open standards, there is an ever-growing suite of software implementations (both free and commercial) that support them. We recommend that Data Publishers assess these open-standard API specifications to determine their suitability to for publishing their Datasets using APIs.

Finally, we're increasingly concerned with providing access to very large Datasets. The OGC has published a series of informative blogs on the subject of cloud-native geospatial data sharing. These are listed among in section 11.4.2 Informative References.

TODO: to be completed

8.2.3.1. Publication and topic selection

When publishing a dataset, a data publisher selects a given topic according to the WIS Topic Hierarchy. Given the multidisciplinary nature of some data, a data publisher must select a single topic for publication purposes, and always uses WCMP2 discovery metadata to provide a fulsome description of their dataset and its relevance to additional disciplines.

8.3. Performance management

8.3.1. Service levels and performance indicators

This section describes the minimum performance criteria for operation of a WIS node.

TODO: to be completed

8.3.2. Provision of system performance metrics

This section describes how a WIS node should provide metrics to the Global Monitor service and its primary GISC.

TODO: to be completed

8.4. WIS Node reference implementation: wis2box

Members may use whichever software components they consider most appropriate to comply with the WIS2 Technical Regulations.

To assist Members participate in WIS2, a freely available, open-source Reference Implementation has been developed: "WIS2 in a box" (referred to as wis2box). It builds on mature and robust free and open-source software components that are widely adopted for operational use.

wis2box provides functionality required for both Data Publisher and Data Consumer roles. It

provides the following technical functions:

- Real-time or archive data and metadata publishing to WIS2 (Publish), including available data transformation and processing pipelines

- MQTT Message Broker and notification message publication (Subscribe)

- Object storage server providing raw data access (Download)

- OGC API server, providing dynamic APIs and Web services for discovery, access, visualization and processing functionality (APIs)

- Discovery metadata curation / editing tools

- Notification subscription and real-time download of data upon receipt of notifications.

- Provision of system performance and data availability metrics

- The modular design of wis2box makes it simple to extend to meet additional requirements or integrate with existing data management systems.

- wis2box already provides a useful set of functionality and will continue to evolve and develop throughout the WIS2 pilot phase and beyond.

Documentation is published in wis2box documentation.

The project in hosted in GitHub: https://github.com/wmo-im/wis2box

Chapter 9. Implementation and operation of a Global Service

9.1. Procedure for registration of a new Global Service

During the initial stages of the WIS2 pilot phase an informal process is used to register interest in operating a Global Service. Candidate WIS Centres should inform their WIS Focal Point and contact the WMO Secretariat to discuss their offer to provide a Global Service.

9.2. Performance management and monitoring of a Global Service

9.2.1. Service levels, performance indicators, and fair-usage policies

TODO: to be completed

9.2.2. Provision of metrics

TODO: to be completed

9.3. Global Broker

9.3.1. Technical considerations

- There will be multiple Global Broker instances to ensure highly available, low latency global provision of messages within WIS.

- A Global Broker instance subscribes to messages from NC/DCPCs and other Global Brokers

- A Global Broker instance will subscribe to messages from a subset of NC/DCPCs and republish them.

- At least one Global Broker will subscribe to messages from every NC/DCPC.

- For full global coverage, a Global Broker instance will subscribe to messages from other Global Broker instances and republish them.

- A Global Broker instance will republish a message only once – noting that a particular message may be received multiple times (e.g., from different sources). Discarding duplicate messages is referred to as "anti-loop".

- It is not required that a Global Broker instance republishes messages from all other Global Brokers (e.g., establishing 'fully meshed' connection). However, it is essential that messages propagate through WIS efficiently and effectively, from originating NC/DCPC to Data Consumers in all Regions. Consequently, it is recommended that topological distance between every Global Broker shall not exceed 3 "hops" (i.e., a message received at a Global Broker shall be republished by no more than 3 other Global Brokers on its route from the originating NC/DCPC). Connectivity between Global Brokers will be recommended by Experts from INFCOM/SC-IMT.

- Global Brokers use distinct "channels" to keep messages from originating NC/DCPC separate from messages originating from Global Cache instances. This is implemented in using the top-level ("channel") of the topic structure (see WIS2 messages

- Standard topic hierarchy

  ◦ A Global Broker will validate notification messages against the standard format (see Notification message format and structure), discarding non-compliant messages and raising an alert.

  ◦ A Global Broker is built around two software components:

    ▪ An off the shelf broker implementing both MQTT 3.1.1 and MQTT 5.0 in a highly-available setup (cluster)Tools such as EMQX, HiveMQ, VerneMQ are compliant with these requirements.

    ▪ Additional features (anti-loop, message format compliance,…) are required. An open source implementation will be made available during the pilot phase.

TODO: to be completed

9.3.2. Practices and procedures

The following procedures will be described here once validated through testing during the WIS2 pilot phase:

- Assigning a Global Broker to a NC or DCPC
- Alerting originating NC or DCPC about malformed or non-compliant messages

TODO: to be completed

9.4. Global Cache

9.4.1. Technical considerations

- The Global Cache will contain copies of real-time and near real-time data designated as "core" within the WMO Unified Data Policy (Resolution 1).

- During the initial stages of the WIS2 pilot phase Global Cache instances will provide open access to their cached content. Access control mechanisms may be added later.

- A Global Cache instance will host data objects copied from NC/DCPCs. These are persisted as files.

- A Global Cache instance will publish notification messages advertising availability of the data objects it holds. The notification messages will follow the standard structure (see 4.3 Notification message format and structure).

- A Global Cache instance will use the standard topic structure in their local message brokers (see WIS2 messages 4.4 Standard topic hierarchy).

- There will be multiple Global Cache instances to ensure highly available, low latency global provision of real-time and near real-time "core" data within WIS.

- Global Cache instances may attempt to download cacheable data objects from all originating centres with "cacheable" content. A Global Cache instance will also download data objects from other instances. This ensures the instance has full global coverage, mitigating where direct download from an originating centre is not possible.

- For full global coverage, a Global Cache instance will download Data Objects and discovery metadata records from other instances.

- Global Cache instance will operate independently of other Global Cache instances. Each Global Cache instance will hold a full copy of the cache – albeit that there may be small differences between Global Cache instances as "data availability" notification messages propagate through WIS to each Global Cache in turn. There is no formal 'synchronisation' between Global Cache instances.

- A Global Cache will store a full set of discovery metadata records. This is not an additional metadata catalogue that Data Consumers can search and browse – it provides a complete set of discovery metadata records to support populating a Global Discovery Catalogue instance.

- A Global Cache is designed to support real-time distribution of content. Data Consumers access data objects from a Global Cache instance by resolving the URL in a "data availability" notification message and downloading the file.

- There is no requirement for a Global Cache to provide a "browse-able" interface to the files in its repository allowing Data Consumers to discover what content is available. However, a Global Cache may choose to provide such a capability (e.g., implemented as a "Web Accessible Folder", or WAF) along with adequate documentation for Data Consumers to understand how the capability works.

