

ET-WDC Meeting Report

WebEx Teleconference on 7 June 2016

Report finalised on 30 June 2016 by Geir Braathen and Jörg Klausen.

Agenda

1. Welcome by chair
2. New GAW IP and relevance for ET-WDC
3. WIGOS Metadata Standard (WMDS) and WIGOS Metadata Representation (WMDR)
4. Information from participants
5. Review next steps (action items) agreed today
6. Next Meeting

Participants

Geir Braathen
Tony Colavecchia
Markus Fiebig
Jörg Klausen
Martin Schultz
Anatoly Tsvetkov
Kjetil Tørseth
Mikio Ueno

Apologies

Ellsworth Judd Welton, Øystein Godøy, Dominic Lowe, Van Bowersox

1. Welcome by the Chair

The ET-WDC Chair, Jörg Klausen, welcomed the participants to the WebEx conference and started out by thanking the SSC Chair, Prof. Greg Carmichael for agreeing on the appointment of three new ET-WDC members. He then presented the four new members of the Team. Unfortunately, three of these new members were not able to attend the WebEx conference.

1. Martin Schultz from FZ Jülich is chair of the RG SAG. He has participated in several ET-WDC meetings already as an invited expert due to his interest in questions around data exchange.
2. Ellsworth Judd Welton from NASA Goddard Space Flight Center runs MPLNet. He played an active role at last year's meeting in Zürich.
3. Øystein Godøy from the Norwegian Meteorological Office (met.no) is Norway's national contact point for Global Cryosphere Watch (GCW) and responsible for the GCW data portal.
4. Dominic Lowe is from the Australian Bureau of Meteorology. He is familiar with OGC standards and he has helped with WIS /WIGOS standards.

In addition there are also two new ex officio members:

5. Mikio Ueno, Japan Meteorological Agency (JMA) is the new head of WDCGG. Before this, he was the head of the Ozone Layer Monitoring Center at JMA.
6. Kjetil Tørseth, Norwegian Institute for Air Research (NILU) is the interim manager of the WMO/GAW World Data Centre for Reactive Gases. He will keep this role for a while and then hand it over later. Sverre Solberg (NILU) will be active in this together with Markus Fiebig.

1b. Approval of Agenda

The agenda was approved without changes.

