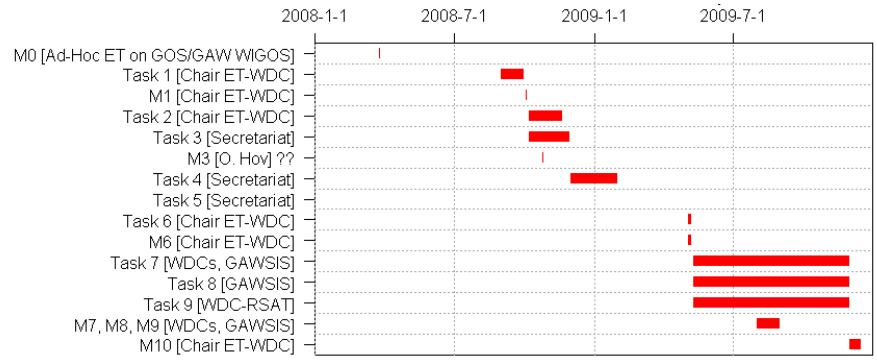
GOS-GAW Pilot Project (PP) Proposal

Improvement of Interoperability of GAW World Data Centres with WIS and Establishment of Prototype Services to Facilitate User Access to GAW Data

Project Name	Improvement of Interoperability of GAW World Data Centres with WIS and Establishment of Prototype Services to Facilitate User Access to GAW Data
Acronym	WDC-PP
Project Type	WIS-WIGOS Pilot
Project Status	Planning, parts presently under development (cf. project summary)
Project Overview	This pilot will improve the interoperability of GAW WDCs with WIS and other clients (e.g., individual users, satellite community) by exposing ISO 19115-compliant metadata representations describing the data archived at each WDC. These metadata will be used to establish a prototype client (human-machine interface) through which data and metadata of a given station that are archived at different WDCs can be combined and retrieved with a single client.
Project Aims	- Improve the possibility for machines to discover data archived at the individual WDCs - Improve the dissemination of data archived at the WDCs - Establish tailored services for the GAW global stations allowing them comfortable access to data and information products available from WDCs
Partners/Participants	Environment Canada (WOUDC), JMA (WDCGG), MGO (WRDC), JRC Ispra (WDCA), NOAA (WDCPC)[1], DLR (WDC-RSAT), Empa/ MeteoSwiss (GAWSIS), WMO GAW Secretariat
Project Costs	CHF 13,000, plus in-kind contributions
Funding Source(s)	WIGOS trust fund, Partners
Project Timescale	An indicative time table of activities is given below.  Tasks and Milestones: **M0** Request to develop a WIGOS GOS/GAW pilot proposal out of ET-WDC [Ad-Hoc ET on GOS/GAW WIGOS]

[1] WDCPC is presently in transition from its present host to another. The exact nature and primary location of WDCPC is not decided, however, NOAA currently maintains primary responsibility.

	Task 1 Development of draft proposal [Chair ET-WDC] **M1** Presentation of draft proposal to ET-WDC in SPb [Chair ET-WDC] **Task 2** Iteration and finalisation of proposal, submission to Secretariat [Chair ET-WDC] **Task 3** Inquire about WMO internal funding sources [Secretariat] **M3** Mentioning of proposal at CAS Mgt. Group meeting (Geneva) [L.Jalkanen] **Task 4** Submission of proposal to EC WG WIGOS/WIS [Secretariat] **Task 5** Request WDCs to participate [Secretariat] **Task 6** Presentation of project at JSC meeting (Geneva) [Chair ET-WDC] **M6** Kick-off meeting during GAW2009 (Geneva), finalisation of Timeline/MS [Chair ET-WDC] **Task 7** Implementation of ISO-compliant metadata [WDCs, GAWSIS] **Task 8** Implementation of data-extraction/repackaging prototype [GAWSIS] **Task 9** Implementation of enhanced services to global GAW stations [WDC-RSAT] **M7-9** Intermediate review of progress and further training [WDCs, GAWSIS] **M10** Prepare and submit progress report for EC WG WIGOS/WIS [Chair ET-WDC]
Expected Key Deliverables	- WIS-compliant representations of metadata for each data set archived at the WDCs. - Definition of dissemination pathways for these metadata. - A web-interface client demonstrating the interoperability of GAW WDCs. - Demonstration of data analysis tools/services for GAW global stations in support of research activities and especially for synergetic data utilization.
Project Links	- http://www.wmo.int/wdc-pp/ should be set up to host information on the project.
Project Summary	**Background** The system of GAW WDCs has developed significantly since its inception. WOUDC was established in 1961, WRDC in 1964, the other existing WDCs later in the 1990s. Except for WDCPC, all centres are well established with a long-term perspective. They are also recognized by WMO, the GAW community, and others as being the primary archives for GAW- and related data. In addition to being a GAW WDC, WDC-RSAT is also recognized as an ICSU WDC. Since 2001, GAWSIS has been integrating metadata from GAW stations (global, regional, contributing) in a web-accessible database system, thereby providing a common data discovery mechanism for the GAW system. Since 2005, some WDCs have established metadata-sharing mechanisms with GAWSIS based on flat ASCII files. With WOUDC, WDCGG, this mechanism ensures weekly updates of the contents of GAWSIS. With WRDC and WDCA, such a mechanism is presently being implemented. With WDCPC and WDC-RSAT, no such mechanisms are currently in place. WOUDC, WDCGG and WDCA have implemented open, easy access to archived data through web interfaces or ftp. WRDC provides visualization of the data on a per-user basis, but the data themselves are difficult to access. WDCPC data are presently not easily available, and most precipitation data are archived by regional

