MINUTES

OF THE

47th MEETING

OF THE

CEOS WORKING GROUP ON   
INFORMATION SYSTEMS AND SERVICES

(WGISS)

Silver Spring, Maryland, USA

29 April to May 2, 2019

Hosted by

National Oceanic and Atmospheric Administration (NOAA)

Table of Contents

[1 WGISS Plenary Session, Part I 7](#_Toc8597653)

[1.1 Introduction 7](#_Toc8597654)

[1.2 Logistics 7](#_Toc8597655)

[1.3 Host Welcome 7](#_Toc8597656)

[1.4 National Centers for Environmental Information (NCEI) Overview 8](#_Toc8597657)

[1.5 Introduction to NCEI 8](#_Toc8597658)

[1.6 WISP Report 9](#_Toc8597659)

[1.7 WGISS Chair Report 9](#_Toc8597660)

[1.8 Review of CEOS 2019-21 Work Plan and WGISS 2018-2020 Work Plan 10](#_Toc8597661)

[1.9 CEOS Executive Officer (CEO) Report 11](#_Toc8597662)

[1.10 SEO Report 11](#_Toc8597663)

[1.11 CEOS Working Groups Cooperation with WGISS 12](#_Toc8597664)

[1.11.1 WGCV 12](#_Toc8597665)

[1.11.2 WGCapD Report 12](#_Toc8597666)

[1.11.3 WGClimate 13](#_Toc8597667)

[1.11.4 Sustainable Development Goals Ad-hoc Team (SDG AHT) 14](#_Toc8597668)

[1.12 GEO Secretariat Report 14](#_Toc8597669)

[1.13 Future Meetings 15](#_Toc8597670)

[2 Future Data Architectures Elements Demonstration and Interoperability Workshop 16](#_Toc8597671)

[2.1 Introduction and Scope 16](#_Toc8597673)

[2.2 GEO AquaWatch 16](#_Toc8597674)

[2.3 CEOS Client Interfaces and Open Data Cube 16](#_Toc8597675)

[2.4 FDA Common Architecture Whitepaper (FDA-08) 16](#_Toc8597676)

[2.5 Inventory and Characterise Existing FDAs (FDA-09) 17](#_Toc8597677)

[2.6 USGS Architecture for Services in the Cloud 17](#_Toc8597678)

[2.7 ESA PDGS Data Cube 17](#_Toc8597679)

[2.8 The European Data Cube Service 18](#_Toc8597680)

[2.9 GEOhazards Thematic Exploitation Platform 18](#_Toc8597681)

[2.10 Copernicus Data and Exploitation Platform – Deutschland 19](#_Toc8597682)

[2.11 Food Security Thematic Exploitation Platform 19](#_Toc8597683)

[2.12 Plateforme d'Exploitation des Produits Sentinel (PEPS) 19](#_Toc8597684)

[2.13 JASMIN Super Data Cluster 20](#_Toc8597685)

[2.14 Big Data Approaches Supporting Market Growth in EO at the Satellite Applications Catapult 20](#_Toc8597686)

[2.15 Concrete Next Steps for FDA Interoperability and Ecosystem Take-up 20](#_Toc8597687)

[3 Data PRESERVATION and STEWARDSHIP 22](#_Toc8597688)

[3.1 ESA EO Data Preservation System 22](#_Toc8597690)

[3.2 Data Preservation Activities at NOAA 22](#_Toc8597691)

[3.3 EO Ontologies 22](#_Toc8597692)

[3.3.1 NASA 22](#_Toc8597693)

[3.3.2 ESA 22](#_Toc8597694)

[3.3.3 CNES 23](#_Toc8597695)

[3.3.4 Discussion 23](#_Toc8597696)

[3.4 RDA Possible Collaboration 23](#_Toc8597697)

[3.5 RDA FAIR Data Maturity Model Working Group 24](#_Toc8597698)

[3.6 WMO Expert Group Possible Collaboration 24](#_Toc8597699)

[3.7 GEOSS DMP IG Proposal 24](#_Toc8597700)

[3.8 PV2020 and Other Relevant Conferences 25](#_Toc8597701)

[4 Data DISCOVERY and ACCESS 26](#_Toc8597702)

