

MEETING OF THE  
Expert Team on WORLD DATA CENTRES  
(JMA, Tokyo, Japan, 21-23 January 2014)

## **GAW World Data Centers: Roles, Responsibilities, Collaborations, Data Flows**

Submitted by Jörg Klausen, with contributions from Tony Colavecchia, Hiroshi Koide

### Table of Contents

1.	Introduction .....	2
2.	Terms of Reference for the GAW WDCs .....	2
3.	Terms of Reference GAW SAGs and ETs: .....	2
4.	WOUDC .....	3
4.1.	Scope of Operation .....	3
4.2.	Existing Collaborations .....	3
4.3.	Data Flow Diagram .....	3
4.4.	Other Related Data Centers .....	3
4.5.	Way forward .....	3
5.	WDCGG .....	4
5.1.	Scope of Operation .....	4
5.2.	Existing Collaborations .....	4
5.3.	Data Flow Diagram .....	4
5.4.	Other Related Data Centers .....	4
5.5.	Way forward .....	4
6.	WDCA .....	5
6.1.	Scope of Operation .....	5
6.2.	Existing Collaborations .....	5
6.3.	Data Flow Diagram .....	5
6.4.	Other Related Data Centers .....	5
6.5.	Way forward .....	5
7.	WRDC .....	5
7.1.	Scope of Operation .....	5
7.2.	Existing Collaborations .....	5
7.3.	Data Flow Diagram .....	5
7.4.	Other Related Data Centers .....	5
7.5.	Way forward .....	5
8.	WDC-RSAT .....	6
8.1.	Scope of Operation .....	6
8.2.	Existing Collaborations .....	6
8.3.	Data Flow Diagram .....	6
8.4.	Other Related Data Centers .....	6
8.5.	Way forward .....	6
9.	GAWSIS .....	6
9.1.	Scope of Operation .....	6
9.2.	Existing Collaborations .....	7
9.3.	Data Flow Diagram .....	7
9.4.	Other Related Data Centers .....	7
9.5.	Way forward .....	8

## 1. Introduction

This document attempts to review the roles and responsibilities of the GAW WDCs, existing collaborations and data flows involving the GAW WDCs, and data flows not involving the GAW WDCs in an environment of rapid change. The purpose of this analysis is to help establish the priorities for the GAW WDCs for the coming years.

This rapid change is due to an ongoing revolution in ICT and internet connectivity, an increased awareness of the impacts of climate change and air pollution on society, and a desire for increased visibility of funded research project, articulated through project data centers that grow into (semi-)persistent archives. As a result, the GAW WDCs are no longer the only players in the field of global archiving of observations, data management, and developing guidelines for data submission and metadata.

For each GAW WDC, the scope of operation, existing collaborations and an overview of the data flows involving and not involving the WDCs is shown.

## 2. Terms of Reference for the GAW WDCs<sup>1</sup>

- (a) Assist users of atmospheric chemistry observations for which GAW has global coordination responsibilities to easily access observations of known quality through an appropriate archiving or archive networking mechanism.
- (b) Establish harmonized guidelines and data formats for data submission and retrieval for each parameter.
- (c) 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.
- (d) Perform plausibility and consistency checks on submitted data, flag data problems, and provide feedback to the stations, when necessary.
- (e) Participate in – and contribute to meeting the goals of – the Expert Team on GAW World Data Centres (ET-GAW WDCs)
- (f) Keep abreast of and evolve WDC operations in line with the development of WIS with particular attention to harmonizing data formats and the needs of NRT data delivery.

## 3. Terms of Reference GAW SAGs and ETs<sup>2</sup>:

- (a) Keep informed of scientific and technical developments in their field;
- (b) Advise the SSC and Members on the developments, priority areas and progress in their field, taking into consideration user requirements;
- (c) Contribute to the review of the GAW Implementation Plan;
- (d) Implement recommendations, tasks and projects as defined in the GAW Implementation Plan;
- (e) Review the status of the relevant part of the GAW network, ensure active collaboration with contributing networks, and advise on further enhancement of the observational capacity of the GAW Programme, fostering the development of QMF;
- (f) Promote timely data delivery;
- (g) Develop a broad spectrum of products and services, taking socio-economical needs into consideration;

---

<sup>1</sup> WMO (2007), GAW Strategic Plan: 2008-2015, WMO/GAW Report No.172

<sup>2</sup> WMO (2013), CAS-XVI Doc. 10 Annex 2 to Resolution 2, approved

- (h) Interact effectively with the WMO Secretariat on matters related to the GAW Programme;
- (i) Take into consideration activities by other UN organizations and institutes in their field and collaborate as appropriate.

