MINUTES

OF THE

46th MEETING

OF THE

CEOS WORKING GROUP ON   
INFORMATION SYSTEMS AND SERVICES

(WGISS)

Oberpfaffenhofen, Germany

22 to 25 October, 2018

Hosted by

[German](http://www.inpe.br/) Aerospace Agency (DLR)

Table of Contents

[1 WGISS Plenary Session, Part I 7](#_Toc531846912)

[1.1 Introduction 7](#_Toc531846913)

[1.2 Host Welcome and Logistics Information 7](#_Toc531846914)

[1.3 WISP Report 7](#_Toc531846915)

[1.4 WGISS Chair Report 7](#_Toc531846916)

[1.4.1 The CEOS SIT Technical Workshop 7](#_Toc531846917)

[1.4.2 32nd CEOS Plenary 8](#_Toc531846918)

[1.4.3 WGISS Terms of Reference 9](#_Toc531846919)

[1.4.4 CEOS Chair 2019 9](#_Toc531846920)

[1.4.5 WGISS Cooperation 10](#_Toc531846921)

[1.4.6 WISP Support 10](#_Toc531846922)

[1.4.7 WGISS Vice-chair Nomination 11](#_Toc531846923)

[1.4.8 CEOS 2019-21 Work Plan 11](#_Toc531846924)

[1.5 CEOS Executive Officer (CEO) Report 11](#_Toc531846925)

[1.6 SEO Report 12](#_Toc531846926)

[1.7 WGCapD Report 13](#_Toc531846927)

[1.8 WGISS Brochure 13](#_Toc531846928)

[2 GEOSS-WGISS Interoperability and Future Data Architectures 14](#_Toc531846929)

[2.1 Introduction 14](#_Toc531846930)

[2.2 CEOS WGISS Overview and Report 14](#_Toc531846931)

[2.3 GEO SEC Overview and Report 14](#_Toc531846932)

[2.4 WGISS Connected Data Assets State of the Art 15](#_Toc531846933)

[2.5 CEOS FDA Activities 16](#_Toc531846934)

[2.6 GEOSS Platform and GEO Portal 16](#_Toc531846935)

[2.7 NextGEOSS 17](#_Toc531846936)

[2.8 Regional GEOSS 17](#_Toc531846937)

[2.8.1 AmeriGEOSS 18](#_Toc531846938)

[2.8.2 AfriGEOSS 18](#_Toc531846939)

[2.8.3 EuroGEOSS 19](#_Toc531846940)

[2.9 GEOSS Data Providers Workshop Outcomes 19](#_Toc531846941)

[2.10 GEOSS Evolve Initiative 20](#_Toc531846942)

[2.11 WGISS Contribution to DMP Evolution 21](#_Toc531846943)

[2.12 Discussion Topics 21](#_Toc531846944)

[3 Data PRESERVATION and STEWARDSHIP 23](#_Toc531846945)

[3.1 Data Stewardship Reference Model – White Paper 23](#_Toc531846946)

[3.2 WMO Stewardship Maturity Matrix for Climate Data 23](#_Toc531846947)

[3.3 EVER-EST to Support the Research Lifecycle Management 24](#_Toc531846948)

[3.4 Knowledge Management System (KMS) 25](#_Toc531846949)

[3.5 Measuring EO Data Usage – Best Practices 25](#_Toc531846950)

[3.6 CERN Archiving Centre 26](#_Toc531846951)

[3.7 ISO 19165-2 Standard Status 27](#_Toc531846952)

[3.8 Knowledge Sharing Opportunities: Conferences 27](#_Toc531846953)

[4 Data DISCOVERY and ACCESS 28](#_Toc531846954)

[4.1 WGISS Connected Data Assets Report 28](#_Toc531846955)

[4.2 CWIC Report 29](#_Toc531846956)

[4.3 IDN Report 29](#_Toc531846957)

[4.4 FedEO Update 30](#_Toc531846958)

[4.5 ESA Pilot Datacube for TPMs/HMs/EEs 31](#_Toc531846959)

[4.6 Partner Reports 32](#_Toc531846960)

[4.6.1 EUMETSAT Report 32](#_Toc531846961)

[4.6.2 JAXA Data Portal 32](#_Toc531846962)

[4.7 CEOS OpenSearch Conformance Test Document 32](#_Toc531846963)

[4.8 WGISS Connected Data Assets Client Guide Document 33](#_Toc531846964)

[5 Data INTEROPERABILITY and USE 34](#_Toc531846965)

[5.1 CEOS Tasks Involving WGISS 34](#_Toc531846966)

[5.1.1 FDA-08: FDA Reference Architecture 34](#_Toc531846967)

[5.1.2 FDA-09: Inventory and Characterise Existing FDAs 34](#_Toc531846968)

[5.1.3 FDA-10 Inventory of Software and Tools (Open Source) 35](#_Toc531846969)

[5.1.4 WGCV-01: Data Formats and Interoperability 35](#_Toc531846970)

[5.1.5 WGCV-03: CEOS Data Cubes and Test Site Data Access 36](#_Toc531846971)

[5.1.6 CARB-15 Carbon Data Portal Prototype 36](#_Toc531846972)

[5.2 Interoperability: GO FAIR, RDA 36](#_Toc531846973)

[6 TECHNOLOGY EXPLORATION 37](#_Toc531846974)

[6.1 Interfacing Tools with Data 37](#_Toc531846975)

[6.2 FOSS Libraries and Tools to Raise EO Data and Applications 37](#_Toc531846976)

[6.3 Google Dataset Search Tool Evaluation (DLR) 38](#_Toc531846977)

[6.4 Google Dataset Search Tool Evaluation (NASA) 38](#_Toc531846978)

[6.5 DATA-14: White Paper on Single Sign-On (SSO) Authentication 39](#_Toc531846979)

[6.6 Artificial Intelligence for Earth Observation - Towards an R&I 39](#_Toc531846980)

[6.7 FDA topics going forward 40](#_Toc531846981)

[7 Agency and Liaison Reports 41](#_Toc531846982)

[7.1 Hungarian Space Office Liaison Report 41](#_Toc531846983)

[7.2 NASA Agency Report 41](#_Toc531846984)

[7.3 ESA Agency Report 41](#_Toc531846985)

[7.4 CSIRO Agency Report 42](#_Toc531846986)

[7.5 CNES Agency Report 42](#_Toc531846987)

[7.6 USGS Agency Report 42](#_Toc531846988)

[7.7 JAXA Agency Report 42](#_Toc531846989)

[8 WGISS Plenary, Part II 43](#_Toc531846990)

[8.1 Future Meetings 43](#_Toc531846991)

[8.2 WGISS Summary 43](#_Toc531846992)

[8.3 WGISS-45 Actions 44](#_Toc531846993)

[8.4 Concluding Remarks 46](#_Toc531846994)

[9 Glossary of Acronyms 47](#_Toc531846995)

List of Participants

CEOS Executive Officer Steven Hosford\*

CERN Jamie Shiers

CNES Richard Moreno

CNR Federica Foglini\*

CONAE Homero Lozza\*

CSIRO Rob (Rob) Woodcock

DLR Katrin Molch, André Twele, Thilo Erbertseder, Hendrik Zwenzner, Felix Bachofer, Stefanie Holzwarth, Sven Kröger, Charlotte When, Maren Künemund, Heiko Damerow, Torsten Heinen, Nicole Pinnel, Max Wegner, Dieter Hausamann, Gina Campuzano, Eberhard Mikusch, Cristian Chereji, Johanna Senft, Christian Strobl, Bernhard Buckl, Oleg Goussev, Matthias Rentsch

EC Anica Huck

ESA Mirko Albani (WGISS Chair), Guido Colangeli, Andrea Della Vecchia, Damiano Guerrucci, Marco Leonardi, Rosemarie Leone\*, Iolanda Maggio, Philippe Mougnaud, Joost van Bemmelen, Piepgrass (WGISS Secretary)

EUMETSAT Uwe Voges

GEO Paola de Salvo\*, Bente Bye

HSO Gábor Remetey-Fülöpp

JAXA Makoto Natsuisaka, Yosuke Ikehata

NASA Andrew (Andy) Mitchell, Yonsook Enloe, Dawn Lowe, Christopher (Chris) Lynnes, Michael Morahan, David Borges\*, Brian Killough\* (CEOS-SEO)\*, Douglas (Doug) Newman\*, Hampapuram Ramapriyan (Rama)\*, Nancy Searby\*, Archibald (Archie) Warnock

NOAA Martin Yapur, Anne Kennerley\*, Ge Peng

NOAA & NASA Liping Di\*, Eugene Yu\*, Gil Heo\*, Li Lin\*

SANSA Imran Saloojee*\*,* Karabo Mithi\*

UKSA Kristi Kline, Eldrich (Rich) Frazier\*

UFSM Manoel de Araujo Sousa Junior



# WGISS Plenary Session, Part I

## Introduction

Mirko Albani, WGISS-Chair (ESA), opened the WGISS-46 meeting, thanking everyone for their participation. He asked each participant to introduce him/herself, and reviewed the agenda. The agenda was adopted.

## Host Welcome and Logistics Information

Gunter Schreier, Deputy Director of the DLR Earth Observation Center and responsible for business development, welcomed the participants to the meeting on behalf of DLR.

Gunther described the Oberpfaffenhofen facility, and discussed the 16 sites of DLR. The core tasks of DLR are the EO ground segment, and the use of EO data. DLR is engaged in data centers, ocean security and research in the areas of atmospheres, land surface, and geo-risks and civil security. Gunther described the terrestrial products as basis for global change research and discussed the goal to merge EO data with social media data.

The International Remote Sensing Data Centers (DFD) ground station network consists of Inuvik, Canada, Neustrelitz and Oberpfaffenhofen Germany, and Antarctica.

In addition to supporting DLR missions, the EO center supports national missions, ESA missions, and government and industrial missions. He displayed statistics of data volume in the German Satellite Data Archive D-SDA.

Gunther also discussed the Copernicus Data Exploitation Platform Deutschland CODE-DE, the Urban Thematic Exploitation Platform, and regional exploitation platforms.

Gunther thanked Katrin Molch and the DLR team for all their work to plan and organize the meeting, and wished WGISS a successful and productive meeting.

Katrin discussed meeting logistics of interest to the participants.

## WISP Report

Michelle Piepgrass gave a report on the status of the WGISS Infrastructure Services Project. She outlined the existing team, noting that the team is currently without a lead.

Michelle described the current WGISS organizational structure, which is displayed on the WGISS website, and includes a list of the agency principals of the agencies active in WGISS. She also mentioned how WGISS fits within the CEOS structure. She noted that it is important to notify the WISP team of any new or departingmembers so the mailing lists can be kept up-to-date, as well as changes to the WGISS webpages. The WISP team is also available to support WGISS outreach activities.

Michelle described the naming convention for presentation files, which are posted on the meeting website within the agenda. She also provided the web conferencing information.

## WGISS Chair Report

Mirko Albani gave a report on key WGISS topics.

### The CEOS SIT Technical Workshop

The CEOS SIT Technical Workshop included two days of meetings of the CEOS Virtual Constellations and Working Groups to foster cooperation. Mirko listed several key outcomes related to WGISS:

1. There is a need to include new FDA elements/platforms (e.g. APACHE, PANGEO, BIOMASS MAP, etc.) in the WGISS inventory and to compile “lessons learned” from a user perspective. In addition, there is a need to address challenges of tool interoperability – WGISS.
2. The transition of the FDA Ad-Hoc team tasks should be clearly addressed through updates/additions to the CEOS groups mandates/ToR and work-plans. Need to ensure coordination – WGISS, SEO, and LSI-VC.
3. The WGISS Technology Exploration subgroup should continue to review “future” data architectures and report back to CEOS leadership on what CEOS should be considering for the future to enhance data use and impact (e.g. Machine Learning, AI platforms). Look at emerging trends over the horizon and disrupting technologies – WGISS.
4. There is a need to review the new “Google Datasets Search Tool” to determine how CEOS datasets are represented – WGISS.
5. There is a need to update the inventory of Virtual Constellations datasets and ensure discoverability/accessibility through WGISS Connected Data Assets infrastructure –VCs/WGISS.

Mirko noted that WGISS was given a slot on the agenda, during which WGISS was given an action: SITTW-2018-04: SIT Chair to propose WGISS participation in the GEO Expert Advisory Group. Due by 30 Sep 2018.

GEO has formed an Expert Advisory Group (EAG) with a mandate until end 2019 and with the task to provide expert advice to the Secretariat Director on how best to design future GEOSS. The purpose is to provide relevant and timely advice with respect to cutting-edge science and technology required for a results-oriented GEOSS. The principal deliverable of the EAG will be a document outlining a strategy for implementation of the vision for GEOSS. To prepare the document to be presented to the GEO Executive Committee, the Expert Group will have three meetings, held in Geneva, in connection with meetings of the GEO Programme Board.

It was suggested that Rob Woodcock, supported by Mirko Albany or Andy Mitchell, could be a suitable representative for WGISS. Rob reported that he was told officially that he can provide that support. Mirko asked if he can inform SIT of this, and Rob agreed.

**WGISS-46-2**: Mirko Albani to inform the SIT Chair that CSIRO has confirmed nomination of Rob Woodcock as WGISS representative in the GEOSS Expert Advisory Group. Mirko Albani and Andy Mitchell will support Rob as needed. WGISS suggests that a letter be sent by the SIT Chair to the GEO Secretariat.

### 32nd CEOS Plenary

The 32nd CEOS Plenary occurred on 17-18 October 2018 in Brussels, Belgium. The Plenary decision that affects WGISS is that the FDA is closed and activities divided among permanent CEOS entities. Confirmation is needed regarding the role of WGISS among the Plenary actions. Since the CEOS Plenary 2017, five themes have progressed:

1. Analysis Ready Data (ARD), key to facilitate and foster EO data uptake especially by user communities, which are not comprised of EO data specialists. Progress has been made and activities need to continue in improving the overall ARD development logic (including nomenclature) development steps and principles continue to emerge from the CEOS LSI-VC WG. Generating interoperable and harmonized data products - continuous improvements are being achieved for the Landsat 8 – Sentinel 2 data products. Developing ARD on-demand demonstrators in close interaction with all stakeholders; this is done with strong involvement of commercial data providers. Developing the CEOS ARD 4 Land in which specifications for three product families have been completed and product assessment is ongoing.

Recommendation: Oversight of ARD activities to the CEOS LSI-VC and close coordination with WGISS. Assure continued visibility in CEOS context. Integration of new space actors.

It is not clear what WGISS involvement is with ARD, except the need to monitor what is happening in that domain. There are still a lot of question son what this means in terms of involvement; LSI was working on it for land, but it is not know who will do it for oceans.

Rob suggested that WGISS should be concerned with interoperability and discovery so platforms can use them, with the goal to communicate that they are available and can connect to CEOS assets. The only caveat is that the CEOS Data Cube also exists, and WGISS is involved with this.

1. Data cubes: A large variety of data cubes have emerged. The "CEOS Data Cube" initiative aims at utilizing CEOS agency resources, global partnerships, satellite data, and CEOS organizational groups for further improving data access, data preparation, and data analysis for all users of satellite data. The CEOS Data Cube (CDC) effort is well managed and supported within the CEOS community (via the CEOS Systems Engineering Office) and also the Open Data Cube (ODC) Partnership. The accumulated knowledge is key for the further evolution of a harmonized data cube domain. There are a number of data cube activities that can still be integrated (e.g. European Data Cube Initiative, various industrial Data Cubes).  There are efforts to integrate data cubes in EO data platform environments and to provide an infrastructure (e.g. cloud) substrate.

Recommendation: Continue oversight of these activities by the CEOS Data Cube (CDC) coordination team. Assure continued visibility in CEOS context.

Gunther asked if the software of the DCs will be available. Rob said that the Open DC is on GitHub. The CEOS DC is also open source.

**WGISS-46-28**: WGISS-Exec (Rob Woodcock, Andrea Della Vecchia) to prepare one to two slides for the SEO on what WGISS is doing for Data Cubes and propose a way forward for cooperation between WGISS and SEO on the FDA Data Cube topic.

1. EO Platforms: The concept of virtual data exploitation environments (providing data, tools, processing options, etc.) based on public/commercial infrastructures (e.g. cloud resources) constitutes a broader set of initiatives with the capacity to integrate ARD and data cube components.

Five Copernicus DIAS (Data and Information Access Services) have become operational in June 2018. DIAS data cube instantiations and ODC open source code are available. DIAS, ESA-NASA MAPs, etc. for future integration of data, tools, standards.

Recommendation: for WGISS to assist standardisation, and assure continued visibility in CEOS context.