[4.1 WGISS Connected Data Assets Status Report 26](#_Toc8597704)

[4.2 IDN Update 26](#_Toc8597705)

[4.3 CWIC Update 26](#_Toc8597706)

[4.4 FedEO Evolution 27](#_Toc8597707)

[4.5 NOAA NCEI 27](#_Toc8597708)

[4.6 IDN Front End Portal, Branding and User Metrics 28](#_Toc8597709)

[5 Data INTEROPERABILITY and USE 29](#_Toc8597710)

[5.1 Earth Observation Services Metadata Use Cases 29](#_Toc8597712)

[5.1.1 ISRO Use Cases 29](#_Toc8597713)

[5.1.2 NASA EOSDIS 29](#_Toc8597714)

[5.2 Inventory of Software and Tools (Open Source) (FDA-10) 30](#_Toc8597715)

[5.3 Carbon Data Portal Prototype (CARB-15) 30](#_Toc8597716)

[5.4 White Paper on Single Sign-On (SSO) Authentication (DATA-14) 31](#_Toc8597717)

[5.5 Recovery Observatory 31](#_Toc8597718)

[5.6 GEO Knowledge Hub and WGISS Contribution 31](#_Toc8597719)

[6 TECHNOLOGY EXPLORATION 33](#_Toc8597720)

[6.1 Data Services in the Cloud 33](#_Toc8597722)

[6.1.1 Advanced Geospatial Techniques: Aiding Earth Observation Applications 33](#_Toc8597723)

[6.1.2 OGC Testbed 14 and 15 33](#_Toc8597724)

[6.1.3 CNES Initiatives on Big Data and Cloud 34](#_Toc8597725)

[6.1.4 EOSDIS Approach to Data Services in the Cloud 34](#_Toc8597726)

[6.2 Artificial Intelligence and Machine Learning 34](#_Toc8597727)

[6.2.1 Workshop on Leveraging AI in the Exploitation of Satellite EO and Numerical Weather Prediction 34](#_Toc8597728)

[6.2.2 AI for EO and EO for AI 35](#_Toc8597729)

[6.2.3 Machine Intelligence towards Tomorrow’s Earth Science Data Systems 35](#_Toc8597730)

[6.2.4 A New Large-Scale Sentinel-2 Benchmark Archive to Drive Deep Learning Studies in Remote Sensing 35](#_Toc8597731)

[6.2.5 Multi-task Deep Learning from Sentinel-1 SAR: Ship Detection, Classification and Length Estimation 35](#_Toc8597732)

[6.2.6 CNES Initiatives on AI 36](#_Toc8597733)

[6.3 Discussion and Way Forward 36](#_Toc8597734)

[7 Agency and Liaison Reports 37](#_Toc8597735)

[7.1 Hungarian Space Office Liaison Report 37](#_Toc8597737)

[7.2 JAXA Agency Report 37](#_Toc8597738)

[7.3 NASA Agency Report 37](#_Toc8597739)

[7.4 ISO TC211 38](#_Toc8597740)

[8 WGISS Plenary, Part II 39](#_Toc8597741)

[8.1 WGISS Summary 39](#_Toc8597743)

[8.2 Review WGISS-46 Actions 39](#_Toc8597744)

[8.3 WGISS-47 Actions 39](#_Toc8597745)

[8.4 Concluding Remarks 41](#_Toc8597746)

[9 Glossary of Acronyms 42](#_Toc8597747)

Analytical Mechanics Associates, Inc. Andrew Cherry\*

CAS/AOE Xuesong Li\*, Ziyang Li\*

CAS/IGSNRR Yinghua Zhang\*

CAS/RADI Mingrui Huang

CEOS Executive Officer Steven Hosford\*

CNES Richard Moreno, Pierre LaSalle, Patrice Henry\*

CONAE Homero Lozza\*

CSA Paul Briand

CSIRO Robert Woodcock (WGISS Vice-chair)

DLR Katrin Molch\*, Christian Strobl\*, Christoph Reck\*

ESA Mirko Albani (WGISS Chair), Francesco Barchetta\*, Andrea Della Vecchia, Damiano Guerrucci, Daniele Iozzino\*, Marco Leonardi\*, Sveinung Loekken\*, Cristiano Lopes, Iolanda Maggio, Philippe Mougnaud\*, Marc Paganini\*, Antonio Romeo\*, Michelle Piepgrass (WGISS Secretary)