## 4. WOUDC

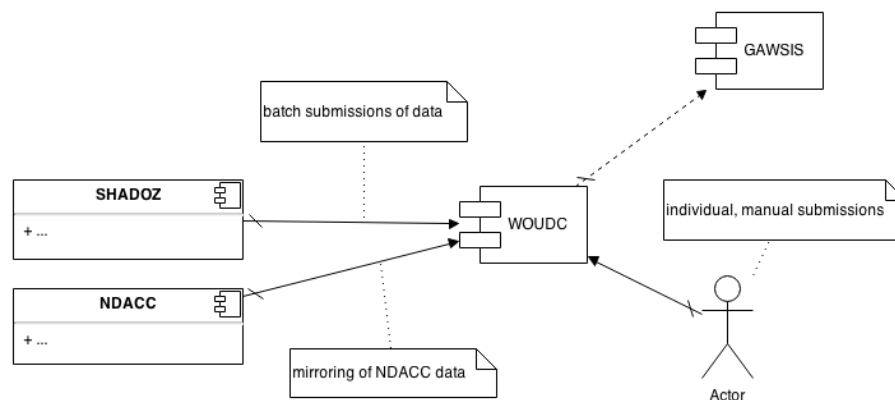
### 4.1. Scope of Operation

[please describe very briefly what kind of data you archive in terms of variables, quality-controlled vs nrt data, most important user group]

### 4.2. Existing Collaborations

- SHADOZ (Southern Hemisphere ADditional OZonesondes) they are sending us the data and we ingest it into WOUDC
- Network for the Detection of Atmospheric Composition Change (NDACC) – given the complexities we are mirroring and making them available through WOUDC

### 4.3. Data Flow Diagram



### 4.4. Other Related Data Centers

[please enumerate other data centers you are aware of that provide similar services, but with whom no data exchange is established. Describe the data flows that exist and give reasons why there is no collaboration]

- European UV Database (EUVDDB) – Exploring mirroring/data management options with FMI

### 4.5. Way forward

Two possible options:

- Decentralized - provide links to related data centres and find some manner (standard exchange format, web services...) to improved interoperability
- Centralized – ingest or mirror data into our data centre

## 5. WDCGG

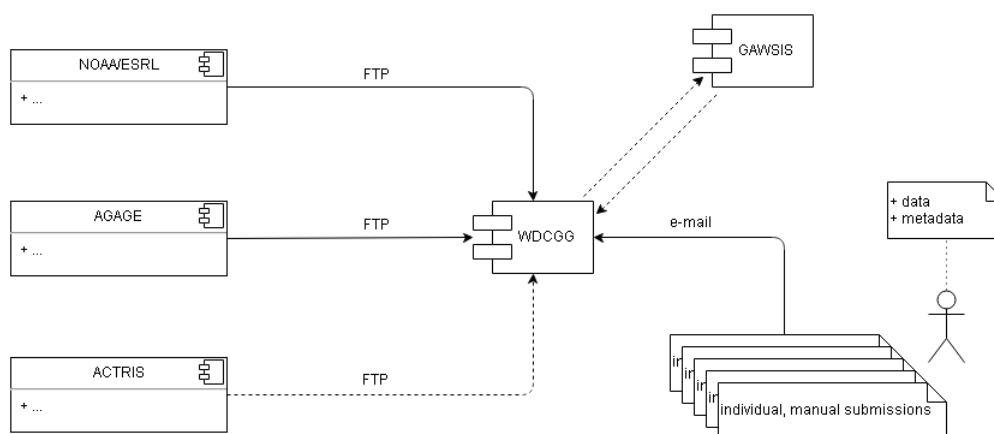
### 5.1. Scope of Operation

WDCGG archives measurement data on greenhouse gases (GHGs:CO<sub>2</sub>, CH<sub>4</sub>, CFCs, N<sub>2</sub>O, surface ozone, etc.) and reactive gases (RGs:CO, NO<sub>x</sub>, SO<sub>2</sub>, VOCs, etc.) in the atmosphere and ocean, as observed under GAW and other programmes. Both ground station and mobile platform (airplane, ship) based measurement data are accommodated. The treatment of column total GHG measurement with satellites (e.g., GOSAT) and ground based FTIR (TCCON) is a grey zone at this moment. The most important user group is GGMT (WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases, and Related Tracer Measurement Techniques).

### 5.2. Existing Collaborations

The largest data contributor to WDCGG is NOAA/ESRL with its global flask sampling network. AGAGE is another core submitter of RG parameters. Since 2013, NILU ACTRIS collaborate with WDCGG on bulk data submission of some RG parameters from European stations. CONTRAIL project submits limited data to WDCGG.

