

GOS-GAW Pilot Project (PP) Proposal

Improvement of Interoperability of GAW World Data Centres with WIS and Establishment of a GAW Data Access Portal

Project Name	Improvement of Interoperability of GAW World Data Centres with WIS and Establishment of a prototype GAW Data Access Portal
Acronym	II-WDC-GAWDAP
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 GAW Data Access Portal (human-machine interface) through which data and metadata of a given station that are archived at different WDCs can be retrieved and combined into a single dataset.
Project Aims	<ul style="list-style-type: none"> - Improve the possibility for machines to discover data archived at the individual WDCs - Improve the dissemination of data archived at the WDCs - Combine data archived at various WDCs for user-friendly download
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	t.b.d.
Funding Source(s)	WIGOS trust fund, Partners
Project Timescale	<ul style="list-style-type: none"> - Development of implementation plan, definition of milestones by April 2009. - Draft implementation of ISO-compliant metadata representations by end of 2009. - Draft implementation of GAWDAP by end of 2009. - Report progress to EC WG on WIGOS/WIS at the end of 2009.
Expected Key Deliverables	<ul style="list-style-type: none"> - WIS-compliant representations of metadata for each data set archived at the WDCs. - Definition of dissemination pathways for these metadata. - A web-interface (GAWDAP) demonstrating the interoperability of GAW WDCs.
Project Links	<ul style="list-style-type: none"> - http://www.wmo.int/ii-wdc-gawdap/ should be set up to host information on the project.

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.

<p>Project Summary</p>	<p>Background</p> <p>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.</p> <p>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.</p> <p>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 appears that most data sets are only available for registered users.</p> <p>Motivation</p> <p>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.</p> <p>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.</p>
-------------------------------	---

	<p>Improving Interoperability and WIS-compliance</p> <p>Interoperability of systems can only be achieved through adherence to strict standards for encoding and communicating information.</p> <p>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.</p> <p>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 geographical² 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.</p> <p>ISO 19139 is the standard for implementing ISO 19115 in XML (extensible mark-up language).</p> <p>2) It is further proposed to develop a prototype-client of this information to facilitate extraction of GAW data across multiple WDCs (GAWDAP).</p> <p>The purpose of such a system is primarily to provide a user-friendly means for integrated retrieval of data from the GAW WDCs. Conceptionally, GAWDAP will focus on providing merged data sets in two dimensions (time, value) for a given location but across various variables. Such data sets are considered useful for cross-validation of observations, model validation, correlation analysis, etc. GAWDAP will also highlight the integrated nature of the GAW WDC system, thereby strengthening each of its components. Technically, GAWDAP will be realized as a web-based client-server application. Upon user request, data sets will be identified using information stored in GAWSIS and retrieved from the respective archive. Data sets (and associated metadata) will then be merged into one two-dimensional table before being served to the user.</p> <p>Achievements</p> <p>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. A very preliminary version of a data retrieval user interface of GAWDAP is located at http://gaw.empa.ch/gawdap.</p>
Date of Last Update	20 Sep 2008
Contact Person	Dr. Jörg Klausen, Chair ET-GAW WDCs

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