EUMETSAT Joerg Schulz\*, Uwe Voges\*

GEO-SEC Paola de Salvo\*

HSO Gábor Remetey-Fülöpp

Imt-atlantique Clement Dechesne\*

ISRO Nitant Dube\*

JAXA Makoto Natsuisaka, Yousuke Ikehata

NASA Andrew Mitchell, David Borges, Lauren Childs, Valerie Dixon, Yonsook Enloe, Brian Killough (CEOS-SEO), Christopher (Chris) Lynnes, Manil Makey\*, Michael Morahan, Douglas Newman, Kenton Ross, Nancy Searby\*, Kurt Thome, Archie Warnock\*, Jennifer Wei\*, Tian Yao\*

NOAA Ken Casey, Martin Yapur, Matthew Austin, Rich Baldwin\*, Sid Boukabara, Ed Kearns\*, Anne Kennerley, Robert Macadaeg, Ge Peng\*, Nancy Richey

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

UKSA Esther Conway\*, Robert Fletcher, Philip Kershaw\*, Daniel Wicks\*

University of Berlin Begum Demir\*

University of Wisconsin Steven Greb

USGS Kristi Kline\*, Renee Pieschke\*



# WGISS Plenary Session, Part I

## Introduction

Mirko Albani, WGISS-Chair (ESA), opened the WGISS-47 meeting, thanking everyone for their participation, and NOAA for hosting. He asked each participant to introduce him/herself, and reviewed the agenda. Mirko gave the highlights of the agenda and the agenda was adopted. He noted that this is a working meeting, and asked everyone to participate.

## Logistics

Martin Yapur welcomed the participants, and thanked his team for arranging the logistics for the meeting. He introduced Dr. Steven Volz, CEOS SIT Chair, NOAA Principal to CEOS, and Assistant Administrator, NOAA Satellite and Information Service (NESDIS).

## [Host Welcome](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_09.00_WGISS-47_Volz_Opening.pptx)

Steven Volz welcomed the participants, acknowledging those who travelled far in order to attend. He noted that WGISS is a stand-out member of CEOS in terms of achievements and contributions. He added that it is well recognised in CEOS that WGISS successfully promotes collaboration in the development of systems and services that manage and supply EO data. He spoke from two perspectives: as CEOS SIT Chair, and as NOAA principal in CEOS.

As CEOS principal for NOAA he discussed the information lifecycle from the NOAA perspective noting that it is key to understand the user requirements to apply information, and to make datasets interoperable and understandable. From the NESDIS perspective the objective is to ingest the data rapidly from multiple satellites for successful operations, to interconnect satellite and in-situ data for ecosystem modelling and forecasting, and to enhance the value of the data archives through interconnection with global archives.

The NOAA connection to WGISS is essential, since WGISS provides a forum for space agencies to interact in terms of their information systems and to provide harmonisation. He described the WGISS interoperable infrastructure, highlighting that less data will be downloaded as users can be more selective of the data they want and need. It is important to select the correct metrics to demonstrate the value of the agencies’ investment into the EO systems. Usability and quality play a central role. Understanding the CEOS/GEO relationship is important, and is a challenge from the information standpoint. The standards that are set are key for interactions with NOAA.

As SIT Chair, Steven noted the importance of working with VCs and WGs as they address the CEOS and GEO objectives in a sustainable manner. It is clear that WGISS has successfully provided a single entry point for data discovery. The common approach that WGISS takes allows solutions to work together, integrating a diverse set of users. The work of FDA shows that WGISS is looking forward.

In the two-year cycle as SIT Chair Steve is focused on enabling CEOS to function effectively, with a goal for the working groups and virtual constellations to work together, determining if they are sustainable and identifying any overlap that can be streamlined. A key question is if the VCs and WGs see WGISS as a resource, and should the VCs identify technical challenges where WGISS may offer help.

The VCs are thematically based, and were created to fill technical gaps. As systems have become healthier, integrated information sets are more available and a different construct of integrated products, like atmosphere/ocean, result from finding commonalities across the VCs. A proposal of restructuring of the VCs asks the question if there is value in an integration of the oceans constellations into a single one, and if so, how do the data systems integrate, and how can the data designers be influenced? Pushing the integration farther upstream lowers the burden of uptake for the users. Steven provided a matrix to visualise this concept.