1. User Metrics Resources Inventory: User metrics are important in the quantitative and qualitative assessments of the attractiveness and impact of the various CEOS FDA initiatives. Together with an up-to-date inventory of EO resources available through the CEOS community: improved harmonization and programmatic planning

a)     The CEOS WGISS WG has taken ownership of the User Metrics and is preparing the corresponding processes and documentation. ESA, DLR, NASA, USGS have already provided substantial information 🡪 highly dynamic parameters (e.g. number users, types of products, available processing capacity).   
  
b)     The EO Resource Inventory is closely linked to User Metrics and the CEOS WGISS WG has also taken ownership. This activity should also be harmonized with the Network of Resources activity, which are handled in conjunction with all ESA Member States in the ESA Data Coordination Body.

Recommendation: Oversight of all User Metrics and EO Resource Inventory by WGISS. Assure continued visibility in CEOS context.

1. Data Analytics: Data Analytics was identified early in the FDA process as a new theme in terms of CEOS coordination. Improved data analysis is also seen as a key driver to increase the usability and use of Earth Observation data, in particular by user communities, which have not been acquainted with EO. It is important to get a more complete picture of the range and the state-of-the-art of EO data analytics in the CEOS context. Specific communities, such as the Artificial Intelligence (AI) community are already formulating specific requirements toward EO data and product providers.

Recommendation: Task WGISS to identify perimeter and mechanisms to be applied to the Data Analytics. Assure continued visibility in CEOS context.

### WGISS Terms of Reference

**WGISS-46-3**: Michelle Piepgrass to send Terms of Reference (ToR) to Rob Woodcock and Mirko Albani. WGISS-Exec to update the WGISS ToR to reflect FDA activities assigned to WGISS and recommendations from the CEOS SIT and Plenary.

### CEOS Chair 2019

The new CEOS Chair is the Vietnam Academy of Science and Technology (VAST/VNSC), who will ensure the priorities and initiatives identified by the 2018 Chair (European Commission) and current SIT Chair (NOAA) are supported and further developed through 2019:

2018 Initiative #1: Laying the foundation for an international CO2 and GHG emission monitoring system

2018 Initiative #2: Bring the benefits of Future Data Architectures to the present - identify new targets

In addition, VAST/VNSC will pursue the following application focused initiatives:

Priority #1: Carbon Observations (forested regions)

Priority #2: Observations for Agriculture (rice)

The Chair initiative is to seek to integrate a number of ongoing CEOS activities and data in support of the target applications for a targeted region (Mekong Delta is proposed, in the areas of Forest Monitoring and Rice Monitoring

How can WGISS contribute to this initiative? Through

* Dedicated Use Case through WGISS Carbon Portal
* New Pilot activities complementing/supporting Vietnam Data Cube (VDC)
* Vietnam User Oriented Training and capacity building activities (with WGCapD)
* SAR Processing and SAR Applications
* Others?

It would be good to have something to present at the next CEOS Plenary about how WGISS was involved. CSIRO is coordinating the training in Vietnam.

### WGISS Cooperation

The following areas for cooperation were listed:

* CEOS Carbon Team - Ongoing development of prototype Carbon Portal (CARB-15)
* WGClimate - Ongoing review of ECV inventory to ensure CDRs access through WGISS Connected Data Assets Infrastructure (DATA-9)
* WGDisasters – Opportunity identified: generic recovery observatory highlighted as possible cooperation area during WGDisasters meeting#10, to be further discussed.
* WGCV – Ongoing around four different topics: Data Formats and Interoperability in the framework of FDA; Quality Indicators in Discovery Metadata, CEOS Data Cubes and CEOS Test Sites Data Access in support of WGCV Activities, and Standardization and Best Practices
* Virtual Constellations – Opportunity identified: update the inventory of Virtual Constellations datasets and ensure discoverability/accessibility through WGISS Connected Data Assets infrastructure
* WGCapD - Ongoing: WGISS Technology webinars and FDA events publicized through WGCapD, upcoming ones organized in cooperation (DATA-11 and FDA-5)
* SEO - Ongoing: COVE tool harvesting metadata from WGISS Connected Data Assets Infrastructure
* CGMS - Ongoing: Working Group (WG) IV Global Data Dissemination
* ESIP - Ongoing: exchange of information and best practices on data stewardship
* GEO - Ongoing:
  + GEOSS Platform accessing WGISS CDA infrastructure for CEOS agencies data discovery and access
  + WGISS contributing to GEOSS-EVOLVE initiative
  + Participation and support to GEOSS Data Providers Workshops (e.g. “User Metrics” session, May 2018)
  + Contribution to CEOS booth at GEO Plenary
  + WGISS represented in NextGEOSS Advisory Board and Working with NextGEOSS on federated authentication technology solutions
  + Joint Workshop on GEOSS-WGISS interoperability and FDA at WGISIS-43 and upcoming WGISS-46
  + WGISS representative will be proposed as member of GEOSS Expert Advisory Group

Andy Mitchell commented that WGDisasters is very interested in FDA; WGISS needs to communicate with more VCs and WGs.

**WGISS-46-8**: WGISS-Exec (Richard Woodcock, Iolanda Maggio) to further explore possible areas of cooperation with WGDisasters, specifically with regards to the generic Recovery Observatory and FDA/CDA.

**WGISS-46-18**: Mirko Albani to liaise with the CEO to trigger update of the inventory of Virtual Constellations datasets and then assess discoverability/accessibility through the WGISS Connected Data Assets infrastructure.

### WISP Support

The WGISS Infrastructure Services Project (WISP) provides technological support and core services for WGISS activities including teleconference and web conference support via GoToMeeting, website support, mailing list and meeting archive support. NOAA will support WISP through WGISS-47. The WGISS Chair will ask the SEO to support these services after April 2019 (in the event that no other agency actively involved in WGISS volunteers to take over WISP).

**WGISS-46-1**: Mirko Albani to bring to the attention of the SEO and the SIT the need for CEOS-globalized (e. g. VCs and WGs) WISP support; suggest this could be provided through the SEO.

### WGISS Vice-chair Nomination

Mirko noted that the WGISS Vice-chair position for the period October 2019 – October 21 is open. Mirko asked the participating agencies to consider filling this position.

**WGISS-46-5**: WGISS-Exec to identify potential WGISS Vice-chair candidates for period 2019-21 (to become Chair in 2021-23) and determine availability.

### CEOS 2019-21 Work Plan

The CEOS 2019-21 Work Plan will be updated with a set of actions for CEOS 2019-21 to address the mandate on FDA (see FDA AHT recommendations). Further details given in the CEO report.

For CEOS agencies’ contribution to WGISS, suggest to attract interest to WGISS by developing a table showing the different groups, with their scope, CEOS agency interest/benefits, and skills of representatives. Mirko displayed an example.

**WGISS-46-29**: Mirko Albani and WGISS-Exec to prepare an updated set of actions for the CEOS 2019-21 Work Plan to address the mandate on FDA (see FDA AHT recommendations) and other WGISS related activities.

**WGISS-46-4**: WGISS interest group leads to provide information for the ‘CEOS Agencies Participation/Contribution to WGISS’ table (a table showing the different WGISS groups, their scope, CEOS agency interest/benefits, and skills of representatives), and to finalize, distribute and post.

## CEOS Executive Officer (CEO) Report

Steven Hosford\* gave a report on general topics related to CEOS. He gave news of the CEOS Entity Leadership changes in one year:

* 2019 CEOS Chair – Vietnam National Space Centre
* 2020 CEOS Chair – Indian Space Research Organisation (ISRO)
* 2019 CEOS SIT Chair – NOAA
* 2020 CEOS SIT Chair – CSIRO
* 2019 October CEOS Executive Officer being sought.
* WGClimate Chair to extend mandate for 1 year
* WGCV Chair changed last week (incoming Cindy Ong, CSIRO)
* All other CEOS WGs looking for new Vice-Chairs in Oct 2019
* Steven listed CEOS upcoming meetings.

Steven gave the CEOS Work Plan items and timeline, and proposed changes:

* Using the WP to improve work traceability
  + Who: CEOS entity (+ nominative), Contributing agencies (+nominative), Resource estimates
  + What: More detailed deliverable descriptions and breakdown into milestones (where relevant)
  + When: Temporal information on milestones/resources
* Links from CEOS WP deliverables to GEO Work Programme elements. As the Space Arm of GEO, we need to know where our work is supporting the implementation of the GEOSS. References in CEOS 2019-2021 WP should be to elements in GEO WP endorsed in GEO-XV Plenary (October 2018)
* Agency-driven top-down requests. During WP development - open call to CEOS agencies to propose new CEOS deliverables
* Requirement for deliverable to be carried by a CEOS entity
* Need for consultation with responsible CEOS entity prior to submission to WP
* Modify the CEOS deliverables.org deliverable tracking tool to record
* Cumulative number of months due date has been set back for every deliverable
* Each WP deliverable should have
  + A responsible CEOS entity and person
  + More detailed description and resourcing plan (option of milestones)
  + Links to GEO Work Programme
* 2019-2021 CEOS Work Plan, two periods for contributors:
  + 9th Nov – 21st Dec
  + 25th Jan – 20th Feb
* Call for top-down agency level inputs (respecting constraints above)

In response to question from Mirko, Steven said that he will be sending out the template for the new elements soon.

Makoto asked about the increasing volume of data and whether the GEO Work Plan is taking this into consideration. Steven was not really sure.

## SEO Report

Brian Killough\* gave the report of the Systems Engineering Office (SEO) activities. He began with background the CEOS Visualization Environment (COVE) which now includes 131 missions and is linked to several mission archives to get metadata and browse images for previously acquired data (Landsat-5/7/8, Sentinel-1, Sentinel-2, and CBERS-4). It also includes new tools for coverage assessments and scene data to support data ordering. The Data Browser provides support for Landsat 5/7/8, Sentinel-1A/1B, Sentinel-2A, and CBERS-4, similar to the Coverage Analyzer, and filtering for cloud cover threshold (optical missions).

Status and Future Plans

* The team has made a number of updates to the core COVE tool to improve security and enhance operating performance. Some of these changes have caused issues with a few features that will be resolved in the coming months.
* Links to archive databases are a constant issue, as things change on the data provider side. Current issues with S1 archive at ASF to be resolved in 6-8 weeks. Other issues with SPOT, Pleaides, Radarsat-2, ALOS, TerraSAR-X and ResourceSAT-2 will be resolved in the coming months.
* Adding a custom mission capability to support future mission concept studies
* Adding coverage and revisit performance calculation capabilities
* Adding cloud cover overlays derived from Landsat historical metadata (13 in total, 1 raw average, 12 monthly average files)
* Adding coincident calculations for both future predictions and historical archive data

Open Data Cube News:

* New Africa Regional Data Cube deployed in May 2018 for five countries. Goal is 49 countries in 20 months.
* Attended SatSummit in Washington on Sept 19-21. The world is interested in Data Cubes, ARD, COG formats and cloud computing.
* Continuing to expand our supply of Jupyter notebooks for data cube applications.
* Strong desire to use “indexed” cloud-based satellite data with data cubes.
* Working on an ODC architecture framework and development of a Jupyter Hub ODC “Sandbox”

WGISS-SEO Collaborations:

* Working with Iolanda Maggio (ESA) to develop an approach for an Open Source Software Inventory (database). The NASA team is investigating options. (CEOS Action FDA-10)
* Maintain and expand the connections from satellite mission archives to the COVE tool.
* Investigate approaches for taking advantage of USGS data hosted in the AWS cloud. Once in place, WGISS could help developing scripts that pre-process the data and build cubes for download.

Mirko asked if WGISS can assist with connecting mission archives with COVE. Brian replied that the process is now seamless.

Kristi asked if flat files were recommended; Brian replied it would help but it puts a burden on agency or on WGISS.

Mirko suggested a discussion with Brian to see how WGISS can support the CEOS Chair initiatives.

## WGCapD Report

Dieter Hausamann gave a report on the Working Group on Capacity Building and Data Democracy. He listed the group’s 11 deliverables, including FDA-5: FDA’s awareness building and outreach – joint with WGISS. WGCapD was represented by Dieter Hausamann and Manoel de Araujo Sousa Junior.

WGCapD is working on three pilot efforts:

* NASA, DLR and CEOS SEO will implement the training calendar for its capacity building program to advertise trainings offered by ARSET and by SERVIR, through EOCollege and for WGCapD, respectively.
* Work with EUMETSAT for revised API and to train someone in how to use and maintain the calendar.
* Each pilot will define their “front end” based on their understanding, compare their approaches, and together agree on the approach that should be recommended for implementation for the WGCapD calendar.
* The WGCapD calendar design will be provided to the WGCapD participating agencies for final review and implementation by Q3 2019.

WGCapD provides training resources on CEOS website and INPE learning center, and is included in the GEOCAB Portal. The WMO Global Campus is implementing an open source library catalog.

Dieter discussed what WGCapD could be doing in the future can where WGISS participate:

* Collection of single agency resources discoverable at each agency’s storage location
* WGCapD collection housed together and discoverable
* User interface tailored for each end user class: decision makers, policy makers, EO professionals and scientists
* Use cases and trainings available by GEO region, national, sub-national scales
* Use cases and trainings available by GEO societal benefit area

Dieter also discussed what could WGCapD be doing to help WGISS:

* Raise awareness of the value of EO data products and services to user communities, including support to locate and access data, products, and tools, and targeted training workshops.
* Support CEOS initiatives and helps WGs and VCs undertake their own capacity building initiatives, e.g. guidance on best practices
* Build awareness of WGISS efforts in WGCapD Workshops
* What else?

Chris noted that WGISS is putting together an inventory of tools, and suggested a pilot exchange with WGCapD.

Rob made three comments:

* Get user experience feedback around FDA elements; but since WGISS does not engage with users, is WGCapD a better conduit?
* There is a lot of work in SEO DC, providing training materials. Jupyter notebooks provide all the necessary information, but it is not ordered – perhaps WGCapD can help.
* Most applications are optical but there is a lot of interest in SAR. Can WGISS work with WGCapD on SAR training materials?

Dieter noted that there is SAR education and expertise in the EOCollege; the number of new SAR applications is exploding.

Andy wondered if a liaison with WGCapD is desirable; Chris is in discussion with Nancy Searby.

**WGISS-46-11**: Rob Woodcock to follow up on a SAR training/meeting in Vietnam with WGCapD/CSIRO.

## WGISS Brochure

Iolanda Maggio presented the newly-developed and printed WGISS brochure and leaflets. Final comments were requested.

**WGISS-46-6**: Interest Group leads to send final comments on the WGISS brochure to Michelle Piepgrass and Iolanda Maggio.

# GEOSS-WGISS Interoperability and Future Data Architectures

## [Introduction](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-46/3.%20Wednesday/2018.10.24_09.00%20Introduction.pptx)

Mirko Albani introduced the workshop to address interoperability aspects between CEOS WGISS and the Global Earth Observation System of Systems (GEOSS) taking stock of the recent development in the frame of CEOS Future Data Architectures (FDA).

Mirko listed CEOS WGISS’ contribution to GEOSS:

* GEOSS Platform accessing WGISS CDA infrastructure for CEOS agencies data discovery and access
* Participation & support to GEOSS Data Providers Workshops (e.g. “User Metrics” session, May 2018)
* WGISS represented in NextGEOSS Advisory Board and Working with NextGEOSS on federated authentication technology solutions
* Joint Workshop on GEOSS-WGISS interoperability and FDA at WGISIS-43 and WGISS-46
* WGISS representative will be proposed as member of the GEOSS Expert Advisory Group
* WGISS contributing to GEOSS-EVOLVE initiative:
  + WP 1: GEOSS Architecture and Evolution
  + WP2: Functional Testing
  + WP 3: Data Management Principles
  + WP 5: Demonstrations Projects
  + WP 6: Community Portals

At WGISS-43 possible synergies were identified between CEOS WGISS and AmeriGEOSS in the areas of outreach and education, data and products, applications and tools, and use cases. Synergies between WGISS and NextGEOSS Architecture and interoperability include: data hub, discovery and access enablers, and enhanced distributed gateway to EO data, processing enablers, publishing appliances, and community portals.

Workshop objectives:

Strengthen interaction and coordination between WGISS and GEOSS at technical level.

Identify additional topics where we could work together and maximise outputs and impact.

Discuss and define CEOS WGISS contribution in shaping the evolution and future vision of GEOSS as a whole and in particular for what concerns ground segment systems and platforms (i.e. Future data Architectures).

Define a WGISS-GEOSS (Action) Plan and schedule periodic teleconferences to monitor progress in the activities.