	data centres that are not (yet) an integrated part of the GAW WDC system. WDC-RSAT provides a plethora of information on a wide range of satellites, including visualization of some data sets. It offers free access to data sets. **Motivation** The WMO Information System (WIS) aims at improving the discovery and retrieval (DAR), and the rapid exchange of data collected primarily by National Meteorological and Hydrological Agencies (NMHSs), but also other sources (EPAs and other government agencies, universities, etc.). Parts of WIS are presently being implemented by Members, and the meteorological part of WIS is expected to become operational in 2009-10. Much of this concerns near-real-time or rapid-delivery data. Presently, the validated data archived and maintained by the GAW WDCs are only mentioned in the WIS implementation plan, but they are not considered to be of highest priority and there is no concerted effort being made for their integration. Discovery of data archived at the WDC is facilitated through the individual archives and – in an all-encompassing manner – through GAWSIS. Data retrieval is only possible through the individual WDCs at present. Users are directed to the archive for a given variable through GAWSIS, but there is presently no way of extracting data for several parameters, let alone distributed at various archives, at once. **Improving Interoperability and WIS-compliance** Interoperability of systems can only be achieved through adherence to strict standards for encoding and communicating information. 1) It is proposed and the aim of this pilot project to ***implement and expose ISO/WIS-compliant representations of metadata for the data archived in the WDCs***. These representations can be established centrally from GAWSIS, or separately at each WDC (in which case GAWSIS would simply be one of the clients of such information). ISO 19115 is a set of standards for the description of geographical [2] metadata. This standard is widely accepted, and even a directive (INSPIRE) in the European Union. The WMO Inter-Programme Expert Team for Metadata implementation (IPET-MI) has recommended this standard for WIS, and it has been endorsed by WMO. The WMO Core Metadata Profile builds upon ISO 19115 and is fully compliant. ISO 19139 is the standard for implementing ISO 19115 in XML (extensible mark-up language). 2) It is further proposed to ***develop a prototype client to facilitate use of GAW data across multiple WDCs***. The purpose of such a client is to demonstrate a user-friendly means of integrated retrieval of data from the GAW WDCs. Conceptually, the client will focus on providing different data sets in a single file as well as the associated metadata for a given location but across various variables. Such data sets are considered useful for cross-validation of observations, model validation, correlation analysis, etc. It will highlight the integrated nature of the GAW WDC system, thereby strengthening each of its components. Technically, it will be realized as a web-based client. Upon user request, data sets will be

[2] ‚Geographical' in this context is everything for which space and time coordinates are relevant.

	identified/found using information stored in GAWSIS and retrieved from the respective archives. Data sets will then be provided in a single file without altering the data before being served to the user. Metadata associated with the data will be served to the user along with the data. 3) It is also proposed to ***establish specific pilot services of WDC-RSAT for the GAW global stations***. The connection of GAW Global Stations to the WDC-RSAT will allow access to information on the current condition of the atmosphere on global, continental and regional scale and will, in some cases, be available in near real time. Further on, it is proposed to establish access to selected numerical atmospheric models for the better interpretation of data. Several services will be realised for the assistance of research activities and especially for synergetic data utilization such as computing-on-demand, video-on-demand or GIS-functionalities. **Achievements** Presently, ISO-compliant representations of the metadata of all data sets registered in GAWSIS are being established. A preliminary version of these files is available at http://gaw.empa.ch/gawsis/xml[3]. A very preliminary version of a data retrieval user interface of the prototype client is presented at http://gaw.empa.ch/gawdap.
Date of Last Update	11 May 2009
Contact Person	Dr. Jörg Klausen, Chair ET-GAW WDCs joerg.klausen@empa.ch

[3] These representations are validated (i.e., they are technically valid ISO-compliant metadata representations), but are currently not always complete and will be refined before being published to WIS. The purpose of coding GAWSIS in XML is to gain experience with the ISO standards, to identify areas where the metadata obtained from the WDCs needs to be improved, and to identify areas in which WMO/GAW needs to interact with IPET-MI (and ISO).