The current proposal is the formation of a new CEOS working group focused on Information Provision to coordinate all activities related to user outreach and applications (e.g., forests, agriculture, freshwater, land degradation, urban, biodiversity, and more in the future). The expertise of this working group would be in finding common best practices for user interfaces, interpreting user needs, generating satellite-based observation requirements, and coordinating with VCs on tailoring observations to better meet user needs. As an example, Steven discussed the SST-VC/GHRSST idea, which combines satellite-derived and in-situ SST to provide pre-processed and merged products (with good estimates of uncertainty) to operational users and the science community.

Two study teams have been set up to understand where the issues and opportunities are:

* Ocean Virtual Constellation Merger Study Team (OVCMST)
* Ad Hoc Working Group Study Team (WGST)

Steven concluded with a reminder of the next SIT Technical Workshop that will include a VC and WG workshop.

Mirko thanked Dr. Volz for the comprehensive description and confirmed that WGISS is happy to participate with the study team activities.

## [National Centers for Environmental Information (NCEI) Overview](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_09.00_Kearns%20-%20CEOS-WIGSS%20Apr%202019.pptx)

Ed Kearns\*, NCEI Chief Data Officer, gave an overview of the National Centers for Environmental Information (NCEI). He discussed recent developments in USA law and mandates for NOAA; the agency is meeting the challenge of complying with the new strategies. One way is with the NOAA Big Data Project; NOAA is seeking a sustainable partnership and has developed a conceptual and functional framework for this.

In response to a question from Robert Woodcock, Ed explained that providing expertise to the users on bands, data formats, limitations of resolution and timing of data can be difficult; but once the partners identify a market for the data and the concepts are explained the users are off and running. The partnership has shown that there is an element of scalability, and the media tools can scale the expertise.

Mirko asked if they are relying on Google and Amazon. Ed replied that they are, but because they are not federal systems, there is a complementary access. For public access NOAA will keep control. Potentially the entire archive will be included. He added that they have insight into usage of the data, but metrics are not part of the contractual agreement so these are very ad-hoc.

Andrew noted that NASA jumped right into the contracts with the large cloud providers, but their data does not show up in the Open Amazon Portal or the Google Big Portal. The fine print on the contracts does matter; it is very important to get those terms (licensing, logo, no data restrictions) set up contractually and to define clearly the business relationship.

Mirko noted that WGISS is working on a white paper on Single Sign-On, and it would be interesting to see what NOAA is doing to complement the input from ESA and NASA. Ed said he would be very happy to do this, and to cooperate in any way.

## [Introduction to NCEI](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_10.10_NCEI_Overview_WGISS_Casey_Apr2019.pptx)

Ken Casey introduced NCEI which hosts and provides access to one of the most significant archives on Earth, with comprehensive oceanic, atmospheric, and geophysical data, and provides science, service, and stewardship to a wide range of stakeholders. He noted that the data holdings encompass weather and climate, coasts, oceans, and geophysics, and that the stewardship of this data plays a key role in their objectives.

NCEI operates on four tiers of data stewardship: national services and international leadership, authoritative records, derived products, scientific improvements, enhanced access and basic quality assurance, and long-term preservation and basic access.

Chris Lynnes asked if authoritative records have a legal standing; Ken replied that certain specific weather reports are certified.

Ken gave details on the NCEI functional organisation, with a focus on the Center for Weather and Climate, and the Center for Coasts, Oceans, and Geophysics. He also described their archive and access volumes and weather and climate products.

Ken noted that he will take Martin Yapur’s role in WGISS; Nancy Ritchey will participate in the DSIG, and Rich Baldwin will participate in the Discovery and Access Interest Group.

Makoto Ikehata asked if NCEI gathers surveys from users. Ken replied that NCEI has a user services group, a large customer relationship data base, attends science conferences, and has direct interactions with users. Their website also has a user survey which provides feedback. Newly developed tools and services result in extensive interactions with the users.