### 5.3. Data Flow Diagram



### 5.4. Other Related Data Centers

- ICOS Atmospheric Thematic Center (ATC): They started their service very recently focusing Europe. ICOS ATC accommodates NRT data but WDCGG has no experience in this area at this moment.

### 5.5. Way forward

In 2012, WDCGG started a 5-year entire reforming process of its service to serve better data contributors and users alike. The reform includes rebuilding the meta-database and the web site interface in consultation with relevant science community (SAG and ET, GGMT, etc.) and the WMO Secretariat. The final goal of WDCGG is to make the contributors and the users happy with its services and to thrive as a core of various information exchanges about gas measurement data, their applications and quality assurance.

## **6. WDCA**

### **6.1. Scope of Operation**

[please describe very briefly what kind of data you archive in terms of variables, quality-controlled vs nrt data, most important user group]

### **6.2. Existing Collaborations**

[please describe what formal collaborations with other data centers exist]

### **6.3. Data Flow Diagram**

[an example is given below; best to provide as a Power Point Slide, so that we can integrate the various contributions]

### **6.4. Other Related Data Centers**

[please enumerate other data centers you are aware of that provide similar services, but with whom no data exchange is established. Describe the data flows that exist and give reasons why there is no collaboration]

### **6.5. Way forward**

[please suggest how the situation can be improved (if there is a need), how you see that collaboration could be fostered and data flows be integrated]

## **7. WRDC**

### **7.1. Scope of Operation**

[please describe very briefly what kind of data you archive in terms of variables, quality-controlled vs nrt data, most important user group]

### **7.2. Existing Collaborations**

[please describe what formal collaborations with other data centers exist]

### **7.3. Data Flow Diagram**

[an example is given below; best to provide as a Power Point Slide, so that we can integrate the various contributions]

### **7.4. Other Related Data Centers**

[please enumerate other data centers you are aware of that provide similar services, but with whom no data exchange is established. Describe the data flows that exist and give reasons why there is no collaboration]

### **7.5. Way forward**

[please suggest how the situation can be improved (if there is a need), how you see that collaboration could be fostered and data flows be integrated]

## **8. WDC-RSAT**

### **8.1. Scope of Operation**

[please describe very briefly what kind of data you archive in terms of variables, quality-controlled vs nrt data, most important user group]

### **8.2. Existing Collaborations**

[please describe what formal collaborations with other data centers exist]

### **8.3. Data Flow Diagram**

[an example is given below; best to provide as a Power Point Slide, so that we can integrate the various contributions]

### **8.4. Other Related Data Centers**

[please enumerate other data centers you are aware of that provide similar services, but with whom no data exchange is established. Describe the data flows that exist and give reasons why there is no collaboration]

### **8.5. Way forward**

[please suggest how the situation can be improved (if there is a need), how you see that collaboration could be fostered and data flows be integrated]

## **9. GAWSIS**

### **9.1. Scope of Operation**

GAWSIS maintains the official catalogue of stations/platforms contributing observations in support of GAW. GAWSIS therefore hosts metadata only, on stations/platforms, measurements performed, contact persons and also all kinds of related publications, technical reports, etc. All stations/platforms accepted into the GAW program (as GAW stations or contributing stations) are eligible for listing in GAWSIS.

So far, the focus has been mostly on fixed stations. In future, GAWSIS will be the source on metadata describing observational capabilities on atmospheric chemical composition and related physical characteristics of the atmosphere for OSCAR in the framework of WIGOS and in full support of the mandate of GAW. Thus, the ambition is to include mobile platforms (in particular, ships and aircraft) to facilitate a comprehensive Rolling Review of Requirements (RRR) process for GAW as stipulated by WIGOS.

Moreover, the next release of GAWSIS will permit to list n.r.t. data streams in addition to documenting the quality-controlled data archived at the GAW World Data Centres and related archives.

The most important user groups of GAWSIS are arguably the WMO Secretariat and individual users looking for an overview of activities at any given station/platform.