## [CEOS WGISS Overview and Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-46/3.%20Wednesday/2018.10.24_09.10_CEOSS%20WGISS%20Overview%20and%20Report.pptx)

Mirko Albani gave an overview of WGISS. He described the WGISS scope and organizational structure and each of the four interest groups:

* Data Preservation and Stewardship
* Data Discovery and Access
* Interoperability and Use (including FDA)
* Technology Exploration

## GEO SEC Overview and Report

Paola de Salvo\* (GEO-SEC) gave an overview of GEO related to the CEOS integrated with future vision of GEO. She noted that Gilberto Camara is the new director of GEO, and he has a new strategy for Results-Oriented GEOSS. She agreed to send the strategy for Results-Oriented GEOSS to WGISS-All.

This new strategy is for a knowledge hub of EO data complemented with in-situ data, models, and algorithms to ensure reproducibility of analysis by the community. Another important element is the importance of ARD with which CEOS could have a very important role, and will lead to additional engagement with CEOS.

Regarding the new GEO Work Plan, the program board will release a call for revisions in December where new activities could be proposed, along with existing activities. CEOS is represented on the program board. The GEO-SEC will play a role as liaison with the board. This framework the GEO-SEC would like to incentivize CEOS to incorporate the ARD element for satellite information through cloud services, in order to provide an easier access to EO data from expert and less expert users. A potential element of discussion is whether ARD should become a foundational task.

Paola encouraged the active participation of CEOS in GEOSS-Evolve. The GEO Plenary is next week, with a side event on GEOS-Evolve, and members of CEOS are invited to the discussion.

The GEO-SEC is actively involved with all flagship initiatives for a deeper interaction and is contacting every flagship initiative to identify the EO data they need for better workflow for discoverability and accessibility.

The GEO Plenary next week will be an important time to understand the needs of the community.

Mirko asked for process of preparation of new GEO WP. Paola replied that in early December there would be a call for inputs to the new WP (new/existing activities). The program board will review and make decisions regarding these.

Mirko has asked for time at the GEOSS-Evolve side event at GEO, and will discuss ARD with the GEO-SEC.

Paola noted that CEOS is data provider ‘par excellence’ for GEOSS.

Mirko highlighted the need for information on the schedule for the GEO WP, since CEOSS is similarly working on renewing the CEOS WP where each activity needs to be linked to its relevance to GEO.

## [WGISS Connected Data Assets State of the Art](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-46/3.%20Wednesday/2018.10.24_10.00%20WGISS%20CDA.pptx)

Yonsook Enloe gave a complete summary of the CEOS WGISS Connected Data Assets. She indicated the metrics page, from which all connected data assets can be identified. She described recent collaborations with GEO.

Yonsook highlighted that the WGISS Connected Data Assets integration with GEOSS continues; the IDN is registered in the GEOSS Yellow Pages. Email and telecon discussion about adding “IDN” to the GEO Web Portal 🡪 “GCMD/IDN” is ongoing. FedEO and CWIC accessible from the GEO Web Portal and integrated via GEO-DAB, and WGISS has continued participation in the GEO Data Providers workshop. The WGISS Connected Data Assets webpage displays metrics; this information has been given to the GEO-SEC team.

Yonsook listed the following challenges:

* Harvesting metadata leads to stale metadata
* Removing or replacing outdated data
* Multiple paths to get to one data collection; even more paths to get to multiple copies/versions of the data
* How will user know which version of the data is the most recent or “the best”?
* Discovering and understanding services across agencies
* Interoperable and discoverable data services: Discover and compare Analysis Ready Data Products and analytic services

Mirko noted that the last two are key for WGISS.

Uwe Voges asked about dynamic search when harvesting metadata. It is very fast, but some of the granules are being copied many places as well.

Andy raised the point of getting the community to standardize using DOI to avoid accessing multiple copies. Because GEOSS is so distributed, no one is really accountable for eliminating multiple copies. This is why it is recommended to access through the IDN where keywords are vetted; if a desired keyword does not exist, it can be requested. The IDN is very reliable and very fast. Those who harvest may stop doing so when confidence in live search is generated; but there needs to be a best practices for DOIs.

NASA has also published papers on BPs for harvesting. Perhaps a CEOS BP for harvesting would be helpful.

Chris Lynnes suggested a short analysis to see if there is a way to recognize if something has changed in the IDN.

Rob raised the issue of linkage being a crowd source challenge, and how it scales is quite important. Guido Colangeli noted that this is handled in GEO in the architecture implementation pilot. Linking data to these ontologies is still being assessed and domain experts are needed to look into these keywords. Satellite data has to be linked to ground truth data.

Rob asked if they have looked at the GCMD keywords; there are many ontologies and one is the GCMD. Uwe added that Spacebel has an ontology for EO products.

Andy suggested having a workshop on ontologies.

Paola suggested organizing a full-day interaction including the GEO-DAB team to address the challenges and to ensure that any improvements can be reflected in the GEO-DAB. Mirko replied that WGISS needs to do some work on this first.

**WGISS-46-7**: WGISS-Exec (Mirko Albani, Rob Woodcock, Andrea Della Vecchia, Yonsook Enloe, and Liping Di) to explore the possibility of a Carbon Portal case study built around the VAST-VNSC Chair Initiative. After confirmation, WGISS-Exec will get in touch with the SEO and VNSC as needed.

**WGISS-46-21**: WGISS-Exec to consider organizing a session at the next WGISS-47 on EO ontologies as part of the DSIG.

**WGISS-46-22**: Michael Morahan, Andrea Della Vecchia, Liping Di to organize a meeting to discuss a metadata model for services description to be circulated and approved by WGISS Model will be used to describe services which will then be discoverable through the IDN SERF interface.

## CEOS FDA Activities

Rob Woodcock gave a presentation on future data access and analysis architectures (FDA), with the objective to see how GEOSS is operating, and to identify areas where GEOS and WGISS could practically work together.

CEOS FDA is driven by the expectation of substantial growth in the EO based digital economy across industry and government, and a step change in EO satellite capability over the next five years leading to new applications. Elements include:

* EO analytics platforms: Data + compute + tools
* Cloud hosted: scalable analysis
* Third-party application development: Common Dev Interfaces (APIs) and Data Cubes (not just files)
* Timely availability
* Ready to Analyse, not just access
* Multi-sensor integration
* Pixel level data discovery and access (refined search)

Rob gave three examples of user experience: Swiss DC, AquaWatch, and Big Data hosts (AWS on Earth, Digital Globe GBDX). Rob described the Swiss DC and AquaWatch in detail, noting that the mission of AquaWatch is to improve the coordination, delivery and utilization of water quality information for the benefit of society.

Rob raised the question: From original data store, to pre-processing, analytics optimized data store, to analysis, to visualization, how can CEOS assist to make this manageable for the users? It will require interaction, integration, interoperability, and interfaces.

Rob listed the following discussion points: Scaling to global analysis is technologically plausible, but scaling of actors is challenging because of changing user expectations, governance, economic, provenance, new participant, and discovery and access impacts.

Kristi noted that the goal for Landsat is to deliver what is happening and changing on the landscape, the way weather is done.

## [GEOSS Platform and GEO Portal](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-46/3.%20Wednesday/2018.10.24_11.50_WGISS%20GEOSS%20Platform%20and%20GEOSS%20Portal.pptx)

Guido Colangeli gave a presentation on the GEOSS Platform and the GEO Portal. He discussed the current resources in GEOSS, the global policies that it supports, and the various initiatives. Guido discussed providing an access point for scientific knowledge bodies*:* Targeted search of Sustainable Development Goals (SDG), Wikipedia, and GEO Programmes.

GEOSS is moving toward a user-centric GEOSS: The generic interface for many users, but to specific searches/needs.

The GEOSS View Subset of the GEOSS resources is defined by applying a set of clauses: Discovery clauses (e.g. spatial envelope, keywords, sources, etc.), and access clauses (e.g. data format, access protocol, CRS, etc.)

The GEOSS user-centric approach is built around three pillars: customize, collaborate, reuse. This is applied in the GEOSS Status Checker, GEOSS View, GEOSS Mirror, GEOSS Like, GEOSS Widget, GEOSS API, GEOSS Like, and GEOSS Yellow Pages.

Mirko commended GEO for the major improvements in GEOSS.

Rob asked how CEOS/WGISS can help the GEOSS ecosystem. Examples would be a single point of connection, relying on the ranking of each knowledge provider, something able to discriminate based on the user needs. ARD is also essential. To help the GEOSS platform:

* Avoid duplication in metadata retrieval,
* Exploit ranking from CEOS in results,
* Work together on ontologies,
* Provision of ARD,
* Provision of services (and make them discoverable).

## NextGEOSS

Bente Lija Bye\* gave a brief overview of NextGEOSS concept and value proposition of supporting Sustainable Development Goals, with a user feedback mechanism integrated in the Data Hub, advanced discovery tools, and community enhancement. She provided a diagram of the conceptual architecture and listed the pilot services.

Bente continued with five user experiences with using NextGEOSS:

* User engagement
* Data preparation
* Platform setup
* Pilots integration
* Service (pre-)Operations

Bente listed a set of personas that reflect the different tasks of the on-boarding. Capacity building is built into the training, and INSPIRE hackathons are also organized.

Bente reported that the first set of external pilot candidates has been selected. NextGEOSS offers [NEW] Earth observation information data hub services. NextGEOSS and GEOSS Evolve coordinate the next GEOSS AIP - opens January 2019. GEO also makes many announcements on social media, and upcoming GEO Plenary and AGU.

Andy asked if the catalog would remain after three years; Bente replied that yes, it will be maintained. There are 27 partners, and some of the partners will be engaged.

Chris asked if they have a way of measuring the impact of NextGEOSS. Bente replied that they are carefully monitoring and reporting KPIs.

Mirko asked for the main source of satellite data. Bente replied that it is the Sentinel Hub, and they are working on a solution that knows the data policies, speed and other elements.

Mirko asked Bente to describe the relationship with the GEOSS Platform. Bente replied that they are discussing APIs for the purpose, and that GEOSS should be able to harvest anything in NextGEOSS. But for NextGEOSS, the center is the user and the goal is to make such data considerations irrelevant to the user. The platform is at the end of the first external pilot and is about to be released to the public.

Richard asked what the link with EuroGEOSS is. Three members of EuroGEOSS are on the GEO Board.

WGISS is creating an inventory of all the platforms for EO satellite data; NextGEOSS contribution would be helpful, in addition to their contribution to the SSO discussions.

**WGISS-46-12**: Iolanda Maggio to communicate with NextGEOSS, AmeriGEOSS, AfriGEOSS and EuroGEOSS about joining the WGISS inventories (FDA and/or SW/Tools).

## Regional GEOSS

AmeriGEOSS, AfriGEOSS, and EuroGEOSS are regional initiatives under the GEO Framework that promote collaboration and coordination among GEO Members.

### AmeriGEOSS

Nancy Searby\* and Rich Frazier\* gave a presentation on AmeriGEOSS. Nancy began with background information on AmeriGEOSS, and on the AmeriGEOSS Community Platform, which is an exploitation platform for collaboration and discovery. Nancy described the draft architecture, which recognizes the need to extend the current architecture to support the development of products and services.

Rich described the architecture of the platform, and a few of its functions. It provides a variety of methods to search for data, find data in various formats and use the resources from a variety of platforms and devices. The AmeriGEOSS Data-Hub uses the CKAN open source platform to provide access and to link resources in the region. The search tools allow the user to search and discover resources through keyword text search, faceted search, and spatial search. Formats can be searched in the faceted search tool. The search results also present the available formats so users can find the exact resources they need more easily.

The platform operates on many devices including desktops, tablets and phones. Users can interact with data and documents, and can create graphs, charts, maps and more. After users discover resources that are of use, they can view reports and dashboards.

One of AmeriGEOSS’ community initiatives is GEOGlows to support capacity building to understand water use, water availability, and water quality. The outcomes of these activities will support clean water and sanitation efforts under SDG 6.

The AmeriGEOSS community acknowledges that imagery is a powerful resource that can improve understanding and decision-making, and is developing a tool that will allow novice and advanced users to leverage imagery for improved understanding and decision making.

WGISS raised questions around the source of their data. If a user finds something on the AmeriGEOSS hub, it will generally also be found on the GEO Portal.

Mirko asked for areas where WGISS can cooperate with AmeriGEOSS. Nancy and Rich replied that the help received so far has been significant, and further assistance with finding resources, products, and services is appreciated. The Water Portal has been very useful. Cooperation with WGCapD is also helpful.

Yonsook mentioned the WGISS tools inventory, which can be made available; Rich noted that categorization by thematic area would be very useful.

Richard suggested adding CKAN to the inventory.

Joost mentioned a related presentation at the CEOS Plenary.

### AfriGEOSS

Karabo Mithi\* gave a presentation on AfriGEOSS. He discussed the physical infrastructure, location and processing, and mentioned the ground receiving station in Gauteng. Imraam gave an overview of the system and listed the processing servers, and the missions that they receive and archive. He also discussed the online catalog that has search and ordering capability.

Karabo introduced the African Research Cloud project, and the data being made available on ARC, and gave a high level description of SANSA’s contribution to the ARC. He discussed South African EO, their National Integrated Cyberinfrastructure system, and SANSA’s involvement in PetaFLOP, Lengau, and DIRISA.

The emerging E&EO Research Data Infrastructure for Earth and Environmental Sciences is an Open Data Platform with shared metadata, data hosting, services, components, portals, gateways, guidance, and physical infrastructure.

Karabo also discussed the role of regional and national GEOSS: AfriGEOSS. He listed the member countries, and the 2017 symposium with key outcomes related to data. He noted that the African Research Cloud Project is evolving into all Africa. The Data Cube is being implemented, and AfriGEOSS is developing capacity on SDG’s.

Mirko asked how WGISS can provide support. Karabo said he would pass the question to Imraan Saloojee and reply by email.

Andy wondered if the SANSA portals should be added to the CEOSS-WGISS list of tools and services.

### EuroGEOSS

Anica Huck gave a presentation on the EuroGEOSS initiative. This new initiative is driven by the EC to increase the benefits of GEO to Europe, and to leverage and make European EO assets visible internationally. She described the governance and strategic activities:

* Inventory of EO development & innovation actions
* Up-scaling of EuroGEOSS pilot applications
* Showcasing of European know-how with relevance to GEO
* Linkages between EuroGEOSS applications and other GEO actions
* Consolidation of national GEO management structures

The characteristics of EuroGEOSS are to provide a simple and flexible framework (voluntary basis). EuroGEOSS is specific, overcoming fragmentation, and working on the downstream part of the EO value chain to

* Leverage existing national Earth Observation services activities
* Strengthen European GEO coordination
* Exploit existing EO European Data Platforms
* Branding and promotion
* Identification of resources

Anica described how EuroGEOSS fits in the GEO context, and explained the specific objectives of EuroGEOSS. The European environmental data platforms are positioned at different levels of the EO value chain and complementary to the delivery of EuroGEOSS services.

EuroGEOSS is addressing current gaps in GEO. Two clear routes for the implementation of EuroGEOSS through new fixed term projects or open innovation partnerships are bringing together existing projects. The current GEOSS platform enables discovery and access of data; the evolution of the platform promotes/leverages EuroGEOSS products. Delivery of services/products/applications through GEOSS and EuroGEOSS, and outcome of the commercial sector engagement is a discussion in GEO.

Anica described the relationship between EuroGEOSS and Copernicus, and the ongoing Horizon 2020 call on EuroGEOSS.

EuroGEOSS is requesting expressions of intent to address the challenge of accelerating the users' uptake of EO data/information for benefit of Europe.

Anica concluded with the following points:

* Successful initial phase of EuroGEOSS; engagement of relevant organisations and countries; structure in place; support through Horizon 2020 ensured; connections with the Copernicus programme are established
* Two clear routes for the implementation of EuroGEOSS: 1) Through new fixed term projects 2) Through open innovation partnerships bringing together existing projects
* In future, improved integration of existing data from European EO assets and initiatives within GEO context into user oriented initiatives envisioned
* Effective use of European Earth observation resources (including space, airborne, in-situ measurements and citizens observatories)
* Scale-up of scientific results and preparation for operational activities
* Part of the Commission proposal for Horizon Europe

Mirko asked Anica to take back to the team a request for areas where WGISS can contribute.

## [GEOSS Data Providers Workshop Outcomes](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-46/3.%20Wednesday/2018.10.24_15.50_%20GEOSS%20Data%20Providers%20Workshop%20Outcomes.pptx)

Paola De Salvo discussed the outcomes of the GEOSS Data Providers Workshop held in May in Frascati. The workshop is the result of major collaborations. Paola explained the GCI journey, GEOSS Portal, GEO DAB, and the new GEOSS Portal; the GCI becomes the GEOSS platform. With increase of resources comes increase of data providers, and participants in the GEOSS Data Providers Workshop. The most recent workshop had 200 participants, 130 organizations, 5 continents and 33 countries. It included participants from governmental, research, education, commercial, and inter-governmental participants.

