



SEO Data Policy Study and GCI Registration Study

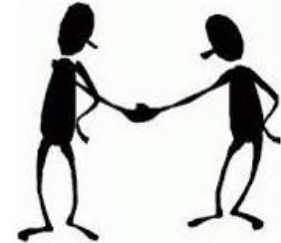
Brian Killough (NASA, SEO)

September 24-28, 2012
WGISS-34 and WGCV-35 Meeting
Hyderabad, India



CEOS Action: (ID-01-C1_1, Due: 31-Oct-2012)

- Conduct an assessment of data sharing policies for current CEOS missions to determine the nature of any restrictions.



Data Access Categories

- **OPEN** -- No restrictions other than simple registration (< 2 days).
- **RESTRICTED** -- Some restrictions to data access including data fees or a proposal process.
- **NO ACCESS** – Data not available for public access and/or no public access source found.

Status

- The first release of this study (Phase-1) was completed on August 30. Thanks to the AMA team for their work !



Next Steps

- Review results with SIT (here) and WGISS (Sept meeting).
- Complete Final Report by CEOS Plenary.
- Work with WGISS to define the way forward for 2013.



- Considered **CURRENT** mission-instrument combinations (**257 total**) out of 100 CEOS missions from the 2012 MIM database.
- Data policies were based on the “majority” of data products from any mission-instrument combination and may not reflect the same data policy for all data products.
- Data Distribution policies were not examined in the study. It is anticipated that some distribution restrictions may exist for products that are OPEN data access. Crediting or referencing the source would be a minor restriction in distribution.
- Laser Reflectors, Data Collection Systems, and Search and Rescue instruments (52 total) were not considered since they are not typical Earth science.
- Decommissioned missions (9 total) were removed (i.e., Envisat, DMSP-F14, GOES-11, SORCE).





Out of 257 current CEOS mission-instrument combinations ...

- **76% - OPEN Data Access**
- 15% - Restricted Data Access
(8% fee, 4% unknown, 3% proposals)
- 9% - No Data Access (see below)



Summary of “No Data Access” findings

- AISSAT-1 (NSC) – No access point found
- CARTOSAT 2A/2B (ISRO) – No access points found. Other similar missions found
- Elektro-L N1 (ROSKOSMOS) – FTP seems to be available, but no access point found
- IMS-1 (ISRO) – May be available through NRSC, but no public access points found
- NMP EO-1 (NASA) – No public access to LEISA-AC instrument data in archive.
- RASAT (TUBITAK) – Contacted TUBITAK. Policy unknown. No access found.
- Resurs DK-1 (ROSKOSMOS) – No public access point found
- SAC-C and SAC-D/Aquarius (CONAE+) – No public access points found for 10 instruments. SAC-C HRTC, SAC-C MMRS, SAC-D Aquarius instruments are all OPEN data access.



CEOS Action: (IN-05-C1_1, Due: 31-Oct-2012)

- Conduct an assessment of CEOS data registration in the GCI to support the GEOSS architecture and interoperability principles. Consider Data-CORE, Data Portals and Broker Services (CWIC).



Definitions

- GEOSS Data Collection of Open Resources for Everyone (**DataCORE**) is a pool of datasets, contributed by the GEO community on the basis of full and open exchange (at no more than the cost of reproduction and distribution) and unrestricted access.
- **Client Portals** are Search / Access tools that connect users to data collections and/or products.
- The CEOS WGISS Integrated Catalog (**CWIC**) provides an interface between client portals and partner data servers

Status

- The first release of this study (Phase-1) was completed Aug 30. Thanks to the AMA team for their work !



Next Steps

- Review results with SIT (here) and WGISS (Sept meeting).
- Complete Final Report by CEOS Plenary.
- Work with GEO and WGISS to define the way forward for 2013.



Out of 257 current CEOS mission-instrument combinations ...

- **35%** produce datasets that are considered **DataCORE**

Some OPEN datasets that are not DataCORE come from these missions:

*FY (NRSCC), DMSP (NOAA, some), GOES (NOAA, some), INSAT-3A (ISRO). **Could these be DataCORE?***

- **44%** have data collections that are accessible through the **CEOS IDN**

*The CEOS International Directory Network (IDN) is the CEOS endorsed location for information about Earth observation data collections. The IDN (idn.ceos.org) is connected to **55%** of the CEOS OPEN datasets. Queries result in Directory Interchange Format (DIF) results from 21,000+ data collections.*

- **27%** utilize **CWIC** for access to data products

*CWIC is utilized by 70 mission-instrument combinations from **28 missions***

NOAA-CLASS, SST Constellation (GHRSSST), and LSI Constellation Portal integration is in progress.