2. New GAW Implementation Plan (IP) and Relevance for ET-WDC

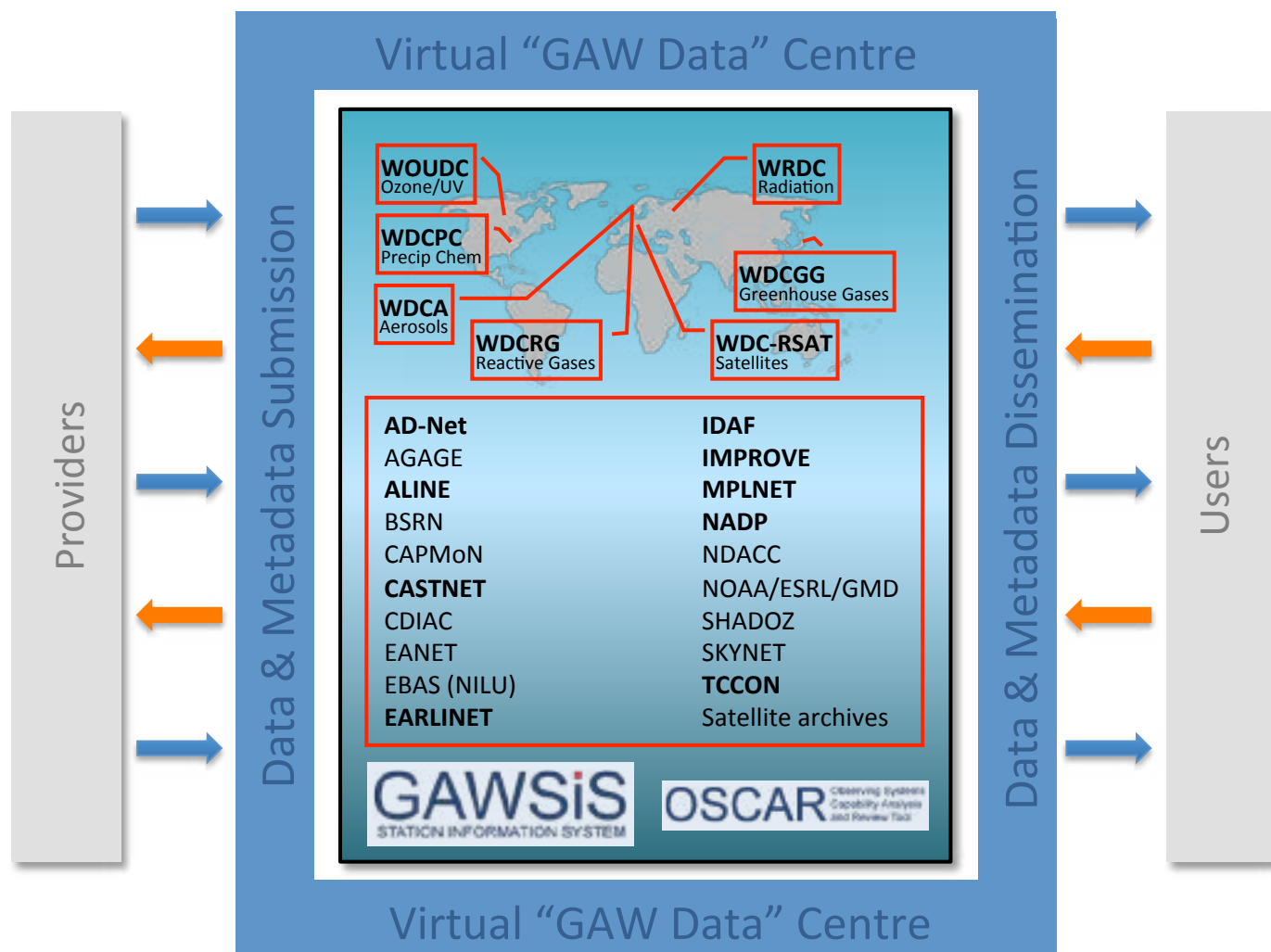
2a. Section 4.3

Jörg Klausen opened up for comments from the ET on section 4.3 in the GAW IP. This section deals with Data Management and contains a list of tasks and a mandate to establish a "Federated GAW Data Management Infrastructure".

The text from section 4.3 is given here:

These WDCs have been working on the harmonization of data submission and data access procedures and will continue these efforts with the joint vision of a GAW federated data management system that will allow fully interoperable access to all GAW data. Through advances in de-centralized web services and ongoing work on the standardization of metadata, the integration of individual GAW stations and contributing networks, as well as satellite data and modelling products shall be facilitated as illustrated in Figure 16. The GAW Station Information System (GAWSIS) will continue to play a central role for data discovery and detailed information on GAW stations and measurements of contributing networks. Through the new WMO OSCAR system and GAWSIS, the World Data Centres and data centres of contributing networks will become integral parts of the implementation of WIGOS, thus also facilitating data exchange across WMO programs and disciplines.

GAW will continue to liaise with other relevant actors (space agencies, environmental agencies, research networks) in order to harmonize metadata and data formats and thus facilitate the use of GAW and other data in various applications. GAW will interact with other WMO programs to ensure consistency of metadata for atmospheric composition information and work towards



Revised Figure 16 for the new GAW Implementation Plan.

providing a complete service chain consisting of the discovery and provision of research driven and operational observations, model data and service applications.

A revised version of Figure 16, naming more contributing networks, and focusing primarily on those that have signed an agreement with GAW, has been provided to the IP writing team. The revised figure is shown below.

2b Section 8.1: Station and Contributing Network Requirements

Markus Fiebig made a comment about the wording concerning Contributing Stations and Networks as given at the end of Section 8.1 in the GAW IP. This issue had also been discussed at the SAG Aerosol meeting in Korea the week before. The original text in the IP is given below.