TODO: to be completed

9.4.2. Practices and procedures

The following procedures will be described here once validated through testing during the WIS2 pilot phase: * Assigning a Global Cache to a NC or DCPC * Lifecycle management of discovery metadata records stored in the Global Cache.

TODO: to be completed

9.5. Global Discovery Catalogue

9.5.1. Technical considerations

- The Global Discovery Catalogue provides Data Consumers with a mechanism to discover the Datasets of interest, as well as, how to interact with and find out more information about those Datasets.

- The Global Discovery Catalogue implements the OGC API – Records – Part 1: Core standard.

- The Global Discovery Catalogue advertises the availability of Datasets and how/where to access them or subscribe to updates, it does not advertise the availability of individual Data Objects that comprise a dataset (i.e., data files).

- A single Global Discovery Catalogue instance is sufficient for WIS2.

- Multiple Global Discovery Catalogue instances may be deployed for resilience.

- Global Discovery Catalogue instances operate independently of each other – each Global Discovery Catalogue instance will hold all discovery metadata records. There is no need to synchronise between Global Discovery Catalogue instances.

- A Global Discovery Catalogue is populated with discovery metadata records from a Global Cache instance – receiving messages about availability of discovery metadata records via a Global Broker.

- A Global Discovery Catalogue should connect to more than one Global Broker instance to ensure that no messages are lost in the event of a Global Broker failure. A Global Discovery Catalogue instance shall discard duplicate messages as needed.

- A Global Discovery Catalogue will validate discovery metadata records against the WMO Core Metadata Profile version 2 (WCMP2), discarding non-compliant records and raising an alert.

- A Global Discovery Catalogue will update discovery metadata records it receives to add links for subscription URLs at Global Broker instances.

- A Global Discovery Catalogue should applying faceting capability as specified in the cataloguing considerations of the WCMP2 specification as defined in OGC API - Records.

- A Global Discovery Catalogue shall provide human-readable Web pages with embedded markup using the schema.org vocabulary, thereby enabling search engines to crawl and index the content of the Global Discovery Catalogue. Consequently, Data Consumers should also be able to discover WIS content via third party search engines.

- A Global Discovery Catalogue shall periodically assess the discovery metadata provided by NCs and DCPCs against a set of key performance indicators (KPIs) in support of continuous improvement. Suggestions for improvement are shared with the originating NC or DCPC and their primary GISC.

TODO: to be completed

9.5.2. Practices and procedures

The following procedures will be described here once validated through testing during the WIS2 pilot phase:

- Alerting originating NC or DCPC about malformed or non-compliant discovery metadata records

- Providing feedback to NC and DCPC about how to improve their discovery metadata

- Removing discovery metadata for a Dataset on request

- 'Bootstrapping' a Global Discovery Catalogue instance from the Global Cache

TODO: to be completed

9.6. Global Monitor

9.6.1. Technical Considerations

- WIS2 standardises how system performance and data availability metrics are published from

WIS nodes and Global Services.

- The Global Monitor will collect metrics as defined in the OpenMetrics standard.

- The Global Monitor will monitor the 'health' (i.e., performance) of components at NC/DCPC as well as Global Service instances.

- The Global Monitor will provide a Web-based 'dashboard' that displays the WIS2 system performance and data availability.

TODO: to be completed

9.6.2. Practices and procedures

Procedures pertinent to the Global Monitor will be described here once validated through testing during the WIS2 pilot phase.

TODO: to be completed

Chapter 10. Coordinating WIS

GISCs ensure the efficient and effective operation of WIS and WIS2 within their Area of Responsibility (AoR) and, in collaboration with the WMO Secretariat, globally. In WIS2 they cooperate with the Global Services and the WIS2 Nodes in their AoR. In the following the term WIS also refers to WIS2.

Guidance is provided here to help GISCs fulfil this vital obligation.

10.1. Supporting WIS centres in their area of responsibility

The primary function of a GISC is to support Members in their Area of Responsibility (AoR) participate effectively in WIS – both as Data Publishers and Data Consumers.

To provide effective support, it is recommended that GISCs operate a helpdesk or ticketing system to manage issues and requests raised from WIS Centres in their AoR.

TODO: to be completed

10.2. Performance management

Monitoring of WIS is intended to improve the exchange of information by WMO Programmes and to ensure that the functions of WIS centres minimize the cost of operating WIS. It is also essential for planning and scaling WIS components to meet changing user needs. It complements monitoring of quality of the information being exchanged, which is the responsibility of the Programmes sponsoring the information within WIS.

To ensure efficient and effective operation of WIS in their Area of Responsibility (AoR), GISCs are responsible for collecting metrics from Centres in their AoR.

A GISC may provide a publicly visible portal displaying performance status of WIS.

TODO: to be completed

10.3. Incident management

GISC are responsible for resolving incidents that impede the operation of WIS.

Appendix E of the Guide to the WMO Information System (WMO No. 1061) describes the key management activities relating to incident management and service continuity in IT systems.

Where the impact arising from an incident may affect WIS Centres outside their Area of Responsibility (AoR), a GISC must coordinate with WMO Secretariat and other GISCs.

Appendix F of the Guide to the WMO Information System (WMO No. 1061) describes the process for responding to an IT security (cybersecurity) incident.

GISCs collaborative provide a "GISC Watch" function wherein, according to a mutually agreed roster, a nominated GISC will actively monitor the global performance of the WIS and coordinate with WMO Secretariat and other GISCs to respond to any issues or incidents arising.

TODO: to be completed

Chapter 11. Information management

11.1. Introduction

11.1.1. Background

The efficient and effective provision of services relying on meteorological, climatological, hydrological and oceanographic information depends on a reliable information infrastructure. This infrastructure should be guided by community best practices and standards, including recommendations and requirements on sourcing, securing, managing, archiving, exchanging, and providing easy access to information. These terms and activities can be grouped under the term "information management" and this part of the Guide aims to provide high-level guidance on those activities. This is done by identifying and describing the fundamental principles of good information management and by highlighting the different stages of the information management lifecycle.

Note: The term "information" is used in a general sense and includes data and products.

11.1.2. Scope

High-level guidance on information management practices that apply in the context of information related to the Earth system is provided in this part of the Guide. Detailed technical information, such as specification of data formats or quality control and assurance methods, is provided in other parts of the Guide and in other WMO publications. These are referenced where applicable.

The principles of information management are described in Section 6.2. Section 6.3 describes the information management lifecycle through the identification of five focus areas. These are:

1. Planning, information creation and acquisition. Creation of information using internal and external data sources and the acquisition of information from various sources.

2. Representation and metadata. Standards to represent metadata, data and information are of primary importance to enable interoperability and long-term usability of the information.

3. Publication and exchange of information. The creation and publication of discovery metadata in a standardized format enabling users to discover, access and retrieve the information.

4. Usage and communication. Publication of guidance material on the use of published information, including on the limitations and suitability of the information and any licensing terms.

5. Storage, archival and disposal. Policies and procedures for business continuity and disaster recovery, as well as retention and disposal.

11.1.3. Intended audience

This guidance is primarily aimed at personnel within WMO centres, with responsibility for planning and undertaking the creation or acquisition, stewardship, exchange and provision of information related to the Earth system.

