Information and Services WG

Portal Team
Terminology Team
Website and Outreach Team
GCW Portal
Data Catalogue
Status and plans

Øystein Godøy
Purpose

- To provide an overview of datasets relevant to GCW
- To provide access to datasets wherever possible
  - Real time data streams
  - Archive access
- To connect GCW with
  - WMO Information System
  - WIGOS
- Distributed Data Management
  - Metadata driven
  - Currently not hosting data
    - Possible, but would potentially require some additional funding
Status (1)

- Manually initiated metadata ingestion, automated harvest
  - National Snow and Ice Data Centre
  - British Antarctic Survey
  - NPI
  - CryoClim
  - CHINARE
  - NIPR
  - CryoNet
    - Davos
    - Sonnblick
      - Via PANGAEA

- Have tested against
  - CCIN
    - Semantic issues
  - WGMS
    - Standardisation issues
  - PANGAEA
    - Granularity issues
    - Further discussions planned
  - IASOA
    - Semantic issues
  - AARI
    - Sea Ice Charts
  - FMI
    - Interfaces issues
  - EUMETSAT
    - Interfaces have been updated, need to recheck
  - GTN-P
    - Standardisation issues
Status (2)

• Discussions with
  • Polish National Polar Data Centre
    • Institute of Oceanology PAS in Sopot,
    • Institute of Geophysics PAS in Warsaw
    • University of Silesia in Sosnowiec
      – Stalled
  • ECDS (SE)
    – New interfaces available, recheck
Status (3)

• Automatic filtering of harvested records
  • Where no dedicated “cryosphere” or “GCW” sets
  • Filters on keywords related to cryosphere
    – Depends on semantic framework

• Manually supervised ingestion of harvested metadata
  • Formats used so far GCMD DIF, ISO19115 (various flavours), FGDC, ad hoc formats, ...
    – Formats or structure is not the issue, semantics are...

• Metadata are exposed through OAI-PMH
  • GCMD DIF and ISO19115
  • WMO validation in place
    – WIS, not WIGOS, need further development
Issues that require attention (1)

• Filtering of harvested metadata
  • Usage of controlled vocabularies
    – Machine readable implementations are required
    – Need software development in system to benefit from crosswalks

• Brokering and cleaning/validation of harvested metadata
  • For WIS compliance
  • WMO controlled vocabularies (and consistency of these)
    – codes.wmo.int is yet not populated enough

• Duplication of metadata
  • Which metadata to expose through machine interfaces
  • How to interact with WIGOS/WIS/GTS
Issues that require attention (2)

- Development of manuals and guidelines
  - Interoperability, Operations, SLAs
  - How to ensure uptake in community?
- Tagging of data?
  - How to identify contributing data centres?
  - CryoNet stations/sites
  - Approval of datasets
- Interaction with contributing communities
  - User communities (e.g. CryoNet)
  - WMO initiatives and data centres
  - Scientific communities through RDA, ICSU (WDS/CODATA), SAON/IASC etc
Issues that require attention (3)

- Development, propagation and utilisation of controlled vocabularies
- Development resources
  - Exploiting synergies with other projects and initiatives
  - Examining funding opportunities
    - Nationally and at the European level
- Ambition level for higher order services
  - Visualisation, transformation etc
Plans - Catalogue software

- Reimplementation of human interfaces
  - Ongoing, CMS solution based on Drupal and web services. Will not be available until 2016
- Reimplementation of basket
  - Ongoing, estimated 2016
- Reimplementation of metadata interfaces
  - Ongoing, estimated 2016
- Semantic interoperability capability
- Interfaces for WIGOS metadata
Tasks identified in work plan

- Integrate relevant data centres
  - Specifically with CryoNet
    - Dedicated effort with Davos
    - Discussions with Sonnblick and Sodankylä
- Establish GCW Catalogue Interoperability Group
  - Related to development of interoperability guidelines
- Develop catalogue interoperability guidelines
  - Preliminary version for discussion
- Develop a GCW Catalogue Operations Manual
  - Preliminary version for discussion
- Software development
  - Continuously
- Establish the GCW catalogue as a WIS DCPC
  - Delayed, need to align timelines and resources internally at METNO
- Analyse the problem of duplication of metadata through harvesting
- Develop bilateral agreements with contributing data centres
- Adapt harvested metadata to WMO requirements
  - Need a pragmatic approach and to guide communities on WIGOS and WIS as well as providing feedback to WIS/WIGOS from the community
- Connect GCW catalogue to WMO GTS for datasets generated by non-WMO members
  - Depends on interaction/interoperability with data centres
The Combined SAON/IASC Arctic Data Committee
Status

- Established autumn 2014
- First meeting in Potsdam, November 2014
- Second meeting in Waterloo, October 2015
- WMO representative
- Coordinated with SCADM
- Coordinating interoperability and metadata exchange/development