Contributing stations are the stations operated as a part of the contributing network. To be

recognized as a contributing network, this network should sign an agreement with GAW. Contributing stations must satisfy the requirements of Regional, Global, Local or Mobile stations except for traceability (point 6) and data submission (point 8). Contributing stations must be operated according to the protocols of contributing networks and share their data through the contributing networks data centres. It is required nevertheless that metadata are provided through GAWSIS. Where appropriate and achievable, network standards shall be directly compared with WMO standards and scales and the network QA principles should be harmonized with the GAW QA principles.

The problem with this text is that it can be understood in such a way that contributing stations don't need to satisfy requirement for traceability and data submission. The point here is that contributing stations have to satisfy the requirements defined by the network they belong to, and these requirements might be somewhat different than the GAW requirements. The new text, which takes these possible misunderstandings into account is given below.

Contributing stations are stations operated as a part of a contributing network. To be recognized as a contributing network, this network should sign an agreement with GAW. Contributing stations must satisfy the requirements of Regional, Global, Local or Mobile stations except for direct traceability to WMO standards (point 6) and data submission to WMO Data Centres (point 8). Contributing stations must be operated according to the protocols of contributing networks, they must have internal traceability within the network and participate in comparisons to establish compatibility between the network standard and WMO standards or scales where such scales or standards exist. Contributing stations must share their data through the contributing networks data centres. It is required nevertheless that metadata are provided through GAWSIS. Where appropriate and achievable, network standards shall be directly compared with WMO standards and scales and the network QA principles should be harmonized with the GAW QA principles.

2c. Section 4.3: Data Management

Jörg Klausen presented the activities related to data management, as they are outlined in the new GAW IP. They are given here below:

A-DM-1. Establish a federated data management infrastructure including GAW Data Centres, data centres of Contributing networks, and GAWSIS.

A-DM-2. Improve open access to data and comprehensive metadata including calibration histories of ground-based, aircraft and satellite observations for the primary GAW variables.

A-DM-3. Implement a data and metadata management strategy that enables interoperable data discovery and access mechanisms and federates existing GAW data centres with data centres of contributing networks.

A-DM-4. Harmonize GAW data management activities with the WIGOS framework, in particular with regards to metadata documentation.

A-DM-5. Develop and promote support of data archiving and analysis centres that address the needs of applications and service delivery.

A-DM-6. Ensure that data collected and archived by WMO/GAW WDCs and archives of Contributing Networks are of known quality, adequate for their intended use and documented comprehensively.

2c Section 6.2.4: Terms of Reference (ToR) for WDCs and CDCs

Jörg Klausen presented the ToRs for ET-WDC as they appear in the new GAW IP.

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

WDC-2. 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.

WDC-3. 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.

WDC-4. 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.

WDC-5. Perform plausibility and consistency checks on submitted data, flag data problems, and provide feedback to the data providers, when necessary.

WDC-6. 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.

During the discussion of these ToRs it was agreed that the six paragraphs ought to be re-ordered like this: 1, 4, 5, 6, 2, 3. It was also agreed that the first paragraph should include data curation and versioning of data sets. This section will hence look like this.

WDC-1. Provide adequate archiving facilities for observational data for which GAW has global coordination responsibilities and ensure data curation and versioning of data sets.

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.

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

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.

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.

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.

3. WIGOS Metadata Standard (WMDS) and WIGOS Metadata Representation (WMDR)

Metadata of GAW stations are to be reported according to the WMDS (GAW IP Annex B Section 8.1 Bullet 8).

Jörg Klausen reminded the Team that we must help and start to implement WIGOS. The Federated Data Centre that we agreed on last year in Zürich is a part of this implementation. Jörg Klausen showed a presentation about WIGOS metadata standards. He showed the ten categories used within the WIGOS Metadata Standard version 1.0. WMO Members are obliged to comply with this for all internationally exchanged data. GAW data falls under this category, so we are obliged to follow this standard.

The WIGOS metadata standard will be implemented in three phases between now and around 2020.