Specifically, the guidance has five main target audiences across the information lifecycle:

1. Information producers or creators (those who produce or acquire the information - they need to ensure the scientific quality of the underpinning information).

2. Information managers (those who manage information).

3. Information providers or publishers (those who publish the information - they are responsible for the provision of the information, and for ensuring that appropriate access is enabled, licensing agreements are in place, etc.).

4. Service providers (those who disseminate the information - they are responsible for ensuring information availability and maintaining capability for easy and secure access to the information).

5. Information consumers (those who utilize the information - they need to understand the restrictions, rights, responsibilities and limitations associated with the information together with the suitability for intended usage or purpose).

11.2. Principles of information management

Effective management of information is essential for WMO Centres to deliver operational services and information that is authoritative, seamless, secure and timely. The principles below underpin this management across the full information lifecycle and provide a framework for information management. The principles are independent of information type and are largely independent of technology, they are therefore expected to remain stable over time.

11.2.1. Principle 1: Information is a valued asset

- An information asset is information that has value. This value may be related to the cost of generating and collecting the information, a value associated with the immediate use or a value associated with the longer term preservation and subsequent reuse of the information.

- 10.2.1.2 This value should be recognizable and quantifiable and the asset should have an identifiable lifecycle. Risks associated with, and to, an information asset should also be identified. As such, information management must be considered an integral part of a WMO centre's responsibilities and needs to be adequately resourced over the full lifecycle of the information.

11.2.2. Principle 2: Information must be managed

- An information asset must be managed throughout its lifecycle, from creation to use to eventual disposal, in a way that makes it valuable, maximizes its benefits and reflects its value in time and its different uses.

- Information managers must consider the entire information lifecycle, from identifying needs and business cases to creating, quality assurance, maintenance, reuse, archiving, and disposal. Careful consideration must be given to disposal, ensuring that information is destroyed only when it has ceased to be useful for all categories of users.

- Professionally qualified and adequately skilled staff with clear roles and responsibilities should apply a sound custodianship framework concerning security, confidentiality and other statutory

requirements of different types of information.

11.2.3. Principle 3: Information must be fit for purpose

- Information should be developed and managed in accordance with its function and use for internal and external users.

- WMO Centres should regularly assess information to ensure that it is fit for its purpose and that processes, procedures, and documentation are adequate.

- Processes should be consistent with the general provisions and principles of quality management as described in the WMO Technical Regulations (WMO-No. 49).

11.2.4. Principle 4: Information must be standardized and interoperable

- Information must be stored and exchanged in standardized formats to ensure wide usability in the short and long term. It is essential for long-term archiving that information is stored in a form that can be understood and used after several decades.

- Standardization is essential for structured information such as dataset definitions and metadata to support interoperability.

- Interoperability is essential for users to utilize information through different systems and software. Open standards help ensure interoperability with their openness and wide adoption across various communities.

- Which standards to use depends on the user community and organizational policies. Interoperability requirements should be considered when selecting the standard for internal use and broader dissemination.

- The use of closed and proprietary standards is strongly discouraged.

11.2.5. Principle 5: Information must be well documented

- WMO centres should comprehensively document information processes, policies, and procedures to facilitate broad and long-term use.

- WMO centres should keep documentation up to date to ensure full traceability of processes along the information lifecycle, particularly for its creation.

- Previous versions of the documentation should be retained, versioned, archived and made readily available for future use. In addition, versions should be assigned a unique and persistent identifier for future unambiguous identification.

11.2.6. Principle 6: Information must be discoverable, accessible and retrievable

- Information should be easy to find through the Web, and for this purpose, the publisher should share discovery metadata with a catalogue service. The catalogue service should include a Web Application Programming Interface (API) to be used by other applications in order to offer user-tailored search portals.

- For information to be easily retrievable once discovered, it should be accessible using standard data exchange protocols.

11.2.7. Principle 7: Information should be reusable

- In order to maximize the economic benefits of an information asset it should be made as widely available and as accessible as possible.

- The WMO Unified Data Policy encourages the reuse of data and information through the open and unrestricted exchange of core WMO data. The WMO encourages the free and unrestricted exchange of information in all circumstances.10.2.7.3 The publisher should provide an explicit and well-defined license for each information item or dataset as part of the associated metadata.

- The Findable, Accessible, Interoperable and Reusable (FAIR) data principles promote open data with the ultimate goal of optimizing reuse of data. These principles should be followed where possible.

Note: Information on the FAIR data principles can be found at: FAIR Principles - GO FAIR [13]

11.2.8. Principle 8: Information management is subject to accountability and governance.

- Information management processes must be governed as the information moves through its lifecycle. All information must have a designated owner, steward, curator and custodian. These roles may be invested in the same person but should be clearly defined at the time of creation. A WMO centre with responsibility for managing information must ascertain:

- information management practices, procedures and protocols, including well-defined roles, responsibilities and restrictions on managing the information;

- definition and enforcement of appropriate retention policy, taking into account stakeholder needs and variations in value over the information lifecycle;

- licensing and defining and enforcing any access restrictions.

- The designated owner should have budget and decision-making authority about preservation and data usage, including passing ownership to another authority.

11.3. The information management lifecycle

11.3.1. Overview

All information should be subject to a well defined and documented lifecycle. The governance of this process is often referred to as the information management lifecycle and this process helps organizations manage information throughout its full lifecycle, from planning, creation and acquisition through usage and exchange to archival and disposal.

The following sections describe two overarching themes, governance and documentation, that apply to all stages of the information lifecycle and then provides high level guidance split into 5 aspects:

- Planning, creation and acquisition
- Representation and metadata

- Publication and exchange

- Usage and communication

- Storage, archival and disposal

Governance covers the rules that apply to managing information in a secure and transparent manner, documentation covers the act of recording the reasons for, and detail of, all operations in the information management process.

11.3.2. Overarching requirements

11.3.2.1. Governance

- Information management governance defines a set of organizational procedures, policies and processes for the management of information. This includes defining accountabilities and compliance mechanisms.

- Effective governance helps ensure that all aspects of the information management process are conducted in a rigorous, standardized and transparent manner and that the information are secure, accessible and usable.

- WMO centres should establish a board or leadership group to develop and regularly review such a governance structure and ensure compliance with its requirements.

11.3.2.2. Documentation

- Documentation describing the who, what, why, when, where and how various actions are undertaken in the management of information is required to ensure the traceability and integrity of the information and to ensure operations can continue if key staff leave.

- This documentation is required for all aspects of the information lifecycle and should be clear, well communicated, regularly updated, and easy to find. Guidance to the documentation should be provided to new staff taking on responsibilities for information management and be a key component of training.

11.3.3. Aspects of the information management lifecycle

11.3.3.1. Planning, information creation and acquisition

Before the creation or acquisition of new information a business case and information management plan should be developed, covering both the input information sources and any derived information. The plans should include:

- Why the information is required

- How it will be collected or created

- How it will be stored

- Whether it will be exchanged with other users and under what policy

- Where it should be submitted for long term archival

- Key roles and responsibilities associated with the management of the information

For externally sourced data the plans should include where the information has come from and what the licensing terms are.