Focal areas connecting resource providers with users, include:

* Uncovering user content and data discovery, access and use needs;
* Highlighting data user and application impact stories;
* Optimizing data for decision and policy making;
* Engaging with citizen observatories;
* Presenting data potential; and
* Entering in discussions.

The goal of the workshop is presenting user-centric GEOSS Platform enhancements, building strategies for strengthening user and provider, and outreach and expanding the user base. Paola described key messages of the workshop:

* The 3rd GEO Data Providers Workshop demonstrated the strengthening of the GEOSS Ecosystem and maturity of the GEOSS Platform with respect to GEO and outside traditional GEO community.
* Present at workshop and engaged to be part of GEOSS Platform:
  + GEO Flagships: GOS4M, GEOGLAM, GEOBON
  + GEO Initiatives: 4 Regional GEOSS; GEOSS Evolve, GEO CRI, GEO Wetlands, Eco-potential (GEO – ECO);
  + GEO CAs: CAMS; C3S; DIAS; GEO-CRADLE; Himalayan GEOSS
* High interest from Data Providers to be part of the GEOSS Platform: Pipeline of 60 New Data Providers to be brokered.
* User needs have become a priority for the present GEOSS Platform and will be an essential driver for future GEOSS evolvement.
* The transition from data to knowledge was a main topic raised in the workshop. This entails specific attention on interaction between GEOSS Platform and knowledge generation systems and services.

## GEOSS Evolve Initiative

Stefano Nativi\* gave a presentation on the GEOSS Evolve initiative, which is one of the official GEO activities. He began with a description of GEOSS; the present result-oriented GEOSS is a resource-oriented GEOSS – collecting, and sharing as many resources as possible. GEO has devoted considerable efforts to building the GEOSS infrastructure and capabilities that have made Earth observations discoverable. To continue leveraging these successes through 2025, GEO will evolve by extending the user audience, placing additional focus on the accessibility and usability of EO, providing a service framework to engage partner and user communities in evolving the current infrastructure, and evolving the current system of systems. There are certain boundary conditions or design constraints: the GEO principles, GEO Governance, Architectural, Data Sharing, and Data Management Principles.

The GEOSS Evolve initiative intends to address the previous needs, with the bjectives of connecting better the data services and products offered, improving data management and sharing, and ensuring the evolution of the infrastructure in the light of the rapid technological development.

GEOSS is progressing along three strands: The development of regional GEOs, the further development of the GCI (now GEOSS Platform) engaging partners and user communities, and the development of a strategy for a “result-oriented” GEOSS led by the GEO-SEC.

The recognized challenge is that it is not yet known what form a “result-oriented” GEOSS will take and how it will evolve from the current “resource-oriented” GEOSS platform. The Expert Advisory Group has been tasked to address a number of questions. The GEOSS Evolve initiative helps such a process having already recognized a set of key issues structured around architecture, interoperability and governance.

Architecture: what do we mean by Knowledge Hub, how best to implement FAOF principles, in particular the reproducibility, and how can the GEOSS Platform best leverage cloud infrastructure and emerging new patterns such as edge and fog computing.

Interoperability: How to ensure interoperability with Regional GEO if and when these develop technological infrastructures to support their mission? How to address the interoperability of multiple data cubes emerging globally?

Governance: What institutional mechanisms need strengthening to increase access and use of in-situ data? What part can Regional GEO play? How will the future GEOSS Platform interact with the Regional GEO initiatives, how can development of GEOSS Platform into a knowledge hub be made in connection with existing structures (Flagships, initiatives, Foundational Tasks?) Shall GEO focus on in-situ data management, only? Possible governance scenarios: three models: collaborative, acknowledged, and autonomous. Star model, mesh model, or Silos model.

Recommendations: The “Silos” model (in which each regional node develops independently with limited or no interoperability with the rest) could not satisfy the GEO Constraints for developing GEOSS. The “Star” and the “Mesh” models (or a hybrid solution) would continue to support GEO Constraints for developing GEOSS. It does not matter so much if the emerging GEOSS is more regionalized or centralized, or whether the tighter or looser interoperability model prevails.

GEO must leverage the new architectural structures and technological developments to extend and increase the level of service of the GEOSS to its community –of over 10,000 data providing organisations and over 14,000 user organisations. GEO must also maintain the GEOSS as a System of Systems based on interoperability and the sharing of data and knowledge –so that the whole continues to be greater than the sum of its parts.

This calls for the following actions:

* Governing the evolution of the GEOSS so that it continues to develop in the desired direction. This implies, for example, the need for a stronger collaboration between GEOSS EVOLVE, the GEOSS platform, the regional nodes and the GEO Flagships
* Reinforcing the process of formalizing, sharing, and making reusable the knowledge existing in the GEO community so that the FAIR principles supporting reproducibility are achieved
* Ensuring interoperability with existing and new analytics systems such as analysis-ready data, data cubes, Big Data analytics platforms, and manage their orchestration
* Reinforcing interoperability, for example through open and standard APIs, between GEOSS and the relevant global systems, platforms and initiatives, such as the UN’s SDG Knowledge Platform, CEOSS/WGISS, UNNGIM, UNEP, ICSU’s World Data System

Stefano mentioned the 11th International Symposium for Digital Earth in Florence, September 2019.

Mirko commented how many of their questions are parallel to WGISS; WGISS will be joining the Expert Advisors Group. Stefano added that WGISS can also contribute through its ESA members.

## WGISS Contribution to DMP Evolution

Iolanda Maggio gave a presentation on the WGISS contribution to DMP evolution. She described the Data Management Principles and the Maturity Matrix – what it is, who will use it, and why it is needed. Iolanda described several exercises, including a DMP Maturity Matrix assessment and its results. A second exercise performed was the WGISS Maturity Matrix with DMP IG simplifying the Data Stewardship Maturity Matrix.

WGISS is preparing a Data Management and Stewardship Maturity Matrix White Paper describing the evolution of a simplified Stewardship Maturity Matrix taking into account input from the DMP IG. The Stewardship Maturity Matrix considers specific EO requirements, resulting in a CEOS Maturity Matrix showing improvement of the Data Management Principles using the CEOS Maturity Matrix.

Bente commented that data quality is not included because it depends on the use of the data. Iolanda noted that the important thing to the MM is that it can be tailored to the environment. Mirko added that quality is determined by the user. This is an assessment to correct the objectives and to monitor.

Ge Peng stated that data quality is multi-dimensional; there is the scientific quality, but this work is about documenting the quality assurance practices, and the maturity of the data management and stewardship practice, independent of the scientific quality; data integrity is from the perspective of the data stewardship quality.

Iolanda added that the NCEI DSMM had three quality columns; here they have been reduced to one.

**WGISS-46-26**: Iolanda Maggio to prepare a proposal for GEO DMP improvement/update based on the work done by WGISS (e. g. Maturity Matrix).

## Discussion Topics

The following discussion followed the workshop

* Rob noted that GEOSS and WGISS are trying to achieve interoperability between countries; it is astonishing what has already been achieved.
* Bente commented that the main argument is that those close to the data are those best suited to advise. The space segment is well organized, but there are still issues to cover. With in-situ data there are larger problems.
* Stefano noted that with networks connecting different machines, the main obstacle was with the application of the infrastructure. Does it make sense to develop structure to share data when the application is not yet known? Is it useful to have mainly a public driven infrastructure, or should the private sector be leveraged.
* Andy raised the topic of creation of create new portals. The small curated portals serve a useful purpose; the big ones less so. Richard replied in defence of community portals: CNES is building a big portal to launch from all themes with a centralized catalog, letting the community portals access that catalog. EO science data are scattered, and a catalog means the portal is a bit easier for the users. It is not whether or not another portal, but rather the cost to build and maintain. This is where the catalog comes in. People use more and more different kinds of data together now. But portal proliferation is a concern.

**WGISS-46-13**: Chris Lynnes, Mirko Albani and Rob Woodcock to put together one page on recommendations for cooperation between GEO and WGISS based on the outcomes of the WGISS-46 joint workshop.

# Data PRESERVATION and STEWARDSHIP

## Data Stewardship Reference Model – White Paper

Iolanda Maggio gave the background and scope of the Data Stewardship Reference Model. The short white paper provides a few guidelines and recommendations for the preservation and improvement of data including a roadmap for scientific data stewardship improvement through WGISS assets support. She discussed the reference model activities, and the WGISS assets supporting data stewardship in terms of best practices, white papers, and systems and services.

Iolanda noted that the white paper has been distributed for comments and was well accepted, and she further clarified understanding of document and contents. The white paper will be re-distributed for final review.

Iolanda also presented a new page on the CEOS website for Lessons Learned and Best Practice documents, to be used by all CEOS working groups.

## WMO Stewardship Maturity Matrix for Climate Data

Ge Peng gave a presentation on the WMO Stewardship Maturity Matrix for Climate Data (SMM-CD). She began with description of the World Meteorological Organization (WMO) Information System (WIS). WMO is a specialized agency of the United Nations (weather, water, and climate) with 191 member countries and territories. It is committed to free exchange of data and products and is dedicated to ensuring the highest possible quality (data, information, and services) and effective access to authoritative, trusted datasets for science, policy and decision-making support.

Ge Peng listed the members of the SMM-CD Working Group, which is under the WMO High-Quality Global Data Management Framework for Climate (HQ-GDMFC), in collaboration with the members of an ad hoc International Expert Group on Climate Data Modernization (IEG-CDM) to help address some of the challenges facing WMO and WIS.

HQ-GDMFC is an inter-programme initiative led by WMO CCl/CBS (Commission for Climatology/Commission for Basic Systems), and is a collaborative framework that enables an effective development and exchange of high-quality climate data based on reliable underpinning infrastructure at the global, regional, and national levels. The building blocks are Data Management Standards, Data Maturity Assessment, and Access to High Quality Datasets.

The framework consists of a catalogue, a matrix, and a discovery and access system. The SMM-CD’s goal is consistent maturity information of data management, stewardship, and governance practices.

Dataset quality is multidimensional. Data product lifecycle –stage-based maturity assessment models involve science, product, stewardship and service. The goal is to ensure that managed datasets are scientifically sound and utilized, fully documented and transparent, well-preserved and integrated, and readily obtainable and usable.

The structure of SMM-CD includes four categories: Data access, usability and usage, quality management, and data management. Peng listed the aspects of each, and defined five levels for the maturity scale, noting that risk decreases, and trust increases as each level is achieved:

* Ad Hoc – not managed
* Minimal – limit managed, not defined
* Intermediate – managed, defined, partially implemented
* Advanced – well-managed, well-defined, fully implemented
* Optimal – level 4 plus measured, controlled, and audited.

Outcomes include a Matrix and a Guidance Booklet. These are reviewed internally by the IEG-CDM team, and external community-wide reviews are performed by invited international domain experts (science, data management, and stewardship), the GCOS secretariat, and the ESIP (Earth Science Information Partner) community a. There is a working session at the ESIP 2018 summer meeting in July. An evaluation template was also developed.

Current status: The SMM-CD documents have been baselined. A use case of 16 global datasets identified by IEG-CDM Sub-Group A is underway. Five assessments have been completed and two more are near completion.

Peng listed the team, and those who have provided feedback and acknowledged their contributions. She also listed references. An EGU 2019 session is being planned for science data centers and repositories, and for end-users.

Rosemarie Leone asked for more detail on the use cases for a clearer understanding. Ge Peng replied that assessment and feedback will be provided. Rosemarie recommended an exchange on architecture.

**WGISS-46-9**: Mirko Albani, Richard Moreno, Ge Peng, Iolanda Maggio to further elaborate possible collaboration with the WMO International Expert Group on Climate Data Modernisation (IEG-CDM), WIS and WIGOS, and possibly organize a joint session at WGISS-47.

## EVER-EST to Support the Research Lifecycle Management

Federica Foglini\* gave a presentation on the EVER-EST to Support the Research Lifecycle Management**.** Research lifecycle management begins with data collection, through re-use and scientific publication. It is a complex work flow, from ISI (International Scientific Indexing) papers, non-ISI and conference papers and reports (grey literature).

She described the ISI publication, with peer reviewed journals with impact factor and citation. She also described the non-ISI papers, such as conference proceedings, but with no DOI and citation index. Grey literature includes report and technical notes, project deliverables, report on data collection and processing (data life cycle).

The main challenges for a CNR scientist are searching for existing data and products, sharing methodologies on the same workflows and data, and adopting shared powerful tools for data processing. Today’s solution:

* searching of existing data and products among many different websites, colleagues and institutional partners
* sharing methodologies through description in scientific papers publication
* working on the same workflows and data mostly with colleagues in the same place and time (laboratory, workshops and meetings)
* adopting shared powerful tools for data processing only if are available in the laboratory

The EVER-EST solution is:

* Remotely access data, software, research results, and documentation
* Organize a scientific workflow in a single digital object, findable and reusable, maintaining attribution through DOI placement
* Collaborate with colleagues located in different parts of the world
* Document scientific work, e.g., encapsulate in a single digital object data and/or results related to a single Supersite event (an eruption)
* Publish grey literature (e.g., project reports, bulletins, etc.) maintaining attribution
* Ensure long term preservation of research work (data, software, results, interpretations)

Federica discussed the following use case: The sea monitoring community is wide and heterogeneous including both multi-disciplinary scientists, national/international agencies and authorities dealing with the adoption of a better way of measuring the quality of the environment. The objective of the Sea Monitoring VRC is to test the EVER-EST platform and provide useful and applicable contributions to the evaluation of the descriptors: D1.Biodiversity, D2.Non-indigenous species, D4. Marine Food Web and D6.Seafloor Integrity. The expected solution in EVER-EST is:

* Data sharing and harmonization- reduction of data and knowledge fragmentation.
* Easy data discovery re-use and re-purposing of open data
* Online data processing – resources and collaboration using a virtual lab
* Implementing fair data principles through the adoption of a research object able to encapsulate, share and reproduce the entire research cycle

Federica also discussed a use case involving cold water corals habitat, jellyfish distribution, correlation between jellyfish distribution and satellite data. Another use case is mapping human impact within lagoons using literature review of 125 papers. A fourth case study is preserving ancient map of the lagoon of Venice for assessing changes of human footprint.

Federica concluded with the following points:

* The EVER-EST project has demonstrated the relevance of research results (Research Objects - ROs) standardisation and interoperability to boost innovation and open science (FAIR principle).
* ROs complemented by Data and Publication DOIs enable the bi-directional link between the data and the research output results and assure the automatic recording and tracking of the quality of the research results and ROs.
* The functionality of Geo-referencing ROs proves invaluable for data providers to assess dataset valorisation requirements including historical maps ingestion to build long term data series from satellite images back to historical ground measurement (e.g. sea-monitoring data cubes, ARDs).

Chris asked if they have tried any re-use. Federica replied that the workflows are shared so this is a form of re-use, providing a powerful way of sharing and knowledge increasing, leaving knowledge in the form of research objects that can be reproduced by others.

## Knowledge Management System (KMS)

Rosemarie Leone\* gave a presentation on the Knowledge Management System (KMS). She began with a discussion of the ESA digital agenda for space information management policy. ESA has been leveraging more than 10 years of LTDP programme mission statements. Their standards and best practices define policy, guidelines, best practices and standards applicable to the longer-term data preservation, discovery and access strategy goals.

Rosemarie described the CCSDS open archive information system data archive and ingestion working group, and the Space Mission Preservation and Stewardship use cases green book proposal:

* Single Mission: EO Mission Preservation and Curation
* Sensors Families Missions Long Term Data Series and Fundamental Climate Records
* Space Mission and In situ Data Fusion
* Multi Space - Domain Mission Data Cross – Valorisation
* From Heritage Mission to New mission conception
* New space mission simulation and test bed

Rosemarie described EO mission preservation and curation graphically, highlighting the various personas. She listed the different experts for longer term data series and CDR and the involved elements in heritage mission archival.

The KMS components include OMNES, based on Open format FITS, open source, IIIF, Dublin Core, METS, and PREMIS. She descried the OMES ingestion, architecture, and workflow. It is very useful for collecting using a common annotation.

Rosemarie described how COGITO DISCOVER works. The COGITO solution reads a text like a human does, deriving meaning from the text as a whole using powerful semantics analysis; cognitive and contextual reasoning and machine learning; and not just applying text analytics and statistical measures. She described the high-level system architecture and workflow.

Rosemarie concluded that addressing the gaps by defining use cases will help determine how agencies are implementing the preservation guidelines, particularly the metadata.

Mirko noted that WGISS is working on a Best Practice for information. The OMNES tool implements the result of this BP, in which archiving for the long term all the supporting materials ingested is done. Andy wondered if WGCapD involvement would be helpful.