What missions could the **IDN** target to add to their database?

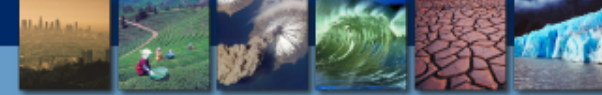
Megha-Tropiques, FY-3B, HY-2A

What missions could **CWIC** target to become partners?

Megha-Tropiques, FY-3A/B, HY-2A, SAC-C/D

Additional missions will also be considered





- **Client Portals** are online tools to Search and Access.
- This study found **30** different portals were used for data access. **19** of those portals were used for more than one mission-instrument combination.

The most common portals were:

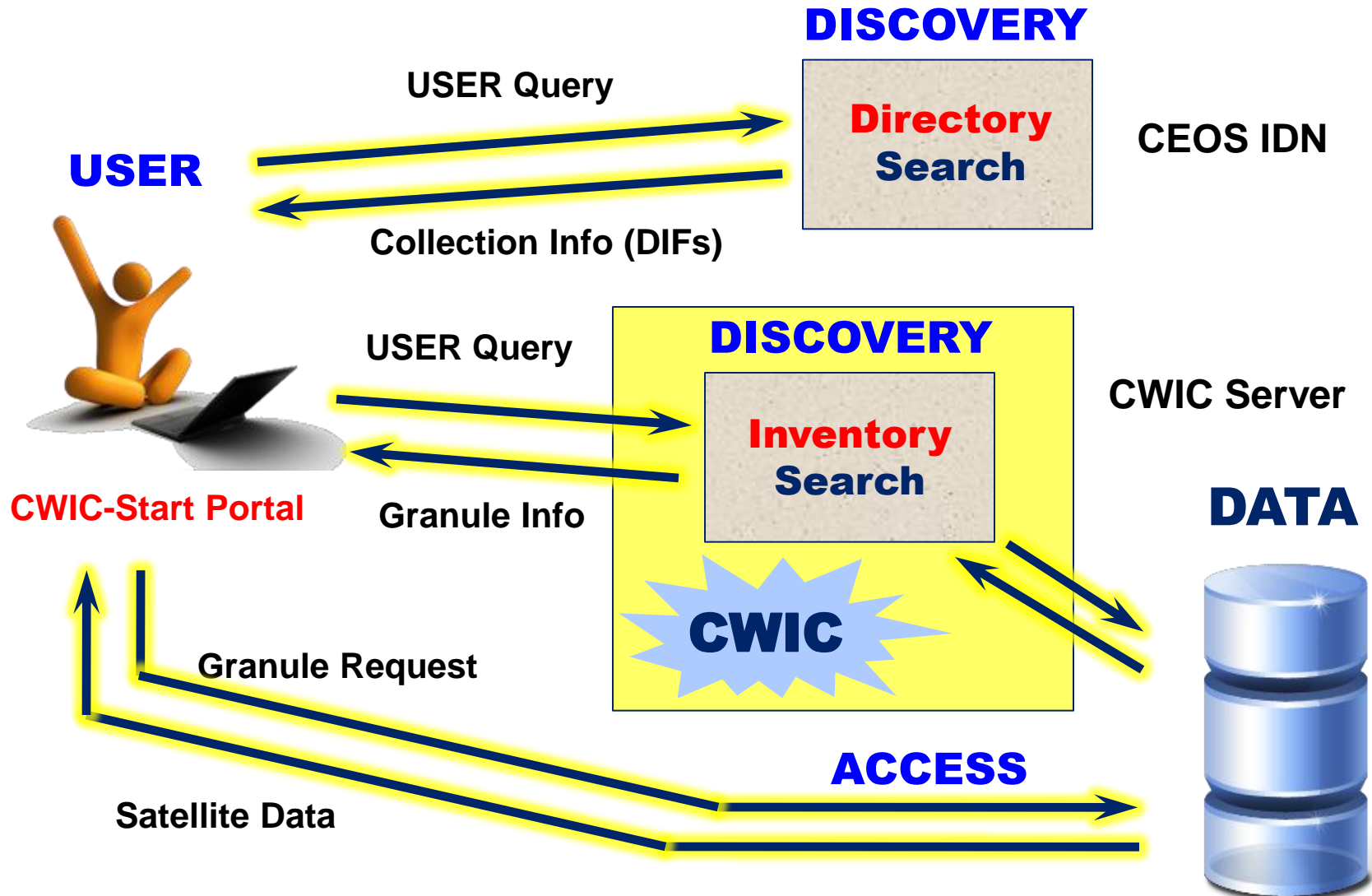
CLASS, FENGYUN, EOLI, NGDC, REVERB, PODAAC

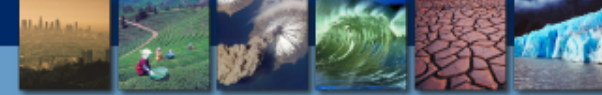
- These client portals are likely the best initial connection for users to find most Earth observation data products for a given mission-instrument.