## [WISP Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_10.20_WISP_Report.pptx)

Michelle Piepgrass gave a report on the status of the WGISS Infrastructure Services Project. She noted that the purpose of the WISP is to make technology an asset for WGISS needs by offering technological support during WGISS meetings and providing solutions for the needs of other WGISS interest groups and projects. The primary activities are maintaining the WGISS webpages found on the CEOS website, maintaining WGISS email lists, providing teleconference support and meeting logistics coordination during the semi-annual in-person WGISS meetings (including teleconferencing for remote attendees) and support WGISS outreach activities.

Michelle outlined the existing team, noting that the team is currently without a lead or agency support.

After some discussion it was understood that the SEO will continue to provide teleconferencing licenses and website support. Other activities can be absorbed into the secretariat. Brian Killough noted that the SEO has a fund and issues some subcontracts to Symbios for certain types of support.

## [WGISS Chair Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_10.30_Chair%20Report%20V1.pptx)

Mirko Albani gave a report on key WGISS topics. He highlighted key activities at SIT-34, noting that WGISS will need to be involved in the SIT Study Team to ensure that WGISS is aligned with any reorganization. There were two side events related to WGISS.

The ARD side event was organised by WGISS and was well attended, with very good discussions which highlighted the need to reinforce collaboration with SEO (e.g. for FDA-10) and LSI-VC on FDA and ARD. There is much interest in the outcomes of the WGISS FDA workshop at WGISS-47 and expectations on tangible results deriving from it. The objective is to define and then implement an FDA-related quick-win project/demonstration to be shown at the next plenary. WGISS has also been asked to organise a session on STAC and one on ARD tools and Data Cubes at WGISS-48.

The GEO Knowledge Hub side event organised by WGISS was also well attended, with excellent discussions on the GEO KH concept and proposed way forward for implementation; Robert is a member of the Expert Advisory Group (EAG) defining the proposal for GEO KH. The GEO KH concept is continuously evolving to align with comments received from GEO EXCOM and GEO member delegations. The GEO Work Plan is also being updated with major changes that will impact WGISS participation (e.g. possible termination of the GEOSS-Evolve initiative). WGISS Exec will need to define contribution to the new GEO Work Plan.

WGISS should provide a representative in the GEOSS Infrastructure Implementation Task Force which will replace the EAG. The CEOS SIT Chair asked WGISS to better understand impacts/implications for CEOS deriving from the GEO KH concept and to develop a response to be used at the CEOS level.

The following SIT-34 actions relate to WGISS:

All CEOS Virtual Constellation Co-Leads, Working Group Chairs and Vice Chairs, and Ad Hoc Team Co-Leads to provide updated content and information for the CEOS website to CEOS SEO, with a cc to the CEO. Due Date: 30 June 2019

WGISS would like a representative to be named on the “GEOSS Infrastructure Development Foundational Task /Implementation Coordination/Task Force”

All WGs to provide feedback on new Deliverables Dashboard and to update CEOS deliverables tool. Due Date: 31 August 2019

The next CEOS Plenary is 14-16 October in Hanoi.

Mirko reported on the following areas of cooperation for WGISS:

CEOS Carbon Team - ongoing development of prototype Carbon Portal

WGClimate - ECV/CDRs registration in IDN in parallel with ECV inventory v3 update agreed during WGClimate-10 meeting

WGDisasters - possible cooperation on generic recovery observatory

WGCV - ongoing with two data cube prototypes (ESA and GA/CSIRO) supporting calibration/validation activities through providing access to ACIX (16 sites), RadCalNet (4 sites) and LPV (3 sites) calibration/ validation data.

Virtual Constellations - inventory of VC data to be updated to ensure discoverability/accessibility through WGISS infrastructure

WGCapD - joint organisation of technology webinars and FDA events

SEO - COVE tool harvesting WGISS metadata, data cubes coordination

SDG AHT - SDGs platforms and data interoperability pilot under discussion

GSICS - WGISS presented at the Global Space Inter-Calibration System annual meeting; cooperation areas with Data and Research WG were identified

GEO

GEOSS Platform accessing CEOS agencies data via WGISS CDA

WGISS membership in GEOSS-EVOLVE initiative and GEO Expert Advisory Group (EAG)

WGISS participation and support to 4th GEO Technology Workshop

WGISS represented in NextGEOSS Advisory Board and working with NextGEOSS on federated authentication technology solutions

Joint Workshops on GEOSS-WGISS interoperability and FDA held at WGISIS-43 and WGISS-46.