The standard is semantic and descriptive and there is room for extension. Remote sensing is less well treated by the standards. The further development of the metadata standards is guided by the Task Team for WIGOS Metadata, co-chaired by Karl Monnik (BoM) and Jörg Klausen. Jörg Klausen gave the Team members the task to inform the SAGs where they are members about the importance of adherence to the WIGOS metadata standards.

He then carried on with an explanation of the physical representation of the metadata standard, the so-called schema. He showed the assembla web-page where one can find examples of how to implement WIGOS metadata. See here:

https://app.assembla.com/spaces/xmirepo/subversion/source/HEAD/tags/wigos_beta2

The schema is currently under review by a small group of experts, and they have until 15 July to complete this. The Members of ET-WDC are also welcome to comment on the schema. The schema will now be tested by prototyping solutions against it.

The schema are building rules for XML files. One can use other formats, but XML is the primary format for now.

Jörg showed Figure 16 from the GAW Implementation Plan (see above under 2a). This figure shows the idea behind the Federated Data Centre, a system where the complexities “behind the scenes” are hidden for the data providers and data users.

Tony Colavecchia wanted to know if the contributing networks are supposed to comply with the metadata standards of WIGOS and the answer to that is: “Yes, that is the long term goal.” We will start with the WMO/GAW World Data Centres. Once these centres know how to accomplish this, then the data centres of the contributing networks should follow suit.

MPLNet and NAPD are two examples of contributing networks that have signed agreements with GAW and although we still have a long way to go, technology is now ripe to accomplish this without reinventing the wheel. For example, we have OGC standards. We have the know-how in the Expert Team to help those data centres that are not yet up to speed on these new technologies.

Jörg Klausen showed diagram that shows the position of GAWSIS in the federated network and how data and metadata flow between the various centres.

JOIN is a tool for visualisation, operated by e.g. Martin Schultz, that can be used for quality control of data.

Martin Schultz mentioned Smart Monitoring where instruments automatically send both data and metadata to data centres without human intervention. Ocean buoys are one example where this technology of Smart Monitoring is implemented. Jörg Klausen replied that this is a development that our Expert Team does not control, but if such technology is being deployed for atmospheric composition monitoring, then we will have to adjust to this. Kipp & Zonen and other instrument manufacturers that are members of HMEI are actively developing sensors that also

provide metadata. Unfortunately, not many manufacturers of instruments measuring atmospheric composition are HMEI members, but we should promote and support a development in this direction. We cannot control this development, but we can involve instrument manufacturers in our development of requirements and hope that they will take this on board.

Geir Braathen mentioned the development of cheap sensors and UNEP's work in this direction. Martin Schultz was of the opinion that this is rather a topic to be discussed in the SAGs and the next RG SAG meeting already has this on the agenda. Jörg Klausen replied that this in the future also will be a topic to be considered in a federated data centre, namely how to discover and deal with crowd-sourced data.

Jörg Klausen then raised the issue of data exchange. Due to lack of time this topic was not discussed at this meeting but Jörg Klausen asked the members to inform themselves about Open Geospatial Consortium (OGC) standards (OGC-WCS, OGC-SWE) and Rasdaman (raster data manager) so that we can have a discussion about this at our next telecon in August.

Martin Schultz advocated strongly that we allow for the use of a very simple and standardised format, for example CSV, and that every data centre should allow for delivery of data in such a simplified format. The degree of expertise around the world is very variable, and by allowing stations to submit data in a simple format, we have a better chance of having data delivered to the data centres.

4. Information from Participants

Jörg Klausen asked for a "tour de table" where all the participants would have a chance to share the latest news from their respective data centres.

Markus Fiebig showed examples of NRT data that one can find at the EBAS data base at NILU. Such data can be found here: <http://ebas-nrt-showcase.nilu.no/>

Tony Colavecchia informed that WOUDC has carried out tests with standard WIGOS XML data and testing of conversion to CSV. EC is working with a vendor to create an open source tool to convert data into NetCDF. At the next meeting Tony would like to see a more comprehensive discussion of data exchange.