Chris commented that knowledge classification might be a potential candidate for WGISS technology exploration.

This is a topic for further cooperation to see if WGISS can support the work on the green book use cases.

**WGISS-46-10**: Rosemarie Leone and Ge Peng to exchange information and further elaborate on use cases to be included in a CCSDS Green Book. Rosemarie to drive input for CCSDS.

## Measuring EO Data Usage – Best Practices

Iolanda Maggio gave a presentation on Best Practices for measuring EO data usage. She began with an outline of the paper that is being developed, and listed the sources (documents, stakeholder roles, and external platforms).

The metrics can be considered from four viewpoints: EO Data Offer, Technologic and Platform, User Engagement, and Strategic and Programmatic. These consist of Earth Observation Data Offer Metrics, Web and Platform Metrics, and User Engagement and Satisfaction Metrics, and can be collected explicitly or implicitly.

Iolanda displayed an example of measuring reporting in terms of metric code, description, parameters to be captured, proposed relevance, and rationale, and described intra-agency and inter-agency use cases.

For inter-agency collaboration activities and projects include data format alignment, metadata improvement, time series generation, thematic platforms development, and data accessibility improvement and re-design.

Andy wondered if WGISS should do an analysis of the metrics that agencies are already collecting. A use case would be toward uniformity. In the long-term, considerations could be that data distribution will no longer be viable in the cloud. Mirko requested suggestions to do this, considering the shift would be to data utilization in the cloud. There will be a lot more categories about data utilized, such as ‘what is accessing the data and for what purpose’.

Since the document is ready and circulated already, it was suggested to remove the gold and platinum rating and replace it with a difficulty rating, continue to take comments, and then analyze. Regarding cloud metrics, they are not formulated enough yet. Different cloud systems may provide different metrics, so ideally the commonalities should be identified.

Mirko noted the applicability to CWIC, FedEO, and IDN.

Data archivists are still very interested in data volume. Ge Peng commented that downloads are a good indicator, but very dependent on domain. The impact also needs to be considered.

Rob suggested an alternative path – skip what agencies are doing, and develop an API that makes it useful to achieve. This would be less work than analyzing.

**WGISS-46-15**: WGISS members to provide comments on the User Metrics Best Practice.

**WGISS-46-16**: WGISS members to perform an analysis of status of implementation of User Metrics at respective organizations and provide input (status check) to Iolanda Maggio.

**WGISS-46-17**: WGISS CDA System Level Team (SLT) to consider/assess feasibility to implement the defined user metrics.

## CERN Archiving Centre

Jamie Shiers\* gave a presentation on data preservation at CERN. He began with a summary of the current state of long-term (multi-decade) data preservation at CERN, and plans for the future. He noted that the LHC experiments have open access policies and have released more than 1PB of data, along with the necessary software + VMs + documentation. He described the history of the data management, the LEP tape management, software, and documentation, and multiple efforts to preserve the data collected up until now.

The DPHEP Study Group produced a Blueprint document outlining the motivation for LTDP in HEP as well as major issues (funding, support), recognizing the urgent action is needed for LTDP in HEP. The preservation of the full capacity to do analysis is recommended such that new scientific output is made possible using the archived data.

Jamie discussed future directions:

* “Analysis preservation” and reproducibility of results are top priorities for Data Producers
* Solving the Open Data funding / resource problem
* Data Preservation for non-CERN experiments
* Completing ISO 16363 certification seen as important for CERN, possibly also ESFRIs + ?
* Tape alternatives? Concerns over size of Enterprise Tape market (we are a small player)
* Use of commercial or other Cloud Back-ends?
* PV2020 and “EIROforum” meetings

In summary, sustainable data preservation is a reality at CERN: 300PB with an outlook to 10EB. They are now addressing “next level” concerns, including reproducibility of results, Open Data at multi-PB level, potential use of (commercial) cloud storage etc. Hopefully, the community can strengthen its links, arrange regular technical meetings and make PV2020 “Simply the Best!”

Mirko noted that there would be a dedicated meeting, with vendors, at ESRIN, to discuss the types of tapes; perhaps WGISS should organize something similar, or a Tape Working Group should be formed, considering that tape technology is changing.

**WGISS-46-27**: DSIG to consider organizing a discussion/session/workshop regarding long term archival tapes and storage media. This could be part of PV2020; consider including vendors.

## ISO 19165-2 Standard Status

Hampapuram Ramapriyan\* gave a presentation the status of the ISO 19165-2 standard: Geographic information -- Preservation of digital data and metadata -- Part 2: Content specifications for Earth observation data and derived digital products. He noted that Part 1 was about the fundamentals.

Rama gave the background, and listed the steps for arriving at an ISO standard. The status of ISO 19165-2 is as follows:

* 00 – Preliminary
* 10 – Proposal – Approval through international balloting; Project Team members (experts from participating member countries) named to form Project Team (PT)
* 20 – Preparatory – New project is registered with Technical Committee (TC 211);
  + 20.20 Working Drafts (WD) exchanged with PT
  + 20.60 Close Comment Period
  + 20.99 WD approved for registration as Committee Draft (CD).
* Items 00 through 20.60 have been completed.
* Three working drafts exchanged with PT and comments accounted for (May 2018 face-to-face meeting and TC 211 Plenary in Copenhagen, Denmark). Drafts were more broadly circulated (beyond PT)
* Approval for registration as CD to be considered at November 2018 TC 211 Plenary in Wuhan, China.

Rama concluded with the table of contents of version 3 of the working draft, and welcomed comments and input from WGISS members.

Mirko commented that this standard is very relevant to WGISS, and it is important for this work to be associated with WGISS.

## Knowledge Sharing Opportunities: Conferences

Iolanda Maggio described the following two conferences that are knowledge sharing opportunities for the data stewardship community:

* PV2018
* The Living Planet

Mirko noted that PV2018 was excellent, well organized and very successful. Dieter added that WGCapD will participate in The Living Planet 2019, which will provide an opportunity for interaction.

# Data DISCOVERY and ACCESS

## WGISS Connected Data Assets Report

Yonsook Enloe summarized recent and future happenings of the WGISS Connected Data Access:

* NOAA datasets being registered/updated in the International Directory Network (IDN)
* EUMETSAT datasets being registered/updated in the IDN
* FedEO data collections being registered in the IDN; FedEO datasets registered in the IDN will be tagged. CEOS OS granule access to FedEO datasets in IDN will be directed to FedEO
* Essential Climate Variables (ECV) and COVERAGE (CEOS Ocean Variables Enabling Research and Applications for GEO) collections will be registered in the IDN (ESA is contacting data providers)
* Continuing integration with GEOSS
* Completed the CEOS OpenSearch Conformance Test document
* CWIC Smart Client Validator for CEOS OS conformance testing. This software is going through the Open Source process at NASA
* Currently the System Level Team (SLT) is writing the WGISS Client Partner Guide

Andy commented that all VCs and WGs seem to have data coordination activities and wondered if WGISS could take central coordination role. Yonsook commented that for smaller data providers the metadata is not as accessible, and requires more effort.

Recent collaborations:

* Essential Climate Variables 2.0: System Level Team analysis of what ECV 2.0 data is discoverable. Facilitate registration of ECV datasets in the IDN. Future: work towards getting granule search access to all ECV 2.0 datasets; Support ECV 3.0 datasets
* COVERAGE: System Level Team analysis of what key COVERAGE datasets are discoverable. Facilitate registration of COVERAGE datasets in the IDN. Future: work towards getting granule search access to all COVERAGE datasets
* Carbon Portal. Future: add additional keywords to GCMD keywords
* Improved IDN dataset registration:
  + NOAA datasets being registered in the IDN. New keywords will be added to the GCMD keywords
  + ESA team registering FedEO datasets in the IDN. More detailed error reports from tools. Better documentation
  + ESA is spearheading registering other European datasets (e.g. Copernicus) for COVERAGE, ECV.
  + NASA and ESA working together on automating the dataset registration/update process. Web Accessible Folders is one potential option

Towards a WGISS federation:

* Data collection registration at the IDN using the GCMD keywords; information about how granule search is supported will be included in the data collection registration; tag for granule access; propose new keywords to the GCMD Keywords set (if needed);
* Data partners need to support or be translated by CWIC or FedEO to WGISS supported standard: CEOS OpenSearch Best Practices (v 2)
* All searchable data must have a data access path: Data download, Data order (free or with cost), Email order (free or with cost)
* Servers must have high availability (99%?)
* A technical POC needed for each data partner

Proposed Data Partners Status webpage: The need for a status information webpage for all data partners is proposed on this page: [http://CEOS.org/wgiss-connected-data-assets-status/](http://ceos.org/wgiss-connected-data-assets-status/). It includes a table of all data partners who offer data to WGISS, have registered datasets in the IDN, provide granule search via CWIC, FedEO, or standalone server, and include a path to data access via data download, data order system, or email order. Status types: Development, Operational (Up or down); daily access testing needed to determine current accessibility!

Client Partners can offer search and access to all the satellite data available through the WGISS Federation. Clients can offer search and access to a limited subset of data available through the WGISS Federation along with other services through tagging (e.g. CWIC, FedEO, LSI, Carbon). Clients can offer support for a 2-step search; all granule search results will contain links to data access.

For tool developers the WGISS Connected Data Assets Client Partner Guide is offered. A new document is being written by the System Level Team, with details how to search for collection and granule data at IDN, CWIC, and FedEO. Expected completion date is first half of 2019. Outreach to potential tool developers includes a one page flyer targeting tool developers; a soft copy of the flyer will be kept on the WGISS website for local printing. The URL should also be advertised: [http://wgiss.CEOS.org/access](http://wgiss.ceos.org/access)

The current WGISS Work Plan Deliverables are:

* Produce How-To Guide to become WGISS Client Partner that describes how to connect and search for data in the IDN, FedEO, and CWIC
* Review and update FedEO How-To Guides for Data Partner
* Review and update CWIC How-To Guides for Data Partner

Yonsook wondered if the FedEO How-To Guide for Client Partners and the CWIC How-To Guide for Data Partners need review and update. Yonsook concluded describing the role of the System Level Team.

Rob asked how ECVs scale; users may want to tag datasets as useful for them, and tag them by user group. Yonsook noted that the tagging of ECVs is for larger groups. The technology exists to identify subsets through keywords alone but it does not scale.

It was noted that there is a lot of discussions around FDA and how to leverage the WGISS infrastructure. Chris added that getting around the security is a hurdle, and GEOSS may be interested in being empowered to do this.

## CWIC Report

Yonsook Enloe listed updates from the CWIC Data Partners:

* National Center for Remote Sensing of China – operational
* Global High Resolution Sea Surface Temperature (GHRSST)/NOAA - operational
* USGS/LSI – operational
* EUMETSAT - operational
* CCMEO - operational
* ISRO – 2 data centers connections are operational (MOSDAC and NRSC)
* INPE – operational (but currently experiencing outage)
* NCEI/NOAA – developing CWIC interface to NOAA One-Stop
* China GEOSS – new partner
* Australian data centers – on hold

Challenges for CWIC include easing dataset registration in the IDN from diverse data providers, evolving the CEOS OpenSearch BP, and incorporating data services within the WGISS Connected Data Assets. Additional challenges are portal access to a predefined subset of data and services interoperable and discoverable data services to discover and compare ARD products and analytic services.

## IDN Report

Michael Morahan gave a report on the IDN. He began with an overview of NASA’s Vision of Services (UMM-S), listing the required fields and the latest capabilities and features. These services provide a method of transforming the data (e.g. subsetting, re-projection or reformatting, or a combination of these).

Michael described SERF which is a legacy metadata standard and discussed the need for SERFs need to be migrated to UMM-S in CMR. Michael gave the list of what is being migrated, and the timeline.

The successor to DocBuilder is the Metadata Management Tool (MMT). Michael described the MMT data flow diagram, and explained the MMT manage collections mode and the draft collection mode.

Michael also discussed the status of the transition from DIF-9 to DIF-10, giving the status for each agency, and the schedule for UMM-C and GCMD/IDN proposed keyword revisions (new fields and topics).

Michael gave the status of two IDN actions from WGISS-45:

* Action: WGISS (Michael) to start defining best approach for representing and including QIs for the selected test case in discovery metadata searchable by end users: WGISS need one specific and one broader example of SST QI to start analysis. Awaiting input from WGCV
* Action: WGISS (Michael) to define how to get this info into the IDN for discover and possibly access by August: Will be started after receiving initial input by WGCV, completed by end November

Rob wondered about the encoding, given the UMM-S with containerization in the cloud, which is a slightly different model. Chris replied that “services” is misleading in the UMM-S. It is not yet known exactly what changes will be needed. Rob said knowing that it is located and works with a certain type of data.

Michael provided input to Iolanda for the Open Source Inventory using IDN content.

**WGISS-46-24**: WGISS-Exec to evolve the OSS inventory target from an inventory sheet to a mechanism for ongoing CEOS publishing and discovery of open source tools important to CEOS agencies. Assess for sustainability and commitment from CEOS agencies prior to development. [OSSW inventory should focus on tools/SW that CEOS wants to make visible for wider use. When published on WGISS web site they should be classified by: 1) Thematic area (problem they solve), datasets usable, targeted users, category (e. g. visualization, etc.) High relevant tools should be highlighted. Second steps will be to approve at WGISS level a service/tools metadata (see WGISS-46-22) and ensure that CEOS agencies register services/tools in IDN using it. Discovery will then be possible via IDN “CEOS branded” entry point]

## FedEO Update

Andrea Della Vecchia gave an update on the Federated Earth Observation (FedEO). He described the gateway system and how FedEO fits within the WGISS CDA. The current software refactoring objectives are optimization of the gateway and the catalog and of the dataset metadata ingestion job, porting all FedEO components to Docker and Kubernetes, and preserving all functional/interoperability requirements. Andrea discussed data ingestion, performance, time response, and results, noting that the new FedEO SW will be deployed at ESA beginning 2019.

Andrea next discussed the ESA collaborative environment to provide the ESA PDGS with a set of interoperable services permitting the users to access, discover, download EO data, discover and access basic, and hosted processing for authorized users/communities.

The ESA Catalogue shall be the centralized metadata repository of ESA Collaborative Environment, with reuse same FedEO SW. It will manage collections DOIs and be part of CEOS WGISS Data Assets via FedEO.

Andrea gave the status of FedEO metadata export into IDN: 20 ESA collections, providing two-step search, are in the IDN via the FedEO DIF-10 generator. The new FedEO Metadata Mediator – an automatic procedure beginning 2019 – will prepare and export the metadata. Andrea also listed the collections ready to be uploaded to IDN, and the partners where technical contacts need to be initiated.

Andrea discussed two open issues with DIF-10 in the areas of Metadata Preparation and NASA DIF-10 Validation.

He suggested these DIF next steps:

* ESA reports to NASA about understanding of mandatory/optional DIF-10 fields, and identified inconsistencies between DIF-10 Writer Guide and DIF-10 validators
* FedEO Ingestion tool shall be enhanced (Q1 2019) to generate a log file showing mapping information and let it generate a human readable “Ingestion report”. This will simplify the internal process of metadata owner to make the metadata IDN ready, passing through FedEO.
* Proceed with systematic European Partner metadata ingestion into FedEO and automatic export into IDN

Andrea presented two CEOS connected data assets scenarios:

* M2M Interface: APIs allow two steps search to external clients aligned to CEOS OpenSearch BP 1.2
* GUI Interface: CEOS Connected Data Access allow the users to discover/access both collections and products metadata

Andrea concluded with metrics of collections and granules of various organization datasets.

Yonsook noted appreciation for the work done to register the datasets. Mirko suggested improving the IDN front end so there is a single entry point, with a harmonized look and feel; this would be a good discussion in the SLT.

**WGISS-46-19**: WGISS CDA SLT to discuss way forward to implement a single front-end/portal within the IDN to access the WGISS CDA for data discovery and access (see proposal by Andrea Della Vecchia).

**WGISS-46-20**: Michael Morahan to provide a one page description of IDN entries (i.e. DIF-10 Writer Page points with fields/sub-fields and reference to external sources) and a DIF-10 metadata validator.

## ESA Pilot Datacube for TPMs/HMs/EEs

Andrea Della Vecchia gave a presentation on ESA Pilot Datacube for TPMs/HMs/EEs. He introduced the concept of an EO collaborative environment, which is a virtual working environment providing access to EO data and analysis tools, processors, information and communication technology resources required to work with them, through one coherent interface. He described examples of three current ones: Thematic Exploitation Platform, Mission Exploitation Platform, and Joint ESA-NASA Multi-Mission Analysis Platform. Andrea noted that a collaborative environment aims to change the paradigm from “data to user “to “user to data”, and transparent access to the content.