Should these portal links be included in the MIM for each instrument?

- NASA is developing a new Client Portal, called **CWIC-Start**, for direct connection to the IDN and the CWIC Server.

CEOS Data Discovery and Access





Conducted multiple telecons with GEO reps to discuss the GCI ...

- **Doug Nebert** (USGS) and **Mirko Albani** (ESA)

Strong support of the CEOS IDN and CWIC.

There is no easy way to get a list of GCI registered components and services.



- **Stefano Nativi** (CNR)

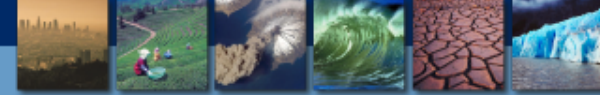
EuroGEOSS project (started in 2009) has created the Discovery and Access Broker (DAB). The DAB will be used by the GEO Portal to query the IDN and many other capacities (OneGeology, WMO Information System or WIS, etc). A demo of DAB will be presented at the GEO Plenary. It is anticipated that both CWIC and HMA are functional with the DAB by that time.



- Recent discussions and emails with Mirko Albani have uncovered some more details about GMES Services and HMA.
- Access to data in the GMES portal is only for GMES services approved by the EU. According to our Data Policy study, the access would be “Restricted” due to the approval process. GMES service approval by the EC requires that the user must be part of a project with an SP7 Space Theme (see cordis.europa.eu)
- Access to catalogs or datasets utilizes the HMA protocol for all of the missions listed on the right. Need to search by instrument to compare with data similar to the SEO data study.
- There are some missions that use CWIC through the IDN and HMA through GMES. Mirko believes it would be easy to have CWIC speak to HMA or CWIC to interface with HMA-based ESA system like FedEO.

GMES Services ~40 Missions and Instruments using HMA

Demios-1,, NigeriaSat-1&2,
 Beijing, COSMO-SkyMed,
 IRS-P5 Cartosat-1,
 LISS-III&IV,
 IRS-1C/1D,
 IRS-P6 1&2 Resourcesat,
 IKONOS archived,
 TerraSAR-X, Tandem-X,
 RadarSat-2, RapidEye,
 Pleiades, Formosat-2,
 Spot-4&5, Landsat ETM+
 GeoEye-1, GOCE,
 Envisat, GOSAT,
 PROBA, ERS archive,
 MODIS, ALOS archive
 Quickbird, Worldview 1&2,
 Parasol, Meteosat,
 MSG, MetOp,
 Scisat, Sentinels



- **Ken Casey (NOAA)**

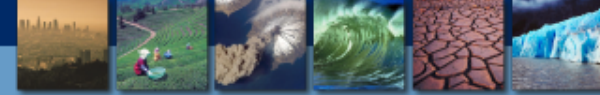
(1) Distinguish between "OPEN-simple registration" and "OPEN-no registration". Interoperable data access is often interrupted by even simple registration systems so it would be good to know the differences.

(2) 77% OPEN data access is misleading. Many L2 and L3 datasets are OPEN, but the L0 and/or L1 are not openly available. With a future focus on Climate Data Records and ECVs, this is important to success.

- **Ivan Petiteville (ESA) ...**

What is missing, in most cases, is the access to the products themselves.

Yonsook Enloe commented that inventories are available through CWIC with access to data download and sometimes via a URL. Petiteville suggests we need to work more on getting direct access to the actual data products.



- DO NOT register datasets into the GCI. Focus on expanding the content of the IDN as the main source of CEOS dataset information (via DIFs). Add tags for CWIC and DataCORE.
- Continue the development of CWIC-Start and include MIM nomenclature.
- Add MIM keywords to the IDN to support CWIC-Start queries.
- Expand the number of CWIC partners using the results of the SEO Study to target potential agencies and missions.
- Continue dialogue with GEO and its use of IDN and CWIC.
- Continue discussions with ESA about the use of HMA and integration with CWIC.
- Continue the CEOS Data Policy study ...
 - (1) Refine and improve existing data for current missions,
 - (2) Add new data from past missions ... could be a large number,
 - (3) Identify the primary Client Portals for all OPEN data combinations,
 - (4) Identify the primary catalog protocols (CWIC, HMA)
 - (5) Review detailed dataset availability and registration details (per Ken Casey)