Once information has been acquired it should be checked to ensure that the contents and format are as expected. This may be done using a compliance checker or validation service. Once these checks have been performed the information content should also undergo quality control checks using well documented procedures to identify any issues. A record of the checks should be kept and any issues detected should be documented and feedback to the originators. It is also important to subscribe to updates from originators so any issues identified externally can be taken into account.

Information created rather than acquired should undergo the same processes as the acquired information. The information created should undergo quality control and the resulting files checked against the specified format requirements. The results of the processes and checks should be documented.

To ensure traceability and reproducibility the information and documents at this, and subsequent stages, should be version controlled and clearly labelled with version information. Similarly, software, or computer code, used to generate or process information should be version controlled with the version information recorded in the documentation and metadata. Where possible, software should be maintained within a code repository.

11.3.3.2. Representation and metadata

The formats used to store and exchange information should be standardized to ensure its usability, both in the short and long term. It is essential that the information can be accessed many years after archival if required. To ensure this usability, the format and version information should be recorded in the metadata record for the information and should be included in the information where the format allows.

Information exchanged on the WMO Information System and between WMO centres is standardized through the use the formats specified in the WMO Manual on Codes (WMO-No. 306, Volume I.2) and the Manual on the WMO Information System (WMO-No. 1060). This includes the GRIB and BUFR formats for numerical weather prediction products and observational data and the WIS Core Metadata Profile for discovery, access and retrieval metadata. The format for the exchange of station and instrumental metadata, the WIGOS Metadata Data Representation, is also defined in the WMO Manual on Codes (WMO-No. 306, Volume I.3).

These formats have been developed within the WMO community to enable the efficient exchange of information between WMO centres and for the information to be interoperable between centres and systems. The formats, including detailed technical information, have also published openly through the WMO manuals, enabling use of the formats and information by other communities, promoting reuse of the information.

The WMO formats specified in the manuals are subject to strong governance processes, and changes to the formats can be traced through the versions of the manuals. The code tables and controlled vocabularies are also maintained in a code repository. To enable future reuse, the technical information, including detailed format specifications, should be archived alongside information for future access. This includes any controlled vocabulary, such as BUFR tables or WIGOS metadata code lists, associated with the format.

11.3.3.3. Publication and exchange of information

To maximize the benefits and return on investment in the acquisition and generation of information there needs to be a clear method as to how the information will be published, exchanged and accessed by users.

Information is published on the WMO Information System through the creation of discovery metadata records. These records are publicly searchable and retrievable via WMO cataloguing services, providing access to the records via the Web and via a Web Application Programming Interface (API). The metadata records should include information on how to access the described datasets and services (see also 10.3.3.3.4) and how to subscribe to receive updates and new data.

Guidance on the creation of these discovery metadata records is included in Part V of this Guide. Technical regulations are provided in the Manual on the WMO Information System (WMO-No. 1060). Before exchange and publication the metadata should be assessed using the WMO Core Metadata Profile Key Performance Indicators to ensure usable and high quality metadata in addition to metadata that conforms with the technical standard.

Note: Further information on the Key Performance indicators can be found on the WMO Community Website at https://community.wmo.int/activity-areas/wis/wis-metadata-kpis

The Web standards and protocols used should be adequately documented to enable users to find and retrieve the information. This should be possible both manually and automatically via machine-to-machine interfaces and should be standardized between centres.

Updates to the information exchanged on the WIS, including the publication of new information or the cessation of previously exchanged information, is published in the WMO Operational Newsletter.

Note: The newsletter is available from: https://community.wmo.int/news/operational-newsletter

11.3.3.4. Usage and communication

For information to have value it must inform users, aid knowledge discovery and have impact through informed decision making. Ensuring that the user can make effective use of the information is an important step in the information management lifecycle. This takes two forms:

1. Provision of suitable information within the discovery metadata (See 10.3.3.4), enabling users to discover and access the information and to assess whether it meets their requirements. This should include licensing information.

2. Provision of user guides and documentation on the suitability of the information for different uses, including any technical caveats or restrictions on the use of the information.

For common types of information the guides may be generic or link to standard documentation. Information on the observations available from the WMO Integrated Global Observing System is provided within the Manual and Guide to the WMO Integrated Global Observing system, WMO-No. 1160 and WMO-No. 1165 respectively. This includes information on the expected uses and quality of the data, either directly or through links within. Similarly, information on the data and products available through the Global Data Processing and Forecasting System is provided in the Manual on the Global Data Processing and Forecasting System (WMO-No. 485).

For non-standard and specialist products targeted user guides may be more appropriate. These should include a plain text summary for the non-technical user and should also be accessible and retrievable via a link within the discovery metadata. Any user guide should be in addition to the technical documentation described under Planning, information creation and acquisition (see 10.3.3.1).

Updates and the availability of new information should be announced and published via the WMO Operational Newsletter (see 10.3.3.4.6). Other communication methods may also be used but these should not be in place of the operational newsletter. It is also recommended to allow users to subscribe to receive updates directly.

The discovery metadata should include a valid point of contact, enabling users to provide feedback and ask questions about the information provided.

11.3.3.5. Storage, archival and disposal

The type of storage used should be appropriate to the type of information stored. Core information exchanged operationally should be stored and made available via high-availability and low latency media and services. For some operation critical information, such as hazard warnings, there is a requirement for the end-to-end global distribution of the information to be completed in two minutes. For other operational data there is a requirement for the global exchange to be completed in 15 minutes.

The storage requirements for non-operational services and information may be different but the guidance provided in this section applies equally. Further information on the performance requirements is provided within the WIS Technical Specifications listed in the Manual on the WMO Information System (WMO-No. 1060).

Backup policies and data recovery plans should be documented as part of the information management plan. These should be implemented either before or when the information is created or acquired and should include both the information and the associated metadata. The backup and recovery process should be routinely tested. Specific guidance on the expectations and requirements for WMO centres is provided under the operational guidance in Part VII of this Guide.

Business rules governing the access to and modification of the information should be clearly documented in the information management plan. This must include the clear specification of roles and responsibilities of those managing the information. Information on who can authorize the archival and disposal of the information and the processes for doing so should be included. The roles associated with an information resource are standardized as part of the WIS Core Metadata Profile, see Part V of this Guide for further information.

The archival and long-term preservation of an information resource should be identified and included in the information management plan. This may be at a national data centre and/or a WMO centre. The WMO centres are recommended for globally exchanged core data and include those centres contributing to the Global Atmosphere Watch, the Global Climate Observing System and the Marine Climate Data System (see Manual on Marine Meteorological Services, WMO-No. 558), as well as the WMO World Data Centres and those defined in the Manual on the WMO Information System (WMO-No. 1060) and those defined in the Manual on the Global Data Processing and Forecasting System (WMO-No. 485).