GEO Knowledge Hub side event organised at SIT-34

GEO Sec invited to attend WGISS-47

CEOS representation in the GEOSS Implementation Task Force

Regarding WGISS internal matters, Mirko reported that the WGISS organisational structure chart is being updated. WGISS has the following sustainable commitments:

Interagency coordination working well, very good cooperation spirit

Healthy participation to meetings (average 20-25)

Some members attending all meetings (e.g. NASA, NOAA, USGS, CNES, ESA, CSIRO, JAXA, RADI, AOE CAS, RSCC, HSO)

Others attending only some meetings depending on location and discussed topics (e.g. EUMETSAT, ROSCOSMOS, INPE, CONAE, DLR, UKSA, ISRO)

Additional agencies expressed interest in WGISS activities with potential future participation (e.g. GISTDA, VNSC/VAST, ASI, CSA, CCMEO)

The WGISS vice-chair October 2019 – October 2021 nomination period is open, and the WISP team is currently without agency leadership.

**Action WGISS-47-01**: Michelle Piepgrass to finalise and post the WGISS organizational structure chart. Due by May 20, 2019.

**Action WGISS-47-04**: Mirko Albani to request SIT Chair Team to include a WGISS representative to the CEOS Study Team on WGs. Due by May 20, 2019.

**Action WGISS-47-05**: Mirko Albani and Robert Woodcock to provide SIT Chair Team with the name of WGISS representative for the GEO Implementation Task Force to be proposed to GEO Secretariat. Due by May 20, 2019.

## [Review of CEOS 2019-21 Work Plan and WGISS 2018-2020 Work Plan](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_12.00_CEOS%20and%20WGISS%20Work%20Plans.pptx)

Mirko Albani gave a review of the CEOS 2019-21 and WGISS 2018-20 work plans. The CEOS 2019-21 contains near-term objectives and deliverables designed to achieve the goals outlined in the CEOS Strategic Guidance document. The main outcomes that involve WGISS are

Carbon Observations, Including Forested Regions,

Capacity Building and Data Democracy,

Data Discovery, Access, Preservation, Usability and Exploitation: approaches, systems, tools and technologies

CEOS Services

The WGISS actions in the CEOS Work Plan are:

CARB-15: Carbon data portal prototype.

FDA-5: Promote awareness of FDAs.

DATA-9: ECVs/CDRs discovery and access through WGISS systems.

DATA-11: Data and technology exploration webinars and workshops.

DATA-13: Develop a white paper on single sign-on (SSO) authentication.

DATA-15: Explore emerging trends and disrupting technologies, evaluate advantages/drawbacks for adoption in EO and identify most relevant use cases. Summarise analysis in the form of white papers.

DATA-16: CEOS data holdings reported and accessible in GEO and other international relevant contexts.

FDA-8: Establish a common description of Future Data Architecture functional locks and identify interfaces and interoperability approaches.

FDA-9: Inventory and characterise existing FDAs operated by both public and private entities including the standards and approaches they use.

FDA-10: Finalise inventory of software and tools available or used at CEOS 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.

FDA-14: Facilitate discovery and access for end users to data analytics and processing tools and services through the WGISS CDA infrastructure.

DATA-2: Full representation and accessibility of CEOS agency datasets through WGISS standards and CDA infrastructure.

The WGISS Work Plan will be updated for alignment with 2019-21 CEOS Work Plan. It contains in addition several internal actions and deliverables necessary to achieve the WGISS mandate (e.g. promotion, cooperation, sharing, etc.) It will be circulated for review/approval by end of May.

**Action WGISS-47-02**: Iolanda Maggio to update the WGISS Work Plan based on the CEOS Work Plan 2019-2021. Due by June 15, 2019.

## [CEOS Executive Officer (CEO) Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_09.00_CEO%20Context.pptx)

Steven Hosford\* gave a report on general topics related to CEOS. He began with a reminder of the VNSC priorities

Carbon Observations (forested regions)

Observations for Agriculture (rice)

The next CEOS Chair will be ISRO, and the next SIT Chair will be CSIRO/GA; the CEO position is vacant.

