
MEETING OF THE
Expert Team on WORLD DATA CENTRES
(NILU, Kjeller, Norway, 2-5 October 2017)

Report of the Chair of ET-WDC

(submitted by Jörg Klausen)

1. Terms of Reference

- ET/WDC-1. Take responsibility for metadata and data management issues in support of the scientific and operational objectives of GAW.
- ET/WDC-2. Work with the SAGs, WMO expert teams and partners to establish harmonized data management guidelines, including standardised data formats to allow for adequate (seamless) interoperability.
- ET/WDC-3. Guide and support the further development of GAWSIS as the central catalogue of observing facilities and observations supporting GAW, linking the WDCs and Contributing Data Centres.
- ET/WDC-4. Keep abreast of and recommend best use of changing technologies affecting information management within GAW.

(Reference: [WMO Global Atmosphere Watch \(GAW\) Implementation Plan: 2016-2023](#), WMO GAW Report No. 228, WMO 2017)

2. Activities since Meeting in 2014

2.1. Ad-hoc meeting 2015

This 1-day meeting took place at MeteoSwiss as part of a larger international workshop to discuss and recommend the “Future Directions of GAW Data Management”. A report of that workshop is available at <https://tinyurl.com/y8kpggce>.

Key outcomes, issues and conclusions of the meeting were as follows

- Participants agreed on the need for global, high-quality atmospheric composition data coordination
- There is support for GAWSIS as the international metadata gateway/hub for access to atmospheric composition data
- The distribution of data among Met Services and other sources is a current reality
- There is a need for the management of data (other than metadata) whether it exists with a WDC or other data centres
- There is a need to define how other data centres/contributing networks could be included as part of the federated approach (e.g., global data gets submitted to WDCs and regional data to regional databases).
- There is a recognition that future data management system should be built as a **federated system** of the data archiving and service providing centers.

MEETING OF THE
Expert Team on WORLD DATA CENTRES
(NILU, Kjeller, Norway, 2-5 October 2017)

The results of this workshop were the basis for the revised ToRs as reflected in section 1 above.

2.2. Telecons

- June 2014 (telecon-8)
- August 2014 (telecon-9)
- October 2014 (telecon-10)
- December 2015 (telecon-11)
- June 2016 (telecon-12)
- August 2016 (telecon-13)
- March 2017 (telecon-14)
- May 2017 (telecon-15)
- July 2017 (telecon-16)

2.3. Membership of ET-WDC

Following CAS-16 in 2013, the membership structure of advisory bodies and expert teams was updated. Since then, the expectation is that SAGs and ET-WDC are composed of 'regular' members (who would have 2x4 year terms) and 'ex officio' members. It took ET-WDC quite a while to implement the new guidelines. Meantime, ET-WDC has successfully recruited a number of 'regular' members that are not directly involved in the operation of the WDC and GAW SIS infrastructures (cf. <https://sites.google.com/site/wmoetwdc/the-team>). This meeting is actually the first time most of them will meet in person.

3. GAW Strategic Plan: Implications for ET-WDC

In addition to the ToRs/Tasks of ET-WDC, the GAW Implementation Plan defines specific tasks for the World Data Centres. The following provides a personal assessment of the author of the current state of WDCs with respect to these tasks.

WDC-1. Provide adequate archiving facilities for observational data for which GAW has global coordination responsibilities.

- Given their long-term operations, most WDCs would seem to provide adequate archiving facilities, albeit not all of them are at the same level of maturity. BCM measures are not published. Provenance and lineage information and service-orientation varies among the centers.
- Action: ET-WDC should establish a catalogue of criteria to define adequacy in an objective manner. Only then can assessments be made.

WDC-2. Check submitted data for necessary format elements and the availability of comprehensive metadata and reject the submission of data that do not meet these formal criteria.

MEETING OF THE

Expert Team on WORLD DATA CENTRES
(NILU, Kjeller, Norway, 2-5 October 2017)

- This is done by all WDCs, but the criteria vary a great deal. For example, the requirements for description of instrumentation and procedures used to generate data are very diverse.
- Action: All WDCs under WMO/GAW need to adopt the core elements of the WIGOS metadata standard (WMDS) as a minimal standard. The WMDS offers a lot of optional metadata elements that significantly increase the possibility of adequate use of observations.

WDC-3. Perform plausibility and consistency checks on submitted data, flag data problems, and provide feedback to the data providers, when necessary.

- It was previously agreed (Tokyo meeting 2014) that all WDCs should publish reports on their internal procedures applied to submissions. To date, such reports have either not been prepared or not been discussed within ET-WDC.
- Action: Provide full transparency about handling of data submissions and bring this to the attention of ET-WDC. Again, it may be useful to agree on a common metric and standardize this reporting to some degree.