The ESA Collaborative environment seeks to provide the ESA PDGS with a set of interoperable and federated services permitting the users to:

* Access to missions/platforms information supported by a common ontology
* Discovery and, if possible, direct download of EO data: Copernicus, Third Party, and Heritage Missions, Earth Explorer, and International repositories.
* Discovery and access to basic services: browse/visualization tools and time series extraction, and EO data extraction, resampling and re-projection
* Hosted Processing for authorised users/communities (e.g., CAL/VAL)

Andrea displayed the architecture, and noted that bringing client to data is efficient at the Petabyte level with hosted processing. Standardised data access interfaces allow connecting a wide range of user interfaces. The deployment of DAS in front of each data source enables effective access services. Data remain at their own location (multiple data centers) with the original data format.

Andrea demonstrated the system using the CEOS CAL/VAL test case. Next steps are to include CEOS Radiometric Calibration Network to ingest ground measurements, and to retrieve multisource timeseries.

Next steps for 2019:

* TTO ESA Catalogue Q1 2019
* TTO ESA TPM/HM/EE Datacube Q2 2019
* Q4 ’18
  + Most of ESA TPM/HM/EE collections available
  + Preliminary support of CEOS WGCV and ESA Cal/Val use cases
  + Web UI available from ESA EO Landing Page
* Q2 ’19
  + Software refactoring to boost performance and manage concurrent users
  + Internal OpenSearch interface aligned with CEOS OS BP 1.2
  + All ESA TPM/HM/EE collections available
  + Full support to ESA Cal/Val initiatives (CEOS WGCV support need to be formalized)
  + Content synchronization with ESA Catalogue and FedEO
* Q4 ‘19
  + CEOS WGISS(/WGCV) international cooperation
  + Evolution about interoperability (e.g., connection with several Open Data Cube instances)

Yonsook commented that this is very interesting, and looks forward to this model.

In response to a question from Richard, Andrea explained that for different instantiation of server, you move the visualization information only. To work on just a subset, or a time series, you move the subset information on the web.

**WGISS-46-30**: Andrea Della Vecchia, Rob Woodcock, WGCV (Medhavy, Philippe Goryl, Cindy Ong) to discuss way forward on ongoing activities on data cube support to WGCV.

## Partner Reports

### EUMETSAT Report

Uwe Voges discussed the status of EUMETSAT integration into CEOS. He reported that they updated the list of collections provided for CWIC (including identifier, links to PN, CMR), and that collections were tested and documented. EUMETSAT is in the process of migrating from DIF9 to DIF10 using the XSD for DIF Version 10.2 b. He listed important changes related to DIF9, and updates done after first DIF10.

Uwe described the workflow for IDN/CWIC metadata (DIF10) provision at EUMETSAT, noting that search is possible on all collections stored in CMR/IDN, though the maximal temporal extent where products are available could only be requested via the collection MD.

Uwe described the EUMETSAT OpenSearch-EO Interface, which has been adapted with minor changes and is backwards compatible. Testing with CWIC Smart Client is no longer possible.

Uwe asked if the CWIC OpenSearch Adapters are updated to OGC OpenSearch-EOP 1.1 (13-026r9: offerings, GeoRSS Simple, error codes, links, parameters, ranges and sets), and if there are there plans on using GeoJSON binding.

Prospects for further development: When OLDA becomes operational, OLDA collections may be provided via a separate OpenSearch endpoint. Currently only one collection overlapping (EO:EUM:DAT:METOP:GLB-SST) -> s. EUMETSAT Collections for CWIC\_IDN and FEDEO V3\_1.xlsx.

### JAXA Data Portal

Yousuke Ikehata gave a status report on the JAXA data portal. He began with an overview of the new G-Portal in which three systems have been merged. The portal integrates catalog services, provides granule level search, supports several response formats, and is compatible with OpenSearch and CSW.

Features of the new system include: Processed data is expanded and immediately provided, and the new system archives processed products. The usability is improved. All the data obtained by the following satellites is processed and uploaded to the new system:

* MOS-1/1b(MESSR\*)
* JERS-1(VNIR, SWIR, SAR)
* ADEOS(AVNIR\*, OCTS)
* ADEOS-II(AMSR, GLI)

MODIS catalogs are harvested from CMR@NASA to fill the gap between ADEOS-II/GLI and SGLI. The data transfer performance is improved. The new portal integrates data access and catalog service, and contributes to GEOSS and CEOS portals by providing catalogs.

Yousuke reported that JAXA has already registered DIF-10, and FedEO collection level information were provided by email. Granule level information is harvested by CSW. There is now a direct connection with the GEOSS Portal.

Makoto suggested sharing information about new dataset availability since they have new data; Iolanda agreed to work with them on that. It was confirmed that the “New Dataset Availability” procedure also applies to re-processed data.

## CEOS OpenSearch Conformance Test Document

Andrea Della Vecchia gave a presentation of the CEOS OpenSearch Conformance Test Document. He listed its objectives, achievements, roadmap, responses from reviewers, and pending activities. Efforts center around the delivery of the CEOS Conformance Test SW: CEOS Conformance Test Software will permit to validate and score external OpenSearch endpoints vs CEOS guideline. This CEOS tool shall be on CEOS OpenSearch Conformance Test Plan guidelines. Within WGISS SLT working group the following has been discussed and agreed:

* NASA to keep OpenSearch-UI (former CWIC-SMART) On-line Validation tool updated with future CEOS OS BP changes (SLT MoM 01/03/2017) – Done
* NASA has plans to make all their software, including OpenSearch-UI On-line Validation tool, open source. Until that happens, the Validation is a publically accessible service (SLT MoM 01/03/2017) – Done
* ESA is available to support NASA to enhance and test OpenSearch-UI Validation tool, to produce a CEOS tool usable by all agencies (SLT MoM 01/03/2017). This also implies the application of an open source license for this Conformance Test Software (Sec. 4.2 WGISS#42 MoM)

Andrea listed the following Synergies with OGC initiatives:

* Requested to OGC to approve that the related document package enters the so-called “public review” phase (3 weeks). Public comments period was closed in August 2018 ([OGC call for public review](http://www.opengeospatial.org/pressroom/pressreleases/2831)).
* Current activity (September-November 2018) is assessment of public comments that were received.
* Documents package resolving the comments has to be delivered next to OGC for   
  final TC Vote for adoption as standard.

Short term next steps – Q1 2019:

* CEOS Conformance Test SW
* To share CEOS Conformance Test SW with OGC, within TeamEngine: <http://opengeospatial.github.io/teamengine/>

Long term next steps– by 2019:

* Search-Engine-Friendly with
  + schema.org (Annex C CEOS OS BP 1.2) to maximize the interaction/visibility with WWW search engine (e.g., [Google Dataset Search engine](https://toolbox.google.com/datasetsearch))
  + Using Microdata, for plain text formatting (e.g., ESA Landing Page)
* Once closed the standardization cycle of OGC 13-026r9 / OGC 17-047 / OGC 17-003, these might be adopted as baseline for a new review cycle of the CEOS Best Practice/Conformance Test Document, in order to:
  + Simplify significantly the CEOS guidelines (e.g., CEOS BP 1.2 is applicable in the new OGC specs)
  + Manage EO data services metadata (e.g., WMS, WCS, WPS, Datacube API, ordering, etc.)
  + Increase gateway/catalogue performance and the level of interoperability (e.g., ranking, flexible string research, light product metadata model, etc.)
  + Ease the GUI development and performance (e.g., moving toward GeoJSON(LD)- encoding)

Rob wondered if the conformance test could be automated. Andrea replied that it would have to be hosted.

**WGISS-46-25**: Andy Mitchell to notify WGISS when the CEOS Open Source Conformance Test Document is approved for Open Source.

## WGISS Connected Data Assets Client Guide Document

Archie Warnock gave a presentation on the WGISS/CDA Client Guide which introduces WGISS Connected Assets system and architecture to potential OpenSearch client developers. It is co-authored by CWIC, IDN and FedEO teams, and is in draft status with official release in 2019. It describes a two-step collection/granule search process, introduces CWIC, FedEO and IDN/GCMD, and reviews WGISS CDA architecture. It also describes the OpenSearch query interface, includes a guide to implementation and includes use cases.

The two-step search process and architecture are described; the search is only of collections registered in IDN. Collection responses contain direct links to CWIC or FedEO granule gateways as appropriate, and utilize CEOS OpenSearch Best Practices:

* Data Collection search via IDN
* Data Granule search via “Granule Gateways” (CWIC and FedEO)

The query interface and implementation interacts with the IDN for Collection searches, performs CWIC and FedEO for granule searches, and defines OpenSearch Requests and Atom responses. OSDDs define available search parameters, describe usage, and describe error handling.

The use cases are included:

* Platform Neutral: Examples for both CWIC and FedEO
* Retrieve IDN Dataset ID via OpenSearch: Example query and response
* Retrieve Granules via OpenSearch: Sample requests and responses (IDN -> FedEO)

# Data INTEROPERABILITY and USE

## CEOS Tasks Involving WGISS

### FDA-08: FDA Reference Architecture

Rob Woodcock gave an update on FDA-08: FDA Whitepaper. The task is to establish a common description of Future Data Architecture functional blocks and identify interfaces and interoperability approaches.

Multiple views per the Reference Model for Open Distributed Processing (RM-ODP) will be used. The RM-ODP framework factors an architecture description into the following five complementary viewpoints:

* Enterprise Viewpoint: (Why) defining the purpose, scope and policies of the system
* Information Viewpoint: (What) describing the semantics of information and information processing within the system
* Computational Viewpoint: (What) a decomposition of the system into computational interfaces
* Engineering Viewpoint: (What) describing the system infrastructure and mechanisms supporting distribution
* Technology Viewpoint: (How) a focus on technology choices to realise the system

Rob displayed the table of contents of the white paper and described the workflow: work out ways to talk about this more effectively. A decision facing the group is whether the white paper should be a long document, or if just a fact sheet is best; can FDA-08 focus only on what is changing and current gaps?

The list of contributors is low, and more are needed for a successful product, with the goal of elevating outputs from agency activities to the CEOS community level, WGISS information systems. It was suggested that SIT or other WGs might provide broader participation.

Chris added that these are good topics for webinars; some of these may help gather contributions.

Rob requested that participants email him with potential FDA themes.

Richard commented that H-2020 have done very interesting work which might be helpful. Regarding references to “from data to information” WGISS could play a role, perhaps in processing, orchestration and monitoring.

**WGISS-46-31**: Mirko Albani, Rob Woodcock, Iolanda Maggio, and Richard Moreno to organize an FDA workshop during the next WGISS meeting to demonstrate FDA elements capabilities (subset selected from FDA inventory) and practically address interoperability aspects to make the ecosystem more manageable for users.

**WGISS-46-32**: Chris Lynnes to investigate and write an information model in FDA (knowledge).

### FDA-09: Inventory and Characterise Existing FDAs

Iolanda Maggio gave a presentation on task FDA-09: Inventory and characterise existing FDAs operated by both public and private entities including the standards and approaches they use (e.g. Data Cubes, Exploitation Platforms, Copernicus DIAS, etc.).

Iolanda gave the background and illustrated the inventory table that is being developed; approximately 65% of the information has been completed. She displayed the EO themes and missions, and listed the four layers: data generation, resource, platform services, and exploitation; the largest is the platform services layer. The four functions are discovery, view, download, and processing.

The team will circulate the inventory and receive feedback from WGISS members by mid-November. Plans are to finalize version 1.0 in December.

Iolanda presented the following ideas for WGISS-47:

* Selection of a subset of FDA inventoried elements (e.g. Data Cubes, TEPs)
* Identification of a set of high priority use cases
* Organization of a 1-2 days hands-on workshop to:
  + Demonstrate capabilities/functionalities with respect to identified use cases
  + Identify commonalities between FDA elements and interfaces
  + Confirm mapping with respect to Common Reference Architecture
  + Define way forward for FDA elements contribution to CEOS objectives and stable visibility in the CEOS context
  + Address FDA elements interoperability aspects

It was noted that there is some overlap between tools and FDA inventories.

### FDA-10 Inventory of Software and Tools (Open Source)

Iolanda Maggio gave a presentation on FDA-10: Inventory of Software and Tools (Open Source). The description of task FDA-10 is to finalise inventory of software and tools available or used at CEPS agencies for EO data exploitation and use focusing on Open Source but remaining as broad and inclusive as possible and implement a mechanism for discovery and access.

Iolanda displayed selections of Open Source Software (OSS) in various WGISS agencies. Currently, the WGISS website has a page for OSS. This page will be moved into the Interoperability and Use Area; an inventory database will be implemented with search capabilities (November). Iolanda displayed a table with a compilation of results from an OSS Survey in CEOS Agencies; it compiles the reports with relevant metadata.

Iolanda suggested a use case for the OSS Inventory, with the goal of moving from user need to OSS Inventory utilization. The Sentinel Application Platform (SNAP) architecture is ideal for Earth Observation processing and analysis. She listed feature highlights and technologies and formats, and sample missions using the tool:

* NetBeans platform – desktop application framework
* Install4J – multi-platform installation builder
* GeoTools – geospatial tools library
* GDAL – reading/writing raster and vector geospatial data formats
* Jira – issue tracker
* Git – version control system, hosted by GitHub

Mirko reminded that it was suggested earlier to classify the inventory by thematic areas, and Kristi added that classification by function would also be helpful. The issue of sustainability will need to be addressed.

### WGCV-01: Data Formats and Interoperability

Kristi Kline gave a presentation on WGCV-01: Data Formats and Interoperability in the frame of FDA. She began with a description of Analysis Ready Data (ARD), which is data processed to a level that enables direct use in applications. Kristi explained that ARD allows

* geospatial, multi-spectral, and multi-temporal manipulations for the purposes of data reduction, analysis, and interpretation,
* consistent radiometric processing scaled to surface reflectance and surface temperature,
* consistent geometry including spatial coverage and cartographic projection – e.g., pixels align through time,
* metadata of sufficient detail on data provenance, geographic extent, scaling coefficients, and data type.

Kristi showed examples of Landsat scene-based products, highlighting natural orbital drifts. She noted that ARD facilitates time-series analysis. Users spend 80% of their time dealing with data (finding, cleaning, reorganizing). USGS wants to reduce this for the users.

Metadata can be created at a variety of levels. Landsat metadata can be grouped into three main tiers:

* collection-level metadata (preservation)
* granule-level metadata (technical/descriptive,
* and pixel-level metadata (technical/descriptive)

Landsat Collection 2 Changes include minor changes for Level-1 product metadata, normalize Level-2 product metadata to be similar to Level-1 product metadata, and provenance.

Kristi described the SpatioTemporal Asset Catalog (STAC) specification aims to standardize the way geospatial assets are exposed online and queried and demonstrated visualization which can only be done with ARD.

Kristi concluded saying that, by moving the data to the cloud, agencies can enhance the exploitation of the data. The goal of USGS is to have the data in multiple cloud provider systems, managed by the agency. USGS can reprocess data in the cloud in four days, a huge improvement over processing on premises. The processing is the affordable part of the cloud. Storage is not a cost benefit. This work is being driven not just by cost but by to increase access to the data.

Richard noted that COGS are compatible, and NetCDF is useful when the storage in the cloud as an archive.

### WGCV-03: CEOS Data Cubes and Test Site Data Access

Rob Woodcock gave a presentation on WGCV-03: CEOS Data Cubes and CEOS Test Sites Data Access in support to WGCV Activities. The objective of the task is to provide an improved user access to the LPV supersite data sets coupled with analytics tools to provide improved results for the users: Open Data Cube on Demand + Jupyter Notebook for Cal/Val = Ready-to-run.

Rob described the calibration and validation workflow, and notebook DEA with field data. The DEA reads like a report, but has live data. Rob displayed tables, graphs, and visualizations to illustrate the workflow. Next steps include clarifying the following points:

* ARD-on-Demand and field data equivalent – service composition is a problem.
* Field data-on-Demand.
* Elevating notebook to CEOS level – CEOS/WGISS Information Systems.

Rob concluded saying that what is missing is the middleware to connect the systems, in spite of different technology and different interfaces. Also missing is a practice to allow the different APIs to be connected to avoid the redundancy of data, working on different installations, minimizing the movement of data.