Steven reported that the CEOS 2019-2021 Work Plan received virtual endorsement on 22nd March and is available online. Thematic changes are the replacement of FDA with Data Quality and CEOS Services; WGISS now has its own section. He noted that the development of a new CEOS Deliverables database with added functionality is underway. For now, updates should be sent to Steven.

Mirko noted that WGISS is starting to address the actions in the new CEOS Work Plan.

## [SEO Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_12.10_SEO_Report_WGISS47.pptx)

Brian Killough gave the report of the Systems Engineering Office (SEO) activities. He began with a summary of the COVE Tool and discussed status, noting that in 2019 they plan to complete enhancements to the coverage analyser and data browsers tools, improve and increase links to mission archives, add a custom mission capability to support future mission concept studies, improve coverage and revisit calculation products, and add historical cloud coverage reports. Brian reported that links to archive databases are a constant issue, as changes occur on the data provider side. The primary focus is to get links to Landsat and Sentinel-1/2 working well, as they are the most desired.

The SEO team is working with Iolanda Maggio to develop an Open Source Software (OSS) Inventory. The SEO received a list of 150 OSS items to add to the new database. This data includes name, description, PoC, mission, instrument, company, release date, operating system and web location.

Discussion on Sentinel data from 2015-2017 that has not been found gave Brian several places to look (DIAS, JASMIN, and collaborative platforms)

Steven Greb noted that in the water community there are special needs; the water community would have to start with the product, but Amazon may not be willing to host it on top of what they are already holding. A better solution might be to put a tool that sits on top of the archive and generates the water product on the fly. Robert added that these are business model problems, not technical problems, and can be solved with a degree of cooperation, and with better understanding of the business models.

## CEOS Working Groups Cooperation with WGISS

### [WGCV](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_11.30_WGISS47_wgcv_wgiss_interaction.pptx)

Kurtis Thome gave an overview of the CEOS Working Group on Calibration and Validation (WGCV). He discussed current activities in the areas of biomass validation protocol, surface albedo validation protocol, greenhouse gas reference standards, and ARD. In April 2018 four topics were identified for cooperation between WGISS and WGCV:

Data Formats and Interoperability in the framework of FDA: Robert Woodcock (WGISS), Medhavy Thankappan (WGCV)

Quality Indicators in Discovery Metadata: Michael Morahan (WGISS), Nigel Fox (WGCV)

CEOS Data Cubes and CEOS Test Sites Data Access in support of WGCV Activities: Robert Woodcock (WGISS), Greg Stensaas (WGCV)

Standardisation and Best Practices (e.g. ISO 19159-3): Iolanda Maggio (WGISS), Cindy Ong (WGCV)

Initial telecons were successful at organising the activities after the joint meeting in Brazil in April 2018. WGCV and WGISS meetings in October 2018 helped to spur further action as did the CEOS SIT meeting in September 2018. The current WGISS meeting and the July WGCV meeting are helping to remind activity leads of next steps. Key next steps are:

To have a WGCV-organised telecon of the activity leads on both working groups prior to July WGCV meeting

Assess the workloads of the activity leads and determine if additional help is needed since workloads can change in 12 months

Prepare for the next WGCV/WGISS joint meeting

Chris asked if they have documents that help with the calibration/validation process; the LPV has done a fantastic job at producing these. Mirko asked if they could add their documents to the BP and Guidelines web page. Paul Briand asked for these as well, Kurt said they are working very hard to try to get these out.

**Action WGISS-47-03**: Michelle Piepgrass to contact Cindy Ong to obtain the BP and Guidelines for WGCV. Due by May 31, 2019.

**Action WGISS-47-15**: Andrea Della Vecchia and Robert Woodcock to define with WGCV colleagues (Medhavy Thankappan, Phillipe Goryl, and Cindy Ong) coordinated roadmap and short term activities on WGISS-WGCV Data Cubes (ESA and ODC DEA) and possibly demo to be shown at SIT TW in September 2019. Due by June 30, 2019.

### [WGCapD Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_09.00_Searby_WGCapD_WGISS_47_v1.pptx)

Nancy Searby\* presented a report on WGCapD activities. She began with a reminder of their purpose to raise awareness of the value of EO and build capacity in user communities, by providing support to CEOS initiatives and help to WGs and VCs to undertake their own capacity building initiatives, and collaborate with GEO, UNOOSA, and other UN agencies.