Kjetil Tørseth gave an update on the status of the new World Data Centre for Reactive Gases, which has been transferred from JMA to NILU (with the exception of data on carbon monoxide). There is good progress on setting up routines for data submission. A temporary web site has been put up with an update on the progress of transfer from JMA to NILU. Templates for data reporting have been developed in collaboration with the GAW SAG on Reactive Gases and also with the ACTRIS consortium. These templates will be gradually released in the new web system. Work is under way to develop programme independent descriptions of data submission. Officially, data submission will be open from the end of this year, but stations have already been invited to start data submission. Stations are encouraged to submit full historic time series. The reason for this is to take advantage of the recent progress in metadata reporting. This is also to avoid duplicates of the same data sets since some data stored at JMA also already exist in EBAS, but maybe as a different version. An NRT service for surface ozone from Arctic stations is being set up. This is being done in collaboration with the IASAO (International Arctic Systems For Observing The Atmosphere) activity (<http://www.esrl.noaa.gov/psd/iasoa/>). NILU is participating in a common nationally-funded research project with Øystein Godøy from met.no. This project will make a link between EBAS and other national research data resources through improved interoperability. There is also an EU proposal submitted that aims at setting up an Arctic observing system. Godøy is the leader of the data management part of this proposal and this can hopefully provide some funding for NILU. Jörg Klausen will ask Øystein Godøy to share information on national and Nordic data infrastructures. He also wants NILU to share the systems put in place.

Markus Fiebig agreed that the concept for the work flow can very well be shared. The source code, on the other hand is very much tailored to the computer system and local configuration and not so straightforward to port from one centre to another. The URL for the data submission tool is here:

<http://Ebas-submit-tool.nilu.no>

Markus Fiebig mentioned the European Horizon 2020 project ENVRI Plus, which brings together Environmental and Earth System Research Infrastructures, projects and networks together with technical specialist partners to create a more coherent, interdisciplinary and interoperable cluster of Environmental Research Infrastructures across Europe. Data citation services and Dols are also dealt with in this project. More information can be found here: <http://www.envriplus.eu/>

Anatoly Tsvetkov wanted to know if the NRT data visualisation service at NILU is based on commercial software. Markus Fiebig informed that the web page is written in .NET but that the data flow is runs with in-house developed code. The choice of tools to be used for such a system depends to a large degree on the in-house expertise.

Martin Schultz informed about the TOAR database, which contains 9600 ozone data series. The database can be accessed here: <https://Join.fz-juelich.de/access/db>

A login name and password are needed in order to gain access, but it is easy to create an account via the JOIN login page: <https://join.fz-juelich.de/accounts/login/>

Markus informed about the work to have NILU accepted as a DCPC in WIS and its connection to the GISC in Offenbach. This is now near completion and just bug fixes are remaining. The connection ought to work in early July some time.

Mikio Ueno was asked by Jörg Klausen about the progress towards web services that will replace the old csv file based exchange of metadata. The reply was that this transition will take another 6 months to a year to accomplish.

Jörg Klausen informed that the new GAWSIS is up and running. The old GAWSIS does not exist anymore. Currently there is a problem with metadata intake from the WDCs. It will still take a while to get this back up and running as before.

Geir Braathen informed about the GAW Symposium, which is planned to take place in Geneva from 10-13 April 2017.

5. Review of Action Items

Due to lack of time, there was no discussion of the old action items. This will be followed up at the next meeting. The new action items emerging from this meeting are given in the table below.

#	Responsible	Action to take	Deadline/status
1	All WDC Managers	Inform the SAGs where they are members about the importance of adherence to the WIGOS metadata standards and comment on the schema	15 July 2016
2	All	Members to inform themselves about Open Geospatial Consortium (OGC) standards (OGC-WCS, http://www.opengeospatial.org/standards/wcs , OGC-SWE, http://www.opengeospatial.org/standards/swecommon) and Rasdaman (raster data manager, www.rasdaman.com) so that we can have a discussion about this at our next telecon in August.	Before next telecon, which is 24 August.
3	Jörg Klausen	Ask Øystein Godøy to share information on national and Nordic data infrastructures.	ASAP
4	Markus Fiebig	Share concept for the [interoperability, web services] work flow employed at NILU.	ASAP