Earth system information, especially observational data, are often irreplaceable. Other information, whilst technically replaceable, is often costly to produce and therefore not easily replaceable. This includes output from numerical models and simulations. Before an information resource is marked for disposal careful consideration must be given to whether long term archival or disposal is more appropriate. This consideration must follow a clearly defined process documented in the information management plan.

When an information resource is marked for disposal the reasons for disposal, including the outcome of the consultation with stakeholders and users, must clearly be documented. The disposal must be authorized by the identified owner and custodian of the information. The information on the disposal must be included in the metadata associated with the information resource. The metadata must be retained for future reference.

11.4. Other considerations

11.4.1. Technology and technology migration

Information managers must be aware of the need to ensure that the technologies, hardware and software used do not become obsolete and must be aware of emerging data issues. This topic is discussed further in the WMO Guide to Emerging Data Issues (WMO-No. 1239).

11.4.2. Information security

Further information on information security and best practices can be found in the WMO Guide to Information Technology Security (WMO-No. 1115).

---

[13] https://go-fair.org

Chapter 12. Other considerations

12.1. Competencies of personnel

The Manual on WIS (WMO No. 1060), Vol II. Part I, 1.6, recommends that Members have access to staff who have the required competencies needed to operate their WIS Centres. This includes both generic Information Technology competencies and those specific to the operation of WIS.

WIS2 embraces use of open standards and widely adopted Web technologies. As such, there are many training and development resources available online. WIS-specific competencies are described in Appendix A: Competencies.

12.2. Audit and certification of WIS Centres

INFCOM recognizes that for WIS to remain fully functional regular reviews of each NC, DCPC, and GISC are required to assess ongoing compliance with the Manual on WIS. Recommended practices for this rolling review are provided in Appendix B: Recommended Practices for Audit and Certification of WIS Centres

12.3. Pilot phase deployment configuration

TODO: To be completed Once the WIS2 pilot phase is established, information will be added describing the location and configuration details required to connect to Global Services

Chapter 13. References

13.1. Normative

- WMO: Manual on WIS (WMO No. 1060), Vol II. WIS 2.0
- WMO: WMO Core Metadata Profile version 2 (WCMP2) [14]
- WMO: WIS2 Topic Hierarchy [15]
- WMO: WIS2 Notification Message format [16]
- WMO: WIS2 Metric Hierarchy [17]
- W3C: Architecture of the World Wide Web, Volume One [18]
- IETF: RFC 3339 Date and Time on the Internet: Timestamps [19]
- IETF: RFC 3986 Uniform Resource Identifier: Generic Syntax [20]
- IETF: RFC 7946 GeoJSON format [21]
- OGC: OGC API-Records - Part 1: Core [22]
- OASIS: MQTT 3.1.1 specification [23]
- OASIS: MQTT 5.0 specification [24]
- : Cloud Native Computing Foundation: OpenMetrics [25]

13.2. Informative

- WMO: WMO Information System 2.0 Strategy (WMO No. 1213) [26]
- WMO: WMO Guidelines on Emerging Data Issues (WMO No. 1239) [27]
- W3C: Data on the Web Best Practices [28]
- [[w3c-sdw-bp] [W3C/OGC Spatial Data on the Web Best Practices [29]
- OGC: Towards a Cloud-Native OGC. Part 1: The Vision: [30]
- OGC: Towards a Cloud-Native OGC. Part 2: The Standards: [31]
- OpenAPI Initiative: OpenAPI Specification v3.1.0 [32]
- OGC: OGC API standards [33]
- OGC: OGC API – Features – Part 1: Core [34]
- OGC: OGC API – Coverages – Part 1: Core draft [35]
- OGC: OGC API – Environmental Data Retrieval Standard [36]
- schema.org: Schema.org [37]
- Spatio Temporal Asset Catalogue: Spatio Temporal Asset Catalogue specification [38]
- wis2box: WIS2 in a box [39]
- Eclipse Foundation: Eclipse Mosquitto [40]
- RabbitMQ: RabbitMQ MQTT plugin [41]

- Prometheus: Prometheus [42]

- Grafana: Grafana [43]

[14] https://wmo-im.github.io/wcmp2

[15] https://github.com/wmo-im/wis2-topic-hierarchy

[16] https://wmo-im.github.io/wis2-notification-message

[17] https://github.com/wmo-im/wis2-metric-hierarchy

[18] https://www.w3.org/TR/webarch

[19] https://www.ietf.org/rfc/rfc3339.txt

[20] https://www.ietf.org/rfc/rfc3986.txt

[21] https://www.ietf.org/rfc/rfc7946.txt

[22] https://docs.ogc.org/DRAFTS/20-004.html

[23] https://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html

[24] https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html

[25] https://openmetrics.io

[26] https://library.wmo.int/index.php?lvl=notice_display&id=20422

[27] https://library.wmo.int/index.php?lvl=notice_display&id=21644

[28] https://www.w3.org/TR/dwbp

[29] https://www.w3.org/TR/sdw-bp

[30] https://www.ogc.org/blog-article/towards-a-cloud-native-ogc

[31] https://www.ogc.org/blog-article/towards-a-cloud-native-geospatial-standards-baseline

[32] https://spec.openapis.org/oas/v3.1.0

[33] https://ogcapi.ogc.org

[34] https://docs.opengeospatial.org/is/17-069r4/17-069r4.html

[35] http://docs.ogc.org/DRAFTS/19-087.html

[36] https://docs.ogc.org/is/19-086r5/19-086r5.html

[37] https://schema.org

[38] https://stacspec.org

[39] https://docs.wis2box.wis.wmo.int

[40] https://mosquitto.org

[41] https://www.rabbitmq.com/mqtt.html

[42] https://prometheus.io

[43] https://grafana.com

Annex A: Competencies

Seven competencies across four basic functional areas have been identified as follows:

1. Infrastructure: manage the physical infrastructure
2. Infrastructure: manage the operational applications
3. Data: manage the data flow
4. Data: manage data discovery
5. External interactions: manage interaction among WIS centres
6. External interactions: manage external user interactions
7. Overall service: manage the operational service

A.1. Competency 1: manage the physical infrastructure

A.1.1. Competency description

Prepare, plan, design, procure, implement and operate the physical infrastructure, networks and applications required to support the WIS centre.

Many of the skills required here are generic ICT skills and will have already been acquired as part of prior education and training or will be provided by hardware and system suppliers.

A.1.2. Performance components

A.1.3. Management of information technology operations

- Maintain the system in optimal operational condition by setting and meeting service levels, including:
  - Configuration;
  - Preventative and corrective maintenance and servicing;
  - Equipment replacement or upgrade;
  - Networking and processing capacity;
  - System monitoring and reporting procedures, and corrective actions;
- Provide contingency planning and operation backup and restoration;

A.1.4. Management of facilities

- Manage physical site security;
- Manage physical site environmental control.

A.1.5. Knowledge and skill requirements

- General ICT skills;

- Operation, configuration and maintenance of equipment and applications;

- Recognized information technology service management frameworks;

- Current technologies and emerging trends;

- Service level agreements.

A.1.6. Learning outcomes

Staff will be able to:

- Maintain the system in optimal operational condition;

- Plan for upgrades and operation backup and restoration;

- Maintain site security and environmental control.

Staff will learn:

- WIS specific systems;

- WIS site security policies;

- Service level agreements for the centre.

A.1.7. Learning activities

To learn how to perform the required tasks staff may:

- Attend training sessions run by providers of systems and other tools or by other training providers;

- Respond to typical monitoring reports;

- Apply WIS site security measures and respond to typical incidents;

- Apply WIS site environmental control measures and respond to typical incidents.

A.1.8. Assessment

Staff must be able to:

- Configure and maintain system components;

- Respond to monitoring reports;

- Apply WIS site security measures and respond to typical incidents;

- Apply WIS site environmental control measures and respond to typical incidents.

A.1.9. Key learning resources

- Manufacturers' handbooks and guides;

- Documentation of centre's facilities;

- WIS manuals and guides;

- Tools to monitor system security;

- WIS security policies;

- WIS environmental control policies.

A.2. Competency 2: manage the operational applications

A.2.1. Competency description

Prepare, plan, design, procure, implement and operate the applications required to support the WIS functions.

Many of the skills required here are generic ICT skills and will have already been acquired as part of prior education and training or will be provided by suppliers of applications.

A.2.2. Performance components

- Meet service levels by maintaining applications in optimal operational condition through:
- Configuration of applications;
    - Monitoring and responding to applications' behaviour;
    - Preventative and corrective maintenance;
    - Replacement or upgrade of applications;
- Provide contingency planning and application backup and restoration;
- Ensure data integrity and completeness in the event of system failure;
- Ensure system security.

A.2.3. Knowledge and skill requirements

- ICT skills;

- Operation, configuration and maintenance of applications;

- Recognized information technology service management frameworks;

- Current technologies and emerging trends;

- WIS functions and requirements;

- WIS security policies.

A.2.4. Learning outcomes

Staff will be able to:

- Operate, configure and maintain applications;

- Monitor applications and take corrective action;

- Apply and test WIS security protocols.

Staff will learn:

- WIS applications specific to the centre;

- WIS system security policies and procedures.

A.2.5. Learning activities

To learn how to perform the required tasks staff may:

- Attend training sessions run by providers of systems and other tools or by other training providers;

- Initiate monitoring and reporting procedures and respond to typical monitoring reports;

- Apply WIS site security measures and respond to typical incidents.

A.2.6. Assessment

Staff must be able to:

- Configure and maintain system components;

- Respond to monitoring reports;

- Apply site security measures and respond to typical incidents.

A.2.7. Key learning resources

Documentation of centre's applications;

- Manual on the WMO Information System (WMO No. 1060), Part IV. WIS Technical Specifications (WIS TechSpec 6);

- Guide to the WMO Information System (WMO No. 1061)

- Tools to monitor system security;

- WIS security policies.

A.3. Competency 3: manage the data flow

A.3.1. Competency description

Manage the collection, processing and distribution of data and products through scheduled and on demand services.

A.3.2. Performance components

- Ensure collection and distribution of data and products as per data policy;

- Publish data and products;

- Subscribe to data and products;

- Encode, decode, validate and package data and products;

- Create, update and maintain data flow catalogues;

- Manage connectivity between centres;

- Control the data flow to meet service levels.

- Knowledge and skill requirements

- System and network monitoring and viewing tools;

- Data formats and protocols;

- Licensing and data policies;

- Message and file switching systems.

A.3.3. Learning outcomes

Staff will be able to:

- Transfer data and products between their centre, other WIS centres and external users;

- Request data and respond to data requests using ad hoc and routine delivery mechanisms;

- Maintain quality standards (service levels) by monitoring, and responding to, traffic flow, missing data and products, errors and service messages;

- Apply relevant data policies to data and products;

- Identify appropriate formats for data and product exchange;

- Write and read data in data formats as agreed by WMO programs using their centre's tools.

Staff will learn:

- Data representations used in WIS and when to apply them;

- WMO data policies and how they apply to data in WIS;

- The structure of the WIS and how to use reference documents to identify and interpret the routing plans and protocols they will need to use;

- The interfaces of their centre's WIS applications, the information they use to modify their behaviour, and the tools available to control the operation of the applications to achieve service levels;

- How to use a WIS centre interface to find and request data for delivery by ad hoc request and by subscription;

- How WIS handles backup and alternative routings to maintain continuity of data flows.

A.3.4. Learning activities

To learn how to perform the required tasks staff may:

- Subscribe to the WIS services and download data from the Global Cache;

- Using a WIS centre platform, create, modify and delete a subscription for routine delivery of a dataset;

- Use the software tools of their centre's WIS application to exchange information between computers;

- Assess data flows by analysing monitoring reports from their applications;

- Investigate how data policy (including WMO Resolutions 1 (Cg-Ext 2021) is applied to data published by their centre;

- Use tools provided at their centre to view information in different formats and convert data between these formats.

A.3.5. Assessment

Staff must be able to:

- Go to a WIS centre, find data, download them immediately, subscribe for regular delivery and cancel the subscription;

A.3.6. Key learning resources

Data policies

- Resolution 1 (Cg-Ext 2021) – WMO Unified Data policy for the exchange of earth system data (weather, climate, hydrology, ocean, atmospheric composition, cryosphere, space weather) including guidelines on the relationships in commercial meteorological activities; The centre's data policies.

Data representations

- Manual on Codes (WMO No. 306), Volume I.1; Volume I.2 and Volume I.3;

- Guidance on migration to table driven code forms available at http://www.wmo.int/pages/prog/www/WMOCodes.html;

- Tools used at the centre to read, write, convert, validate and display information in Table Driven Code Forms;

- Sample data for reading and writing in Table Driven Code Forms.

WIS discovery, access and retrieval

- Manual on the WMO Information System (WMO No. 1060), Part IV. WIS Technical Specifications (WIS TechSpec 2, 3, 4, and 5);

- Guide to the WMO Information System (WMO No. 1061);

- User account at a GISC accessible via a Web browser.

Managing data exchange

- Manual on the WMO Information System;

- Weather Reporting (WMO No. 9), Volume C1Volume C1;

- Training environment on message and file switch;

- World Weather Watch quantity monitoring statistics

- WIGOS Data Quality Monitoring.

Security of data exchange

- Guide to Information Technology Security (WMO No. 1115).

Network management

- Network management tool and associated documentation;

- System error reports and event viewing tools.

A.4. Competency 4: manage data discovery

A.4.1. Competency description

Create and maintain discovery metadata records describing services and information, and upload them to the WIS Discovery Metadata Catalogue.

Each datum and product record held within WIS must have metadata associated with it so that it can be found and understood. These metadata records are held in a catalogue for discovery, access and retrieval (DAR).

A.4.2. Performance components

- Create and maintain discovery metadata records describing products and services;

- Add, replace or delete metadata records within the catalogue;

- Ensure that all information and service offerings from a WIS centre have complete, valid and meaningful discovery metadata records uploaded to the catalogue.

A.4.3. Knowledge and skill requirements

- Knowledge of WMO and ISO documentation sufficient to create complete and valid metadata;

- Metadata entry and management tools;

- Policies;

- Discovery metadata concepts and formats;

- Written English.