WDC-4. Continually improve the ease of access to data of known quality by evolving WDC operations in line with the development of WIGOS and with particular attention to the increasing needs for NRT data services.

- Most WDCs have improved access to data through the web or are in the process of upgrading their websites. These developments almost exclusively take place without mutual advice/feedback from ET-WDC. Often, the results of such improvements are presented to the team after the fact. It is difficult to see how convergence of the WDCs can be achieved under such circumstances. With respect to NRT data management, this is in its infancy at most WDCs, partly because NRT data submissions are still relatively rare. In Europe at least, with the advent of Copernicus Atmospheric Service (CAMS), but surely also elsewhere, the need for NRT data is increasing ever more, and service contracts are being established to facilitate the availability of NRT observation data. WDCs have a role to play, and ET-WDC needs to develop a strategy.
- Action: Explore within ET-WDC to what extent WDCs can benefit from each other with respect to adopting technological but also conceptual advances. Also, to what extent WDCs could benefit from the operational data delivery mechanisms of the GTS/WIS.

WDC-5. Contribute to the agreement of standards for interoperability of data archives through the Expert Team on GAW World Data Centres (ET-GAW WDCs). This also includes support for the establishment of harmonized guidelines and data formats for the submission and dissemination of atmospheric composition data, metadata and products.

- While developments are ongoing at most WDCs, there is little evidence so far that interoperability among them is taken seriously at present. In Europe, INSPIRE provides some guidelines, and WMO/WIS has adopted OAI-PMH. No common data model for data exists, and metadata are not interoperable in general. At the same time, WMO Congress (Cg-17) has adopted the WIGOS Metadata Standard. The H2020 NextGEOSS project provides some support to a concerted effort.
- Action: Develop a common data model based on international standards and software tools and guidance for transformations of various data formats. OM_Observations, INSPIRE, C3S CDM and other standards should be considered.

MEETING OF THE
Expert Team on WORLD DATA CENTRES
(NILU, Kjeller, Norway, 2-5 October 2017)

- WDC-6. Support and participate in the establishment of a distributed data management system involving all WDCs, the archives of contributing networks, and GAWSIS as the central metadata repository for discovery and access purposes.
- This strategic goal in the GAW Implementation Plan summarizes the objectives at a high level.
 - Action: Develop a work plan with specific milestones and deliverables that can be used to guide the implementation at all WDCs, as well as interested archives of Contributing Networks.

4. Outlook

The following is quoted almost verbatim from the same document (Chair's report) submitted to the ET-WDC 2014 meeting in Tokyo. I believe it is still valid and reflects what lies ahead.

Our common understanding of data management, metadata and interoperability have developed during the last 10 years ... but technology and the numbers of stakeholders have advanced even more. For GAW and GAW WDCs, the challenge of remaining a leader in the actual implementation and operation of observations is increasing. Unfortunately, even as the number of funding opportunities increases, the WDCs are not always or often even not at all part of these projects, programmes, initiatives. If GAW WDCs want to be more than the archives of old programs such as Dobson, or secondary archive for NOAA, or a labeling facility for EC project data, or a host for data of stations that are not involved in other projects/programs, then we need to

- Provide a service of high quality
- Be attractive for data users and science (so that no project data bases cum archives are needed, but instead our WDCs are being involved)

How can this be achieved? In my mind, we need a common vision that translates into common goals that translate into common action. Elements of this are

- ➔ Increased interoperability
- ➔ Increased standardization
- ➔ Improved data policy that takes into account the needs of projects (and perhaps restricts data access for some time)
- ➔ Involvement of some sort in the NRT data delivery services and strengthening our role in establishing mechanisms so that NRT data make it to our WDCs after adequate QA/QC and in a timely manner.
- ➔ develop a common (from the provider/user point of view) input/output facility for data and metadata related to GAW based on a distributed system of WDCs / DCs / GAWSIS with automatic conversion of formats ('virtual GAW data centre' with common look-and-feel), where
 - data can be submitted to and where the metadata determine where the data are processed and archived
 - data can be retrieved from in a number of formats with standardized metadata

MEETING OF THE

Expert Team on WORLD DATA CENTRES
(NILU, Kjeller, Norway, 2-5 October 2017)

- Ensure the data quality in the archives, including strict harmonisation of formats ("no tolerance policy for inconsistencies or errors in formats, units etc"), standardisation of metadata within and among WDCs
- Establish a system for archiving original raw data for ground based total ozone measurements, including the associated calibration history information to facilitate accurate traceability of the instruments' performance (same holds true for all WDCs)
- getting data from 'large urban complexes' on the radar of GAW WDCs / GAWSIS
- improve and formalize the documentation of traceability of observations and the ensuing data and their versions

The vision:

