

Data Integration and Analysis Systems

DA-09-02a Task Update

Rick Lawford and Toshio Koike

Presented at the
ADC meeting
Melbourne, Australia

September, 2009

DA-09-02a Task Report

(September, 2009)

In the winter of 2008-2009 a survey questionnaire was sent to 11 large data centres.

The questionnaire asked for information related to:

1. Name of System or Project:
2. Leading Organization(s):
3. Associated Organization(s):
4. Purposes of the System or Projects:
5. Relation to GEOSS SBAs:
6. Involvement in GEO Tasks (Name Tasks):
7. Brief Description of Archived Data:
8. Brief Description of Data Storage Capabilities:
9. Brief Description of Data Interoperability Function:
10. Brief Description of Data Integration and Analysis Capabilities:
11. Level of use and characteristics of users:
12. [Contributions to GEO Registries](#)
13. Homepage Address:
14. Contact:

DA-09-02a Task Report

(September, 2009)

Nine of the eleven centres responded with completed questionnaires (see Table 1)

These responses have been analyzed and summarized in tables.

The centres have been given a chance to make adjustments to their responses based on the responses of all of the centres. (This step was needed to ensure that the centres all had a common understanding of the nature of the questions.)

A written report is being developed and will be submitted to the ADC when it has been completed.

Any inconsistencies in the questionnaire will be resolved before the Phase II survey is distributed.

Centre	Affiliation	Web Site	Contact
World Data Center for Glaciology, Boulder	National Snow and Ice Data Center (NSIDC)	http://nsidc.org and http://nsidc.org/wdc	rbarry@nsidc.org fetterer@nsidc.org
World Data Centre for Glaciology and Geocryology at Lanzhou	Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences	http://wdcdgg.westgis.ac.cn	lixin@lzb.ac.cn
World Data Center Climate (WDCC)	Max Planck Institut for Meteorology (MPI-M) & Deutsches Klimarechenzentrum (DKRZ)	www.wdc-climate.de	data@dkrz.de
Global Runoff Data Centre (GRDC)	Federal Institute of Hydrology (BfG) Am Mainzer Tor 1 56068 Koblenz, Germany	http://grdc.bafg.de	grdc@bafg.de
The Global Observing Systems Information Center (GOSIC)	National Oceanic and Atmospheric Administration (NOAA) National Climatic Data Center's (NCDC) US Global Climate Observing System (GCOS) program.	http://gosis.org	christina.lief@noaa.gov

Centre	Affiliation	Web Site	Contact
Goddard Interactive Online Visualization AND aNalysis Infrastructure (GIOVANI)	GES-DISC, a NASA center program operating under the Global Change Data Center (GCDC) at the Goddard Space Flight Center (GSFC)	http://giovanni.gsfc.nasa.gov/ http://disc.sci.gsfc.nasa.gov/	Gregory.leptoukh@nasa.gov
NASA Goddard DAAC	GES-DISC, a NASA center program operating under the Global Change Data Center (GCDC) at the Goddard Space Flight Center (GSFC)	http://disc.gsfc.nasa.gov/	Gregory.leptoukh@nasa.gov
Data Integration and Analysis System (DIAS)	The University of Tokyo	http://www.editoria.u-tokyo.ac.jp/dias/ (Japanese only now)	tkoike@hydra.t.u-tokyo.ac.jp
GENESI-DR: Ground European Network for Earth Science Interoperations – Digital Repositories	European Space Agency, Earth Observation Programmes Directorate, Frascati, Italy	http://www.genesi-dr.eu/	luigi.fusco@esa.int joost.van.bemmelen@esa.int

Brief description of Data Storage Capabilities

	theme	storage	Data base
WDC Boulder	Glaciology	32TB(RAID)+ 40TB (MAID) 300TB(Tape Storage) +75 TB disk array	<ul style="list-style-type: none"> •Sybase Enterprise Server relational database •MySQL database server
WDC Lanzhou	Glaciology & Geocryology	10TB (RAID)	<ul style="list-style-type: none"> •Catalog metadata database
WDC Hamburg	Climate	Growth of DB system of 100 TB/year will increase to 1 PB/year after 2009	
GRDC: Germany	River Runoff	sufficient storage capabilities	commercial database
GOSIC NOAA/USA	Earth obs. data portal	No storage system but link function	
GIOVANI NASA/USA	Web-based application	20 TB	
GES-DISC NASA/USA	integration and analysis	500 TB Peak daily numbers:3.3 TB for processing 1.7 TB for archiving	
DIAS Japan	integration and analysis	700TB (disk)	
GENESI-DR ESA	e-science infrastructure		OGC Web <i>Services</i> , GridFTP, HTTPS, OPeNDAP, BitTorrent for file-based DRs.

Types of Data Users

	Project Participants	Academic Users	Government research	Government general	Commercial	Operators	Level of website use	Data Base Downloads	Fully Open Access	Approval before access
	Project Participants	Academic	Gvt Research	Gvt Operations	Commercial	Operators	Website Use	Data Base Downloads	Fully Open Access	App before access
NSIDC/ WDC Boulder	X	X	X	X	X		3.2M/y	7.2 million files/yr*	X	
WDC Lanzhou		X	X	X						X
WDC Hamburg		X	X							
GRDC Germany	X	X	X	X				120/y		X
GOSIC NOAA/NCDC/US	X	X	X	X	X			2500/d	X	
GIOVANI NASA/US		X	X	X	X		2000/m		X	
GES-DIS NASA/US		X	X	X	X		150K/m		X	
DIAS Japan	X	X								X
GENESI-DR ESA/EC	X	X	X	X		X		Restricted by policies		X

Next Steps

A Phase II SBA-specific Survey will be undertaken in the last half of 2009 of major Water Cycle Data Centres (including international, regional and national data centres). This should lead to the identification of possible alliances between centres with shared goals and capabilities, and similar stages of development.

A joint workshop between task DA-09-02a on Alliances and DA-09-01b on Harmonization will be held in Washington DC on November 11 and 12, 2009. ADC members are invited to attend and are asked to contact Rick Lawford or Toshio Koike if they plan to attend.