A.4.4. Learning outcomes

Staff will be able to:

- Use standard WIS tools to create discovery metadata from descriptions supplied by users;
- Add, replace or delete metadata records within the catalogue.

Staff will learn:

- The role of metadata in discovery, access and retrieval of data and products;
- Approved metadata formats;
- How to identify content that is mandatory, acceptable or inapplicable;
- Use of metadata creation tools;
- How to access and modify a catalogue;
- How data flow within, to and from their centre;
- About the tools that allow users to input descriptions.

A.4.5. Learning activities

To learn how to perform the required tasks staff may:

- Create metadata records based on sample descriptions for a range of data and products typical of their WIS centre;
- Insert such records into a catalogue, replace them with records that have been changed and delete them.

A.4.6. Assessment

Staff must be able to demonstrate:

- Successful creation of metadata records for typical products;
- Competence in publishing and deleting metadata catalogue records.

Key learning resources

- Manual on the WMO Information System (WMO No. 1060), Part IV WIS Technical Specifications (WIS TechSpec 1, 2, and 5), and Part V WIS Discovery Metadata;
- WCMP Homepage on WMO Community website
- WIS metadata guidance;
- Metadata entry and management tools;
- Samples of how to complete typical metadata records;
- Metadata policies and WIS metadata guidelines;
- ISO 19100 series: ISO standards on geographic information.

A.5. Competency 5: manage interaction among WIS Centres

A.5.1. Competency description

Manage relationships and compliance between the participants' centre and other WIS centres.

A.5.2. Performance components

- Exchange information with other centres on operational matters;
- Facilitate registration of new WIS centres;
- Facilitate registration of new data and products by other WIS centres;
- Create and respond to WIS service messages.

A.5.3. Knowledge and skill requirements

- Knowledge of current exchanges and requirements for notification of operational changes;
- Procedures and practices for registration of other centres and their data and products;
- Service level agreements;
- Written English.

A.5.4. Learning outcomes

Staff will be able to:

- Facilitate registration of new WIS centres and their data and products;
- Keep other WIS centres informed of the status of services, incidents and requests;
- Monitor and respond to service level reports;
- Manage subscriptions.

Staff will learn:

- About current exchanges and requirements for notification of operational changes;
- What type of data, products and services are available at their centre;
- Procedures and practices for registration of other centres and their data and products;
- Procedures and practices for notifying other centres about operational changes and service availability.

A.5.5. Learning activities

To learn how to perform the required tasks staff may carry out the above activities with the help of software, tools and guidance as used in their operational environment, either in a classroom or under supervision on the job.

A.5.6. Assessment

Staff must be able to:

- Respond to a request for registration of a new centre and its data and products;

- Prepare notifications of typical operational scenarios;

- Respond to typical notifications from other WIS centres.

A.5.7. Key learning resources

- Manual on the WMO Information System (WMO No. 1060), Part III Functions of WIS, and Part IV WIS Technical Specifications;

- Guide to the WMO Information System (WMO No. 1061);

- Weather Reporting (WMO No. 9), Volume C1;

- Exchanging Meteorological Data: Guidelines on Relationships in Commercial Meteorological Activities – WMO Policy and Practice (WMO No. 837).

A.5.8. Local resources

- Service level agreements (as used by the participants' centre);

- Frequently Asked Questions (FAQ) documents (for the user);

- WIS software user guides;

- Guidelines for services available at WIS centre;

- Data policy and associated guidance material;

- First line support procedures and guides;

- User database (for contact information);

- Case tracking and customer management;

- WIS user management;

- WIS subscription management;

- Monitoring dashboard for WIS components.

A.6. Competency 6: manage external user interactions

A.6.1. Competency description

Ensure that users, including other centres, data providers and subscribers, can publish and access data and products through WIS.

A.6.2. Performance components

- Register data providers and subscribers and maintain a service agreement;

- Set and register access criteria;

- Provide systems and support for users to publish and access data and products;

- Manage user relations to ensure a high satisfaction level.

A.6.3. Knowledge and skill requirements

- Data policies;

- External WIS interface;

- WIS registration and monitoring tools and policies;

- User support documentation and help files;

- Written English.

A.6.4. Learning outcomes

Staff will be able to:

- Register new WIS users and providers, setting roles, access authorizations and levels;

- Create and amend WIS user subscriptions;

- Use WIS tools to assist users and providers in resolving problems;

- Create and respond to WIS service messages;

- Undertake first line investigation and diagnosis;

- Manage incidents and requests: log them, categorize and prioritize them, escalate as appropriate and close them when the user is satisfied;

- Keep users informed of the status of services, incidents and requests;

- Gather information and report on user and provider satisfaction;

- Assist users in uploading and accessing data;

- Identify potential problems in services and implement improvements.

Staff will learn:

- What type of data, products and services are available at their centre;

- How WIS applications, including discovery, access and retrieval (DAR) should be used;

- How to apply data policies;

- How to interact effectively with users and providers.

A.6.5. Learning activities

To learn how to perform the required tasks staff may:

- Register users (data providers and subscribers) and set access authorizations and levels using the same software, tools and guidance as in their operational environment;

- Role play user interactions.

A.6.6. Assessment

Staff must be able to:

- Register typical data providers and users;

- Ensure that users are able to upload and access data;

- Respond to typical incidents.

A.6.7. Key learning resources

- Manual on the WMO Information System (WMO No. 1060), Part IV WIS Technical Specifications (WIS Tech Spec 2, 3, 4, and 5);

- Guide to the WMO Information System (WMO No. 1061);

- Weather Reporting (WMO No. 9), Volume C1;

- Exchanging Meteorological Data: Guidelines on Relationships in Commercial Meteorological Activities – WMO Policy and Practice (WMO No. 837).

Local resources

- Service level agreements (as used by their centre);

- FAQ documents (for the user);

- WIS software user guides;

- Guidelines for services available at WIS centre;

- Data policy and associated guidance material;

- First line support procedures and guides;

- User database (for contact information);

- Case tracking and customer management;

- WIS user management;

- WIS subscription management;

- Monitoring dashboard for WIS components.

A.7. Competency 7: manage the operational service

A.7.1. Competency description

Ensure the quality and continuity of the service.

This is essentially a management role ensuring that the WIS system operates as required, now and in the future. Some of the skills required are generic management skills, rather than WIS specific, and would be taught or learnt elsewhere.

A.7.2. Performance components

- Coordinate all WIS functions and activities of the centre;

- Ensure and demonstrate compliance with regulations and policies;

- Monitor and meet quality and service performance standards;

- Ensure service continuity through risk management and planning and implementation of service contingency, backup and restoration. Ensure data continuity in the event of system failure;

- Plan and coordinate the delivery of new functionalities.

A.7.3. Knowledge and skill requirements

- General management skills;

- Overview of local and external WIS operations and associated service agreements;

- WIS regulations and policies;

- Functional specifications;

- Written English.

A.7.4. Learning outcomes

Staff will be able to:

- Ensure that the WIS centre meets quality and service performance standards;

- Identify the challenges and issues to be addressed;

- Foster compliance with WIS framework.

Staff will learn:

- Functions and responsibilities of the WIS centre;

- WIS quality and service performance standards;

- Methods to manage quality, risk and operational service;

- How to monitor quality and service performance standards;

- How to analyse, demonstrate and report quality and service performance at the WIS centre;

- How to maintain troubleshooting, backup and restoration procedures;

- How to plan and coordinate the delivery of new functionalities and improvements;

- How to integrate new technologies and developments;

- How to update the regulatory documents;

- How to maintain service agreements;

- How to plan monitoring resources;

- How to align budget restrictions with human resources demands.

A.7.5. Learning activities

To learn how to perform the required tasks staff may:

Follow quality and service performance standards; Analyse quality and service performance in the WIS centre; Demonstrate and report quality and service performance; Maintain troubleshooting, backup and restoration procedures; Plan and coordinate the delivery of new functionalities; Keep timely records, as required.

A.7.6. Assessment

Staff must be able to:

- Demonstrate successful WIS service;
- Plan replacement and upgrade of equipment and applications to meet new functionalities and requirements.

A.7.7. Key learning resources

- Technical Regulations (WMO No. 49), Volume I;
- Resolution 1 (Cg-Ext 2021) – WMO Unified Data Policy for the exchange of earth system data (weather, climate, hydrology, ocean, atmospheric composition, cryosphere, space weather) including guidelines on the relationships in commercial meteorological activities;
- Manual on the WMO Information System (WMO No. 1060), Part IV WIS Technical Specifications (WIS TechSpec 6);
- Guide to the WMO Information System (WMO No. 1061);
- WIS demonstration procedures and guidelines;
- Monitoring reports;
- Audit reports.

Annex B: Recommended practices for audit and certification of WIS Centres

B.1. Background

The Infrastructure Commission is responsible for certification of WIS centres' compliance with the WIS technical specifications defined in the Manual on WIS. The INFCOM will maintain, within the structure of its SC-IMT on Information Management and Technology (SC-IMT), and the Expert Team on Audit and Certification (ET-AC) to coordinate audits and certification of WIS centres.

Audits and certifications will be carried out in line with the principles established in the Technical Regulations (WMO-No. 49), Volume I: General Meteorological Standards and Recommended Practices, Part VII.

B.2. Auditing and certification

Auditors and certifiers shall be or shall become members of ET-AC. New members must have relevant technical or auditing experience. They must be members (core or associate) of an SC-IMT expert team or have written commitment of the Permanent Representative of their country with WMO allowing them to participate as members of the ET-AC. New members will be mentored by a nominated existing expert. Note that regional diversity of members of ET-AC is essential. Access to ET-AC Workspace and online databases is restricted to ET-AC and the WMO Secretariat.

B.2.1. GISC audits

The Expert Team on Audit and Certification, on behalf of INFCOM, is responsible for auditing and certification of GISCs.

A GISC should be audited by two experts, one of whom must have previous experience of auditing GISCs. Auditors should be from a different region than that of the GISC.

Travel and per diem should be at the GISC's expense and arranged through WMO.

B.2.1.1. Scope of GISC audits

Full audits will cover all aspects of WIS compliance and shall include site visits using practices inline with those of the ISO 9000 series standards.

Interim audits will focus on a particular subset of topics. Actual elements to be focused on will be determined by the Standing Committee on Information Management and Technology (SC-IMT) or its delegated expert team. Centres will be told in advance on which subset of topics the interim audit will focus. Possible areas for review in interim audits include:

a) GISC operations; b) Metadata management; c) Monitoring; d) Incident management e) Quality of service provided by the GISC; e) Coordination within its Area of Responsibility: i) Management of the GISC Area of Responsibility; ii) Capacity development; iii) Participation in WIS coordination and planning mechanisms

B.3. DCPC certification

Data Collection or Production Centres are to be certified by the ET-AC. Where a DCPC is not using the infrastructure of its principal GISC, and its principal GISC is operational, it can be certified by ET-AC once the principal GISC has performed the necessary tests. However, if the principal GISC is not operational, the ET-CAC will arrange for a suitable GISC to perform the tests. Where a DCPC uses the infrastructure of its principal GISC, it is certified as a part of the GISC certification process.

The certification of a DCPC requires only one ET-AC coordinator, who will ask a GISC to undertake tests with the DCPC. It is expected that the centre's principal GISC will undertake those tests.

B.3.1. Verification of compliance of NCs

Compliance of NCs is the responsibility of the Permanent Representative with WMO of the Member accountable for the centre. Verification of compliance of an NC should be done by its principal GISC. The Expert Team on Audit and Certification will monitor the NC compliance process in consultation with NCs and GISCs.

B.4. The review cycle

The review cycle should start from the date of INFCOM endorsement. Audits should take place within the calendar year in which the cycle ends, and their timing will need to be coordinated with the experts called upon to undertake them.

The INFCOM endorsement date should be recorded in the WIS centre database. The date on which the centre became operational should also be recorded if known. Similarly, to an ISO 9001:2015 audit process, the GISC audit will follow the principle of alternating intermediate and full audits aligned with the INFCOM/EC four-year cycle:

a) Intermediate audit (interim, four years): a mid-cycle review of performance and compliance to provide, if necessary, opportunities to introduce corrective actions well in advance of a full audit; b) Full audit (every second audit, i.e., every eight years): this audit will result in a recommendation for confirmation or cancellation of endorsement.

B.4.1. Review of DCPCs

The DCPC review cycle will be eight years. Reviews will cover all aspects of WIS compliance.

B.4.2. Review of NCs

Review of NC compliance is the responsibility of the Permanent Representative with WMO of the Member responsible for the Centre in liaison with the NC and its principal GISC.

B.5. Ad hoc audits or reviews

An ad hoc audit or review can be requested by the president of INFCOM due, for example, to non-conformance causing problems with WIS operations.

B.6. Audit or review outcome

The outcome of the audit or review will be categorized as "endorsed", "endorsed with qualification" or "not endorsed". Audit or review recommendations will be provided to the president of INFCOM and to the Director of WIS.

B.7. Format of report

The Expert Team on Audit and Certification will use a template for final reports, although the content will reflect the areas audited.

B.8. Public notification of type of INFCOM endorsement

The endorsement of INFCOM is based on continued successful audit outcomes. Centre endorsements are published only as "INFCOM endorsed" with no public declaration of whether endorsement was with "qualifications".

Details of reviews and audits of centres are confidential. Auditors will have access to the previous reports on a centre to perform their role.

B.9. Review of audits with qualification

Global Information System Centres that were "endorsed with qualifications" have two years from the date of the audit to demonstrate that they have taken remedial action on the points of qualification.

The Expert Team on Audit and Certification will investigate GISCs that were "endorsed with qualifications" and have not demonstrated that they have taken remedial action within two years of the date of audit. The Expert Team should report to INFCOM on progress in addressing the aspects that incurred the "qualification" and can recommend to INFCOM that it revokes its endorsement.