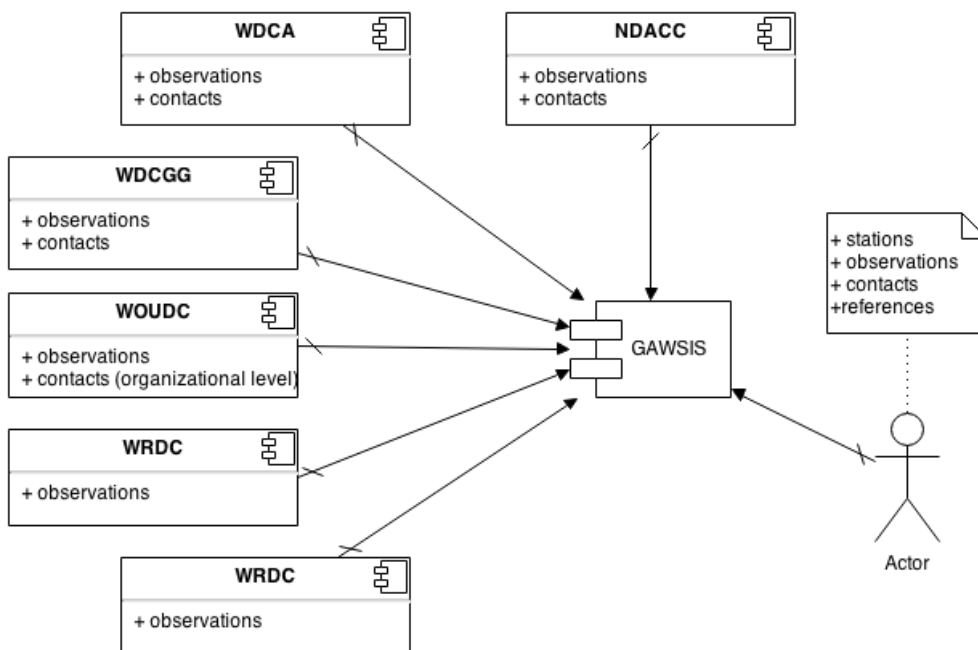
Furthermore, GAWSIS support network development on a global and regional scale (while, at the national level, comprehensive catalogues usually exist.)

GAWSIS links GAW observations to the WMO Information System (WIS) by regularly publishing WIS-compatible metadata records (cf. <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx3bW9ldHdkY3xneDo5OGMwMTE3NDc0MDk0YTQ.>)

## 9.2. Existing Collaborations

Collaborations exist primarily in the context of the Expert Team on GAW World Data Centres (ET-WDC) and the related data archives. These include at present: WOUDC, WDCGG, WRDC, WDCA, NDACC, and EMEP. Moreover, relatively loose collaborations exist with contributing programs such as LALINET, EARLINET, ALINE, AD-Net, TCCON, IMPROVE.

## 9.3. Data Flow Diagram



## 9.4. Other Related Data Centers

WDC-RSAT is a portal integrating many of the satellite observations of relevance to GAW. Satellite platforms have been outside the scope of GAWSIS in the past, however, there is no principle reason for that.

WDCPC is being developed as a global portal for precipitation chemistry. WDCPC is presently unable to share metadata with GAWSIS. NADP is currently negotiating an MoU with WMO that will also include some possibility for metadata exchange.

Several networks for aerosol optical depth exist that are only partially reflected in GAWSIS, and with which no automatic metadata exchange has been established so far. These include AERONET, SKYNET, EANET, and possibly others.

Most of the stations/platforms associated to the various networks forming GALION are included in GAWSIS, but no automatic metadata exchange has been established so far.

EMEP is a significant regional network in RA VI. The scope of EMEP partially overlaps with the scope of GAW, and certain metadata from EMEP are included in GAWSIS.

BSRN is a contributing network to GAW. Metadata exchange has not been a priority in the past, and has not been developed due to resource constraints.

### **9.5. Way forward**

GAWSIS is being re-developed in 2014 and will be fully deployed within the MeteoSwiss ICT infrastructure. As part of this, the formal acceptance procedure for GAWSIS as a DCPC in WIS will be completed (GAWSIS already publishes WIS-compatible metadata records.) The use of state-of-the-art technologies will alleviate some of the technical restrictions encountered in recent years and will positively impact the ability to consume and offer (more standardized) web services. Also, significant improvements of the user experience (speed, useability, consistency) are expected. Moreover, it is planned to markedly improve the usefulness of GAWSIS to answer questions related to 'health of networks' and capabilities of networks. GAWSIS will remain a stand-alone application with a focus on GAW, i.e., atmospheric composition and related physical characteristics of the atmosphere.) Moreover, it will serve as the source of information on capabilities of GAW in the context of the Observation Systems Capabilities and Review (OSCAR) tool for WIGOS.

Specific improvements are foreseen in consuming metadata from the existing sources. Closer collaboration with WDCPC (and/or directly with the individual programs), BSRN, and several South-East Asian programs is expected.

In order to provide a more holistic view on GAW measurements, n.r.t. data streams will be listed in future.

Closer collaboration / integration with the GEO AQ CoP and ESIP will be attempted and is expected to support convergence of approaches.