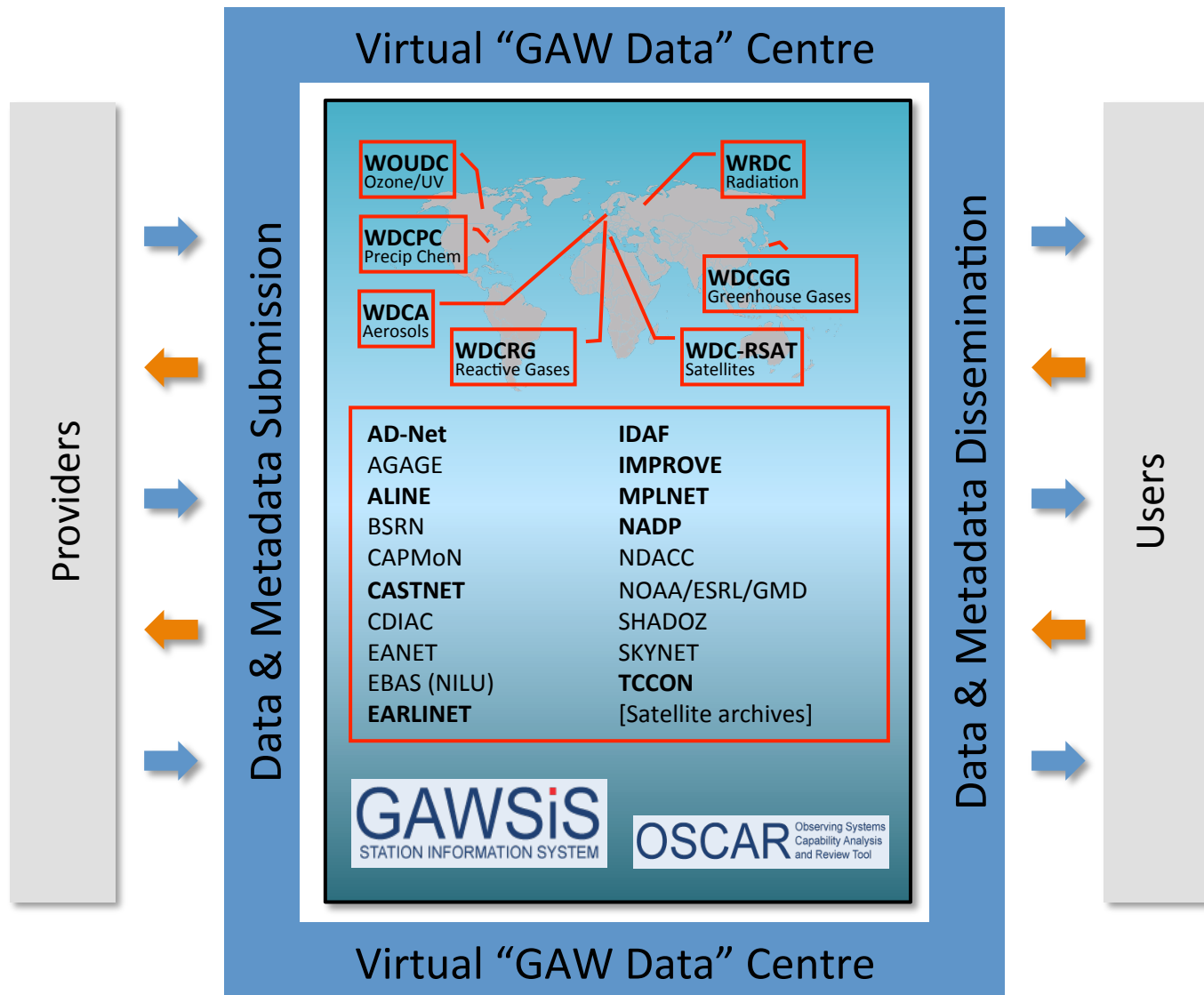
6. Next meeting

It was agreed to have a telecon towards the end of August and a physical meeting in late October or early November. Doodles have been set up for this. The telecon will take place on 24 August at 14:00 CET DST. The exact dates for the October meeting have not yet been defined. The next physical meeting could take place either at NILU or at met.no.