### CARB-15 Carbon Data Portal Prototype

Liping Di\* gave a presentation on CARB-15: Carbon Data Portal Prototype. The goal of the portal is to support the CEOS Carbon Science Mission by providing easy discovery of and access to carbon-related data resources in CEOS member agency collections, brokered by both CWIC and FedEO. The discovery should depend on keywords, spatial and/or temporal constraints, targeting at the CEOS Carbon Community

Eugene Yu gave a thorough demonstration of the portal prototype, showing collection level and granule level search and access within space and time constraints, predefined carbon regions, carbon topics, and GCMD keywords. Eugene demonstrated three scenarios: Discover, access, and display coverage (CWIC), discover, access, and display coverage (FedEO), and data collection found for Carbon using 2016 May wildfire of Fort McMurray, Alberta, Canada as a scenario. This scenario used as an example the analysis of monthly global Carbon Dioxide (CO2) product, demonstrating functionality of carbon related data searching, retrieving, processing, visualization, and analysis of CWIC Carbon portal. The demonstration slides are included in the presentation.

Mirko thanked NOAA and the team for the excellent work to support the activity; Mark Dowell is very happy with the portal. He added that the CEOS Chair Initiative for forest and rice monitoring in the Mekong area may benefit from the portal. Mark Dowell suggested exploring a case study in this context.

## Interoperability: GO FAIR, RDA

Richard Moreno gave a presentation on interoperability: GO FAIR, RDA. He began by listing and explaining in detail the FAIR principles: to be findable, accessible, interoperable and re-usable. Richard described the European Science Cloud Initiative (EOSC) which will allow for universal access to data and a new level playing field for EU researchers; a seamless environment enabling interdisciplinary research. GO FAIR is a bottom-up international approach. Richard explained the vision and strategy; it is a similar concept to the AGU.

Richard also described the Research Data Alliance (RDA), which has 93 working groups. Its vision is for research and innovation stakeholders to freely share the data, regardless of the technologies, disciplines and countries, in order to respond to major societal challenges. The RDA builds both social and technical gateways to allow free sharing of data.

Richard concluded with the question: Should WGISS follow what is done in the context of FAIR and RDA?

Ge Peng noted that she will be participating in the RDA FAIR Data Maturity Model Working Group, and she explained the process of joining working groups. Per Chris’s recommendation, Peng has agreed to update the progress to WGISS.

**WGISS-46-14**: WGISS-Exec to investigate the Research Data Alliance (RDA) working groups and identify those that may have relevance to the work of WGISS. Identify next steps.

# TECHNOLOGY EXPLORATION

## Interfacing Tools with Data

Chris Lynnes discussed interfacing tools with data.  He began with a tool called “Panoply” that uses the NetCDF library. He demonstrated an example, showing various features, such as working with the data remotely using OPeNDAP.

Chris noted that standards are wonderful; they allow Panoply to understand the geolocation coordinates of many files that follow the Climate Forecast (CF) convention. However, standards can be difficult to get exactly right. For example, an initial look at the Sentinel-1 Ocean Wind product in Panoply shows the tool unable to display the data on a map. However, the NCO tool turns out to be useful for “fixing” data by labelling the coordinate information in the metadata, allowing the data to be plotted on the map. A similar problem was found with Icesat-2 Lidar data, which can be fixed with NCO - a simple problem in the metadata to make the data visible on the map.

As a community it is desirable to figure out how to make this not happen with standards development. If not possible, how can the tools be used to correct the problems?

Chris urged the participants to download Panoply; it is free, easy to use, and great for students, scientists and managers (<https://www.giss.nasa.gov/tools/panoply/>

## FOSS Libraries and Tools to Raise EO Data and Applications

Homero Lozza\* gave a presentation on FOSS libraries and tools to raise EO data and applications. He began saying that, until middle of 2000s, the established paradigm dictated that space agencies were in charge of the data generation and responsible for the facilities and infrastructure (ground segment) to download the data. Users were well-trained experts that locally produced high level products. Agencies and users were accustomed to using proprietary software.

Space agencies began to provide free access to environmental and remotely sensed data of highest quality. GNU/Linux was very valuable for this change. A few examples include GDAL, PostGIS, and Q began their development. The OGC Foundation was formed in February 2006 and QGIS became an incubator project of the OSGeo Foundation in 2007. In 2005, Google launched and offered an API that allows maps to be embedded on third-party websites, making GIS popular and accessible to everybody.

All this led to a paradigm shift at the SAOCOM. In September 2006, the IDB approved a loan to Argentina to partially finance the Program for the Development of a Satellite System and Applications Based on Earth Observation. A novel aspect of the SAOCOM project was that, in parallel to the design and construction of the satellites, a set of applications dedicated to agriculture and hydrology were developed. The IDB showed that the simultaneous development of these three Strategic Applications would exceed the incurred costs for the project.

In order to fulfill requirements for IDB, CONAE faced the development of new interactive web products intended for final users, such as farmers and others decision makers, who had rarely been exposed to GIS and remote sensing. Homero displayed the SAOCOM Strategic Application architecture. Proprietary GIS neither suited a client-server architecture, nor had the flexibility and reliability for the development of operational web applications.

Although processor units are mainly implemented over Virtual Machines (VSphere), dependencies between different versions of libraries within a VM often interfere. Docker is used to run software packages called “containers”. Containers are isolated from each other and bundle their own tools, libraries and configuration files being lighter than VMs. Cloud computing with Docker is widely supported (Amazon EC2 Container Service, Google Compute Engine, Microsoft Azure). Conda is an open source, cross-platform, language-agnostic package manager and environment management system. The Conda package and environment manager is included in all versions of Anaconda and Miniconda.

Homero discussed the status of the SAOCOM mission, focused on soil moisture, supporting more than 20 projects, 10 community projects, and several initiatives. The FOSS4G annual conference regularly attracts over 1,000 practitioners.

The Open Science is a paradigm to make scientific research, data and results transparent and accessible for science and society. It consists of several tiers, including Open Access, Open Data and Open Source. Within geoscience, the new role model of “Open Science” or “Science 2.0” is emerging. Suited for the new trends are the following environments:

* R is great not only for doing statistics, but also for many other tasks, including GIS analysis and working with spatial data.
* Python offers similar facilities; it is just a matter of taste.
* There are many useful packages (for example, the GDAL/OGR bindings for R and Python are available through CRAN and PyPI, respectively).
* Reproducible Research is easier to achieve. The Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and much more.

Non-expert users expect higher level EO products to support decision making processes. FOSS foundations, standards organizations, companies, and many others boosted remote sensing and GIS, and revolutionized the way EO data is applied. Some FOSS projects have become a *de facto* standard. Agencies, experts and developers may need guidance to cope with the assorted options FOSS provides, and the rapid changing ITs. Best practices guides, recommendations and assistance will undoubtedly be welcomed by the community of EO.

## Google Dataset Search Tool Evaluation (DLR)

André Twele gave a presentation on an evaluation of the Google Dataset Search Tool. He began with a few quick facts of the tool for facilitating the discoverability of datasets from thousands of repositories across the web, relying on dataset providers to embed structured (meta-)data into their websites using schema.org dataset or equivalent structures (W3C DCAT) for mark-up.

First impressions: For many search results, only a “Description”-field (equiv. to *gmd:abstract* in ISO) is shown. For some search results, only “Dataset published, created or updated”, “Dataset provided by”, “Time period covered”, etc. is shown. There is rarely information on spatial properties, geographical coverage, etc. The dataset search can detect if a dataset is present in more than one repository.

The DLR EOC catalogue, which currently contains 184 entries (ISO 19115/19139), was evaluated. From a snapshot of 20 catalogue entries, 18 entries were discoverable through Google Dataset Search. However, DLR currently does not provide metadata which can be directly processed by Google’s Search Engine (🡪 schema.org/DCAT).

Selected ESA and EUMETSAT collections have all been retrievable through Google Dataset Search. Most frequent metadata sources: geo.spacebel.be, fedeo.esa.int, cmr.earthdata.nasa.gov, data.nasa.gov, [www.europeandataportal.eu](http://www.europeandataportal.eu). Ranking of search results is sometimes questionable: single value-added datasets are ranked higher than the original collection/dataset series. Level of detail for individual search results is quite heterogeneous.

The conclusions reached are that completeness and quality of search results strongly depends on structured metadata added by dataset providers on their sites. Adoption of open standards for describing structured data (schema.org, DCAT, JSON-LD) by Google will further encourage their usage and boost their adoption. Availability of metadata replicas in several catalogues indexed by Google makes it sometimes hard to retrieve the original dataset from its portal. The tendency is that metadata properties/attributes get lost in translation as a result of different schema transformations or portals through which they are offered.

Uwe Vogues commented that Google as a search engine is not reliable because they may not index the data for very large datasets; instead, it should be annotated in html.

## Google Dataset Search Tool Evaluation (NASA)

Doug Newman\* gave a presentation on NASA’s Google Dataset Search Tool evaluation describing the work CMR has done to leverage Google Dataset Search capabilities with respect to CMR HTML landing pages.

NASA uses RDFa tags in the html markup of a CMR landing page to convey semantic information to a search engine.

Key aspects of a collection for discovery that can be conveyed include collection name, version, description, spatial and temporal extent, and science keywords. Usage and provenance information, such as DOI, citations, provider, and curation information can also be conveyed.

Doug explained that the collection landing pages have a link to a granule search, scoped to that collection, in Earthdata Search. This presents opportunities for linkage between a commercial search engine like Google and NASA’s ‘search engine’. Schema.org provides a framework for efficient and intuitive user experiences in the form of the search action concept. A feature in Earthdata Search allows the user to transfer its context to other clients that cater for that collection.

Smart handoffs between Google Dataset Search and Earthdata Search are being explored. A search action embedded in the CMR HTML landing pages gives Google Dataset Search enough information to transfer the textual search to an EDSC context. If a user searched for ‘MOD02QKM Beijing summer 2010’ in Google Dataset Search and then clicked on the CMR result, that link could take the user to EDSC with context intact.

**WGISS-46-33**: WGISS (Doug Newman, André Twele, Michael Morahan, Andrea Della Vecchia) to draft a short summary analysis on Google Data Search Tool capabilities with respect to CEOS agency’s data and differences/complementarity with respect to WGISS CDA Infrastructure Include recommended actions, if any, to be carried out on WGISS side (e. g. in IDN).

## DATA-14: White Paper on Single Sign-On (SSO) Authentication

Marco Leonardi\* gave a presentation on single sign-on authentication. He began with a description of EO SSO at ESA, showing the architecture. He continued with achievements in 2018, such as a successful Cloud services access pilot, a successful ESA Earth Observation federation pilot, and a successful federation pilot between space organizations.

The ESA plan for standard and interoperable authentication solutions includes:

* ESA is performing an evolution in its user and identity management infrastructure aiming at standardising architectures and processes in line with the results of the most recent initiatives in this field like in example the AARC Blueprint Architecture.
* The new ESA Earth Observation Identity and Access Management Infrastructure (EO-IAM) will allow user access to satellite data and to the Exploitation Platforms’ services by supporting standard digital identities federations (and interfederations like eduGAIN).
* The new ESA EO-IAM will be able to make the federated user identification an enabler for the Exploitation Platforms in the context of the Network of EO Resources.

The DATA-14 white paper on SSO authentication aims to promote the best practices for the (federated) single sign on authentication. International working groups and research organizations have been working on this topic for many years with the objective of creating a shared approach to the identity and access management needs. “Interoperability”is one of the main drivers and «federation» is one of the most promising solutions. Interoperable federations for single sign on authentication need to share best prectices and standard architectures. International space organizations and scientific communities active in the field of Earth Observation can cooperate in order to improve the way the EO Data resources can be accessed by the users.

Kristi asked about privacy information. Marco replied that ESA is obligated to apply the GDPR rules. In eduGAIN this is already well considered and handled so it serves as a basis.

Andy wondered what the next steps should be; Kristi would focus more on registration than on authentication. Marco suggested collecting high-level requirements toward an operational concept.

**WGISS-46-23**: Marco Leonardi to trigger a request for information on an SSO White Paper to different partners including NEXT-GEOSS (Bente Bye).

## Artificial Intelligence for Earth Observation - Towards an R&I

Sveinung Loekken gave a presentation on Artificial Intelligence for Earth Observation (AI4EO). He introduced the topic saying that the community is now at a cross road of opportunities, where on the one hand AI is becoming one of the most transformative technologies of the 21st century, while European EO capability and the overall observing system is delivering a unique and comprehensive picture of the planet, generating massive, open data sets. Making the most of this window of opportunity is the challenge, and an urgent one.

Changes in the technologies, tools, operations concepts, and business models addressing Big Data. These include Cloud and Fog computing (XaaS); Platforms / user to the data; Open source frameworks and tools; Distributed value chains; Distributed ledgers; IoT; IDEs; Analysis ready data and advanced data analytics: AI, ML, DL.

Sveinung outlined a few of the challenges, such as data volume, diversity, and complexity; exploitation of AI must be in line with the observing system.

Community recommendations are to conduct exploratory, demonstration, and capacity building activities. There is a timely need for European initiatives to enable rapid transfer of AI knowledge, techniques and expertise from data scientists to the world of EO research (in both directions) and business applications, and foster new partnerships with non-space and ICT players.

Now is a unique and timely opportunity now to play a major enabling role in the scientific and societal AI revolution that is underway by enabling AI4EO at a large European scale.

## FDA topics going forward

Rob Woodcock the following FDA topics for present and future consideration:

* User experience – how? (WGISS Chair report)
  + Work with WGCapD – better access to broad users through training materials and feedback?
  + Instant gratification – working code/tools?
* Maturity assessment (WMO)
  + An FDA Self-assessment tool?
* Do we need an IGSN/DOI/equivalent for CEOS to track provenance and use of data regardless of replication (EVER-EST)?
  + What else would WGISS need to do to support these types of systems?
* Metrics system API (Measuring EO data usage)
  + Is a CEOS API for the metrics data gathering and contribution needed?
  + Metrics collection in Cloud?
  + Impact metrics – DOIs and an h-index for Data? Does CEOS (FDA) create the necessary ecosystem for a Journal/end user to create impact metrics CEOS can use?
* Robustness of the CEOS information systems - failover in a distributed system (CEOS SEO – Cove tool)
  + At least a monitoring tool should be in place to show CEOS Information Systems status?
  + A cache of the metadata for performance and failover?
* Automating Governance (Connected Data Assets…)
  + Service/data source/metadata out of date, notifying managers, etc. – robots to help run our world?
  + Scaling discovery annotation for non-CEOS user-communities. Can CEOS services support/not prevent this?

# Agency and Liaison Reports

## Hungarian Space Office Liaison Report

Gabor Remetey\* gave a report on the Hungarian Space Office (HSO). He described the HSO and EO activities in Hungary, including the Land Parcel Identification System (MEPAR) and Control of Agricultural Subsidies with Remote Sensing (TámELL), Agricultural Risk Management System (MKR), Country-wide mapping and status assessment of ecosystem services (NÖSZTÉP).

Gabor described the National Reference Institution of Land Cover Mapping and listed key players and contributors in EO in Hungary, such as educational institutions and research centers.

Gabor discussed promoting the potential and challenges related to the use of Geospatial Data and Earth Observation for support the achievement of the Sustainable Development Goals, listing domestic and cross-border events, and sharing and outreach activities. He also described the project on Earth Observation Information System (FIR).

Gabor concluded by describing the impact of the awareness actions, which are empowering and motivating students, start-ups, young professionals and engagement of stakeholders. Impacts also include increasing the number of interdisciplinary (Nexus) approaches, GI/EO4SDG issues embedded in the higher education curricula, improved links between the statistical and geospatial communities, and more effective development and use of EO/GI in the UN SDG target and indicator monitoring and yearly national reporting.

## NASA Agency Report

Andy Mitchell gave a report on the National Aeronautics and Space Administration (NASA). He described the Earth Science Data Operations, adding that the Earth Observing System Data and Information System (EOSDIS) is a key core capability in NASA’s Earth Science Data Systems (ESDS) Program. It provides end-to-end capabilities for managing NASA’s Earth science data from various sources – satellites, aircraft, field measurements, and various other programs. These capabilities include distributed data archives, data processing and management, access and user services, network data transport and promotion. There is no period of exclusive access, and data is available at no cost to all users on a non-discriminatory basis, except where agreed upon with international partners.

Andy displayed the EOSDIS components and data collection, and listed NASA’s missions through 2023, describing recent and near-term planned launches and resulting expected data volumes, highlighting the size of SWOT and NISAR data. He also described the current architecture for data centers, giving positive and negative characteristics.

Andy noted that ESDIS sponsored an independent review, which found that commercial cloud environments offer potential for storage, processing and operations efficiencies; improved cross-DAAC collaboration; and potential for new data access and service paradigms. The recommendation is that ESDIS Project should develop, implement and report on the outcome of prototypes to explore the advantages, risks, and costs of using commercial cloud environments for storage and data transfer, processing, and improved data access. A simplified commercial cloud architecture was displayed, along with the EOSDIS Cloud Evolution Project achievements, goals, and milestones.

Andy concluded saying that NASA’s policy is open data, open source, and open services; he concluded with archive size and data distribution metrics.