7. Appendix

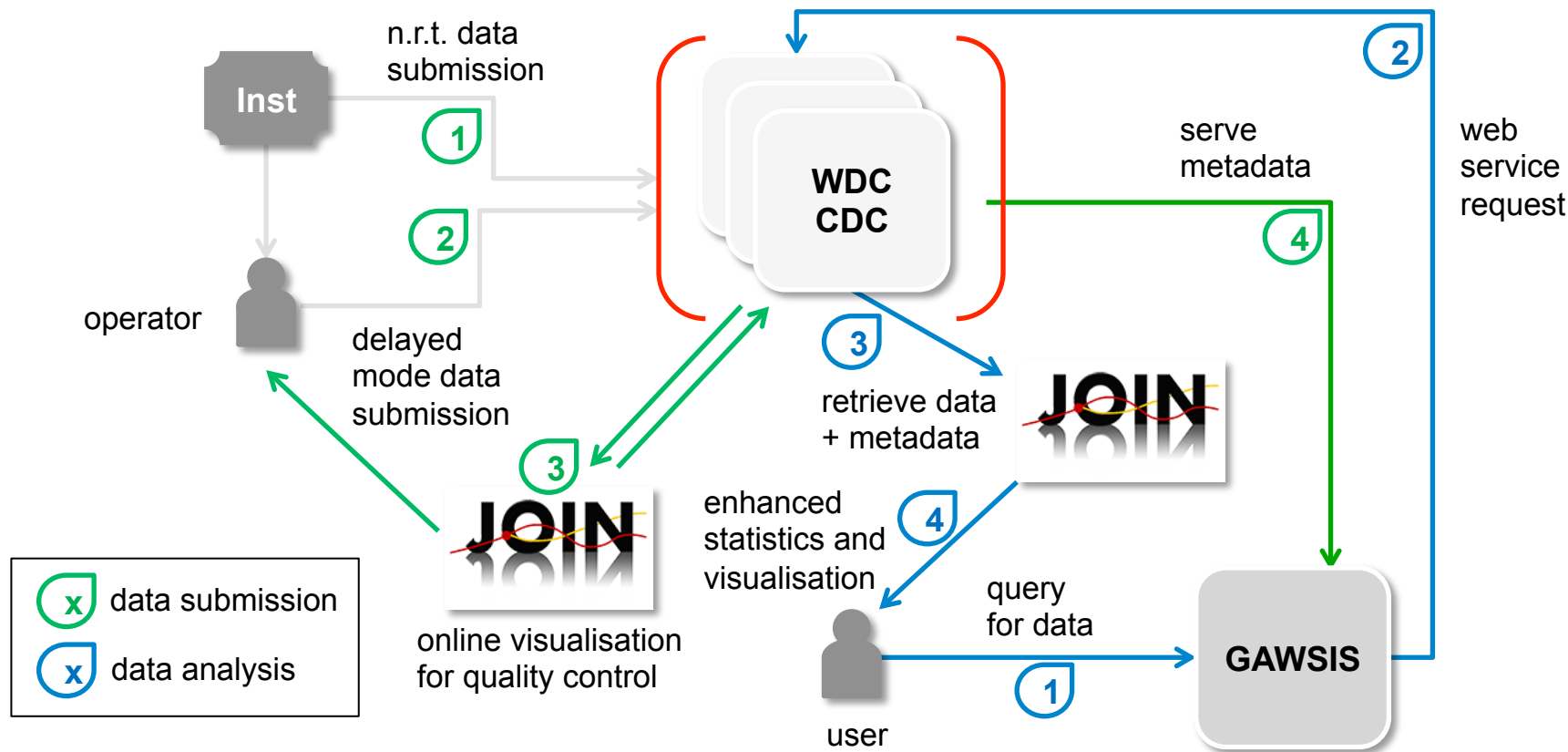
The presentation given by Jörg Klausen can be found on the next pages.

NextGAW

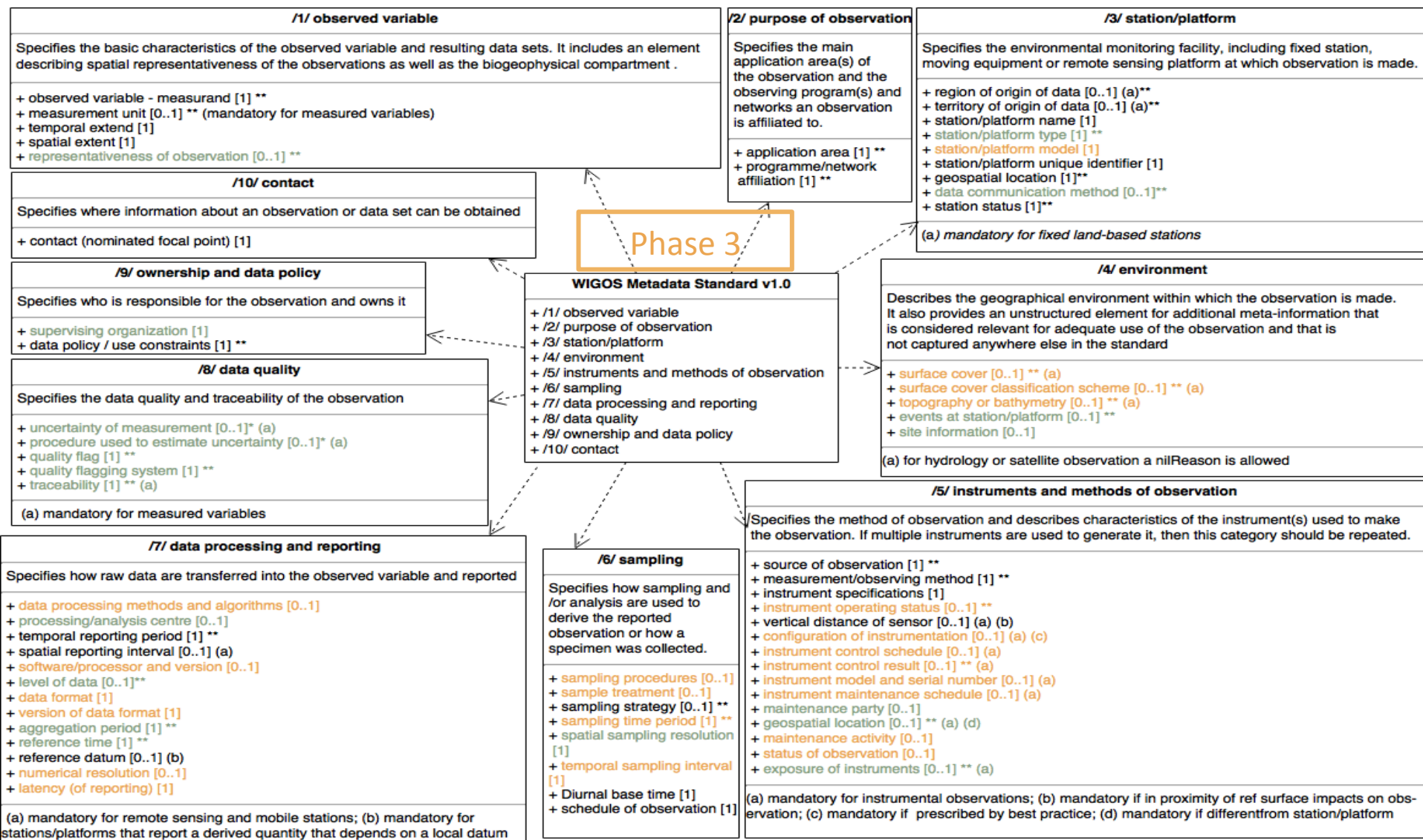




Federated GAW data architecture



WMDS



Data exchange

- Custom formats
- OGC-WCS
- OGC-SWE
- Rasdaman