## ESA Agency Report

Mirko Albani gave a report on the European Space Agency (ESA). He discussed the ESA-developed EO missions for science, Copernicus, and meteorology, noting the successful launch of the seventh Sentinel (Sentinel-3B), and the Aeolus wind satellite. Mirko listed key features of the six Sentinels.

Mirko noted that ESA offers free access for all users to Sentinel products (most recent as well as complete long term archive) and any user can self-register at sentinels.copernicus.eu.

ESA delivers on 24/7 basis Near Real Time products (3 hours from sensing) as well as Non Time Critical products (24 hours from sensing). Mirko gave global statistics of Sentinels data access and distribution. He also highlighted Sentinel results for nitrogen dioxide and the Palu earthquake, and listed statistics of ESA principal investigators projects.

Mirko concluded noting the 50th anniversary of ESRIN, and announcing the Living Planet Symposium 2019.

## CSIRO Agency Report

Rob Woodcock gave a report on the Commonwealth Scientific and Industrial Research Organisation (CSIRO). He noted the successful launch of NovaSAR in September, and presented their ground station architecture.

Rob gave the roadmap for Australia, which includes CSIROSat-1 with CubeSat and 3U SWIR and on-board processing. CSIRO is also participating in small business engagement, the EO for Pacific Workshop, the CSIRO Data Cube, and the New Zealand Centre for Space Science Technology-CSIRO.

## CNES Agency Report

Richard Moreno gave a report on the Centre National d’Etudes Spatiales (CNES). He introduced the French Research Infrastructure ‘Data & Services for the Earth System’, giving the context, goals, structure and organization, and infrastructure.

Richard also discussed the Space Climate Observatory (SCO), a world observatory of the climate change and its impacts from Earth Observation data. Richard described the international framework, noting that the French government identifies 12 key One Planet commitments. The scope of the SCO is monitoring climate change, tracking the impacts of climate change, and mitigating and adapting to climate change. Richard concluded with the schedule for the SCO through 2023.

## USGS Agency Report

Kristi Kline gave a report on the United States Geological Survey (USGS). She began with a description of the Sustainable Land Imaging (SLI) Program, an agreement between NASA and USGS in 2016 that enables the development of a multi-decade, space borne system that will provide users worldwide with high-quality, global, land-imaging measurements compatible with the existing record since 1972.

Kristi described Landsat data continuity since 1972, with volume increasing now to 1 TB/day. She also gave statistics of annual data distribution and the size of the data holdings.

Kristi discussed the USGS Cloud Project, giving its key objectives and scope. Currently USGS contracts with AWS (US-West Region) for Cloud Services. She concluded with a description of the workflow for processing Landsat from L1 to L2 to COGs to Viewer.

**WGISS-46-34**: WGISS-Exec to consider having a special session on cloud services at WGISS-47.

## JAXA Agency Report

Makoto Natsuisaka gave a report on the Japanese Aerospace Exploration Agency (JAXA). He began with a description of JAXA’s past, current and future EO satellites and sensors, and the JAXA EO strategy. The “Unified Climate Change Program” will focus on greenhouse gases observation, Global Satellite Precipitation Map (GSMaP), and Global Forest Monitoring. JAXA will promote related activities through cooperation with CEOS. Makoto outlined the elements of JAXA’s cooperation with CEOS.

Makoto described Greenhouse Gases Observing SATellite-2 (GOSAT-2) “IBUKI-2” for global monitoring, describing the observation targets, instruments, and other specifications.

Makoto also described diagrammatically the ground segments for the JAXA EO satellites, the EO Portals, and the connections with GEO/CEOS portals (IDN, FedEO).

The GCOM-C product release will be in December, via the G-Portal, and METI is developing open and free platform for EO data “Tellus”.

Makoto concluded with JAXA’s long term plan.

# WGISS Plenary, Part II

## Future Meetings

Rob Woodcock announced that the WGISS-47 meeting would be held in Silver Spring, Maryland (MD), USA, from April 29 to May 2, and would be hosted by NOAA. The meeting will be held at Silver Spring Civic Building. Transportation and accommodation details were given, along with a tentative agenda. Information can be found at [http://CEOS.org/meetings/wgiss-47/](http://ceos.org/meetings/wgiss-47/).

WGISS-48 is tentatively set for September/October 2019 in Asia.

## WGISS Summary

Mirko Albani summarized WGISS-46 highlighting the following sessions:

* Plenary session:
  + CEOS Plenary
  + SEO Report: Coverage tool, Data Cubes, Collaboration with WGISS
  + WGCapD Report: Deliverables, collaborations, training calendar and resources, collaboration with WGISS
  + CEO Report
  + WISP Report
  + WGISS Vice Chair (October 2019 – October 2021)
  + WGISS brochure
* Data Preservation and Stewardship session
  + WGISS Data Stewardship Reference Model
  + The structure of SMM-CD
  + Measuring Earth Observation Data Usage - Best Practices
  + CERN - LEP Era: Tape Archive
  + ISO 19165-2 Status
  + Living Planet 2019 Conference
  + PV2020
* Data Discovery and Access session
  + WGISS Connected Data Assets architecture and elements
  + WGISS CDA GUI interface
  + EUMETSAT and JAXA connectivity reports
  + ESA TPM/HM Data Cube
* Technology Exploration session
  + Single sign on
  + FOSS
  + AI
  + Example “path” of metadata from its origin
* GEOSS – WGISS Interoperability workshop
  + Workshop objectives
  + Ongoing CEOS WGISS contribution to GEOSS
  + WGISS CDA architecture
  + GEOSS Platform CWIC/FedEO discoverable data product metrics
  + Inventories: 1) FDA Elements; 2) Open Source SW and Tools. Consensus that an inventory of “EO use environments” is a useful exercise that will provide insight into the EO ecosystem
  + Innovation rate: Rate of technology change vs rate of CEOS (WGISS) change?
  + Clearly, extraordinary EO analytics scales can be achieved for multiple domains and applications across multiple CEOS data sets – at what cost? Can CEOS make it cheaper and faster
  + GEOSS vision
  + GEOSS Platform
  + GEOSS Regional Initiatives: AmeriGEOSS, AfriGEOSS, EuroGEOSS
  + GEOSS Evolve initiative architecture
  + WGISS contribution to DMP evolution
* Data Interoperability and Use session
  + WGISS actions under FDA, WGCV, and CARB
  + Landsat Cloud Optimized GEOTIFF (COG)
  + Carbon Portal endpoints
  + Research Data Alliance (RDA)
* Agency reports
* CEOS upcoming meetings

## WGISS-45 Actions

The following WGISS-45 actions are still ongoing:

**WGISS-45-7**: Yonsook Enloe, Richard Moreno, Rob Woodcock, Mirko Albani, and Chris Lynnes to identify and submit to Michelle the valid versions of their Best Practice and White Papers. Michelle to display these on the Document Management Table on the WGISS website by end of June 2018.

**WGISS-45-8**: Mirko Albani, Iolanda Maggio and Michelle Piepgrass to address Best Practice and White Papers (not lessons learned) Web page (to be put at highest level in CEOS Web Site) population with other WGs material after completion of action on WGISS Interest Groups. Maintain and populate page by September, 2018.

**WGISS-45-13**: Michael Morahan to follow up with Greg Stensaas to update the test site information in the IDN by end of June.

**WGISS-45-15**: Esther Conway, Andrea Della Vecchia, and Michael Morahan to ensure that UKSA’s data collections are accessible through FedEO.

The following actions resulted from the meeting.

**WGISS-46-1**: Mirko Albani to bring to the attention of the SEO and the SIT the need for CEOS-globalized (e. g. VCs and WGs) WISP support; suggest this could be provided through the SEO.

**WGISS-46-2**: Mirko Albani to inform the SIT Chair that CSIRO has confirmed nomination of Rob Woodcock as WGISS representative in the GEOSS Expert Advisory Group. Mirko Albani and Andy Mitchell will support Rob as needed. WGISS suggests that a letter be sent by the SIT Chair to the GEO Secretariat.

**WGISS-46-3**: Michelle Piepgrass to send Terms of Reference (ToR) to Rob Woodcock and Mirko Albani. WGISS-Exec to update the WGISS ToR to reflect FDA activities assigned to WGISS and recommendations from the CEOS SIT and Plenary.

**WGISS-46-4**: WGISS interest group leads to provide information for the ‘CEOS Agencies Participation/Contribution to WGISS’ table (a table showing the different WGISS groups, their scope, CEOS agency interest/benefits, and skills of representatives), and to finalize, distribute and post.

**WGISS-46-5**: WGISS-Exec to identify potential WGISS Vice-chair candidates for period 2019-21 (to become Chair in 2021-23) and determine availability.

**WGISS-46-6**: Interest Group leads to send final comments on the WGISS brochure to Michelle Piepgrass and Iolanda Maggio.

**WGISS-46-7**: WGISS-Exec (Mirko Albani, Rob Woodcock, Andrea Della Vecchia, Yonsook Enloe, and Liping Di) to explore the possibility of a Carbon Portal case study built around the VAST-VNSC Chair Initiative. After confirmation, WGISS-Exec will get in touch with the SEO and VNSC as needed.

**WGISS-46-8**: WGISS-Exec (Richard Woodcock, Iolanda Maggio) to further explore possible areas of cooperation with WGDisasters, specifically with regards to the generic Recovery Observatory and FDA/CDA.

**WGISS-46-9**: Mirko Albani, Richard Moreno, Ge Peng, Iolanda Maggio to further elaborate possible collaboration with the WMO International Expert Group on Climate Data Modernisation (IEG-CDM), WIS and WIGOS, and possibly organize a joint session at WGISS-47.

**WGISS-46-10**: Rosemarie Leone and Ge Peng to exchange information and further elaborate on use cases to be included in a CCSDS Green Book. Rosemarie to drive input for CCSDS.

**WGISS-46-11**: Rob Woodcock to follow up on a SAR training/meeting in Vietnam with WGCapD/CSIRO.

**WGISS-46-12**: Iolanda Maggio to communicate with NextGEOSS, AmeriGEOSS, AfriGEOSS and EuroGEOSS about joining the WGISS inventories (FDA and/or SW/Tools).

**WGISS-46-13**: Chris Lynnes, Mirko Albani and Rob Woodcock to put together one page on recommendations for cooperation between GEO and WGISS based on the outcomes of the WGISS-46 joint workshop.

**WGISS-46-14**: WGISS-Exec to investigate the Research Data Alliance (RDA) working groups and identify those that may have relevance to the work of WGISS. Identify next steps.

**WGISS-46-15**: WGISS members to provide comments on the User Metrics Best Practice.

**WGISS-46-16**: WGISS members to perform an analysis of status of implementation of User Metrics at respective organizations and provide input (status check) to Iolanda Maggio.

**WGISS-46-17**: WGISS CDA System Level Team (SLT) to consider/assess feasibility to implement the defined user metrics.

**WGISS-46-18**: Mirko Albani to liaise with the CEO to trigger update of the inventory of Virtual Constellations datasets and then assess discoverability/accessibility through the WGISS Connected Data Assets infrastructure.

**WGISS-46-19**: WGISS CDA SLT to discuss way forward to implement a single front-end/portal within the IDN to access the WGISS CDA for data discovery and access (see proposal by Andrea Della Vecchia).

**WGISS-46-20**: Michael Morahan to provide a one page description of IDN entries (i.e. DIF-10 Writer Page points with fields/sub-fields and reference to external sources) and a DIF-10 metadata validator.

**WGISS-46-21**: WGISS-Exec to consider organizing a session at the next WGISS-47 on EO ontologies as part of the DSIG.

**WGISS-46-22**: Michael Morahan, Andrea Della Vecchia, Liping Di to organize a meeting to discuss a metadata model for services description to be circulated and approved by WGISS Model will be used to describe services which will then be discoverable through the IDN SERF interface.

**WGISS-46-23**: Marco Leonardi to trigger a request for information on an SSO White Paper to different partners including NEXT-GEOSS (Bente Bye).

**WGISS-46-24**: WGISS-Exec to evolve the OSS inventory target from an inventory sheet to a mechanism for ongoing CEOS publishing and discovery of open source tools important to CEOS agencies. Assess for sustainability and commitment from CEOS agencies prior to development. [OSSW inventory should focus on tools/SW that CEOS wants to make visible for wider use. When published on WGISS web site they should be classified by: 1) Thematic area (problem they solve), datasets usable, targeted users, category (e. g. visualization, etc.) High relevant tools should be highlighted. Second steps will be to approve at WGISS level a service/tools metadata (see WGISS-46-22) and ensure that CEOS agencies register services/tools in IDN using it. Discovery will then be possible via IDN “CEOS branded” entry point]

**WGISS-46-25**: Andy Mitchell to notify WGISS when the CEOS Open Source Conformance Test Document is approved for Open Source.

**WGISS-46-26**: Iolanda Maggio to prepare a proposal for GEO DMP improvement/update based on the work done by WGISS (e. g. Maturity Matrix).

**WGISS-46-27**: DSIG to consider organizing a discussion/session/workshop regarding long term archival tapes and storage media. This could be part of PV2020; consider including vendors.

**WGISS-46-28**: WGISS-Exec (Rob Woodcock, Andrea Della Vecchia) to prepare one to two slides for the SEO on what WGISS is doing for Data Cubes and propose a way forward for cooperation between WGISS and SEO on the FDA Data Cube topic.

**WGISS-46-29**: Mirko Albani and WGISS-Exec to prepare an updated set of actions for the CEOS 2019-21 Work Plan to address the mandate on FDA (see FDA AHT recommendations) and other WGISS related activities.

**WGISS-46-30**: Andrea Della Vecchia, Rob Woodcock, WGCV (Medhavy, Philippe Goryl, Cindy Ong) to discuss way forward on ongoing activities on data cube support to WGCV.

**WGISS-46-31**: Mirko Albani, Rob Woodcock, Iolanda Maggio, and Richard Moreno to organize an FDA workshop during the next WGISS meeting to demonstrate FDA elements capabilities (subset selected from FDA inventory) and practically address interoperability aspects to make the ecosystem more manageable for users.

**WGISS-46-32**: Chris Lynnes to investigate and write an information model in FDA (knowledge).

**WGISS-46-33**: WGISS (Doug Newman, André Twele, Michael Morahan, Andrea Della Vecchia) to draft a short summary analysis on Google Data Search Tool capabilities with respect to CEOS agency’s data and differences/complementarity with respect to WGISS CDA Infrastructure Include recommended actions, if any, to be carried out on WGISS side (e. g. in IDN).

**WGISS-46-34**: WGISS-Exec to consider having a special session on cloud services at WGISS-47.

## Concluding Remarks

Mirko Albani concluded the meeting thanking DLR for excellent hosting; logistics, facilities, and activities were fabulous. Mirko also thanked the participants for their contributions toward an excellent meeting.

.

# Glossary of Acronyms

API Application Programming Interface

ARD Analysis Ready Data

AWS Amazon Web Services

CEO CEOS Executive Officer

CEOS Committee on Earth Observation Satellites

COTS Commercial Off-the-Shelf

CSW Catalogue Service for the Web

CWIC CEOS WGISS Integrated Catalogue

DAAC Distributed Active Archive Center

DC data cube

DIF Directory Interchange Format

DOI Digital Object Identifier

ECV Essential Climate Variable

EO Earth Observation

ESIP Federation of Earth Science Information Partners

GCI GEOSS Common Infrastructure

GCMD Global Change Master Directory

GEO Group on Earth Observations

GEO-GLAM Global Agricultural Monitoring

GEOSS Global Earth Observation System of Systems

GFOI Global Forest Observations Initiative

GHG Greenhouse Gas

GIS Geospatial Information System

GPM Global Precipitation Mission

GPU Graphics Processing Unit

GSDI Global Spatial Data Infrastructure

GUI Graphical User Interface

HPC High Performance Computing

ICT Information and Communication Technology

IDN International Directory Network

ISO International Standards Organization

LSI Land Surface Imaging

LTO Linear Tape-Open

MOU Memorandum of Understanding

NRT Near real-time

NWIP New Work Item Proposal

OGC Open Geospatial Consortium

PI Persistent Identifier

POC Point of Contact

RSS Rich Site Summary

SEO Systems Engineering Office

SDCG Space Data Coordination Group

SIT Strategic Implementation Team

SLT System Level Team

SWG Standards Working Group.

TOA Top of the Atmosphere

ToR Terms of Reference

UML Unified Modelling Language

UMM Unified Metadata Model

VC Virtual Constellation

WCS Web Coverage Service

WG Working Group

WGCV Working Group on Calibration and Validation

WGCapD Working Group on Capacity Building & Data Democracy

WGClimate Working Group on Climate

WGDisasters Working Group on Disasters