Nancy presented the WG’s work plan, showing linkages to the CEOS Work Plan. She also detailed regional activities (AmeriGEO, AfriGEO), such as national in-person training, regional activities and global activities, and developing infrastructure for online trainings and schedules.

With growing interest in Future Data Architectures, WGCapD will identify ways of promoting the use of Future Data Architectures and organise outreach capacity building activities for end users and decision makers (e.g. webinars, workshops, etc.)

Nancy was asked if WGCapD communicates with the academic community; Nancy said that they work with graduate students, trainings, and conferences, and work on relationships for degree program development.

Mirko reminded that for FDA-5 WGISS and WGCapD are tasked to organise webinars together. One possibility would be short demonstrations of the FDA elements presented in the FDA workshop. Kenton agreed, and Chris noted that some webinars have already occurred for these topics. Enough time has passed that it would make sense to present them again.

Nancy reminded that the target audience and purpose should be identified first; a webinar for a data scientist would be different than for an atmospheric scientist; one for the academic sector would be different from one for developing country working professionals. WGISS has more experience with the academic so they should be targeted first, and then the webinars can be adapted for the working professionals. Kenton added that the audiences are often academics transitioning into professional, and can give us feedback to tailor these.

Chris suggested that the first webinar should be something somewhat operational.

Gabor added that HSO is a stakeholder in this, and has had interesting presentations – the stakeholders in Hungary have been active for two years and have expressed interest in staying in contact. Gabor will serve as liaison.

**Action WGISS-47-18**: Chris Lynnes, Yousuke Ikehata and Kent Ross to organise a webinar in September focusing on existing (pre)operational Exploitation and Application Platforms targeting the academic community. Following webinars should target the new developing countries on the same topics. Due by September 30, 2019.

### [WGClimate](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_09.00_WGClimate_Schulz.pptx)

Joerg Schultz\* began his presentation with a short history and objectives of the WGClimate. He explained the architecture for climate monitoring from space and the relationship between WGClimate, CGMS, and GCOS.

Joerg discussed the ECV Inventory, a resource for coordinated response. The ECV Inventory fully describes current and planned implementation arrangements (ECV Product-by-ECV Product) within the architecture; it is verified information about almost 1000 (ECV#2) data records including direct access points to the data. The inventory is continuously updated with gap analysis and action plans created annually with approval from CEOS and CGMS. The Recommendations and Coordinated Actions inform space agency planning, improve availability and interoperability of climate data and the inventory feeds material for all future responses to the GCOS IP.

Users can download the ECV Inventory content ([http://climatemonitoring.info](http://climatemonitoring.info/)) for their own analysis, find direct access points to all CDRs in the inventory, get access to WGClimate gap analysis results and planned actions, and access case studies analysing the use of CDRs for applications. The team is currently updating version 3 of the inventory.

Current work with WGISS relates to CEOS DATA-9: Collaborate on facilitating discoverability and accessibility of ECV Products and space-born CDRs relevant for the CEOS Carbon Action via WGISS Interoperability Systems & Standards (FedEO/CWIC/IDN, OpenSearch):

ECV Inventory #3 contains a question on data record registration with IDN; WGClimate will monitor answers in its gap analysis process.

Recommend that agencies should register CDRs in IDN; it is noted that some agencies providing input to ECV Inventory are not part of CWIC or FedEO federations.

Agreed at WGClimate-10 in March 2019 that WGISS will take ECV Inventory #2 and analyse what essential information is missing (ECV/CDR collection entries not in alignment with IDN guideline on information content).

This should be followed by provision of this information to the WGClimate identifying the “delta” needed to fulfil the needs to complete DATA-9. Based on this description a way forward can be developed.

Joerg suggested that new collaborative activity with WGISS could address potential automation of ECV Inventory population from existing agency common metadata repositories:

Need to map which ECV questions can be covered by metadata repositories (ECV Inventory is not only a metadata base but provides a logical representation of information on CDRs traceable to GCOS guidelines, etc. to facilitate test against it).

Need to analyse if and how ECV Inventory questionnaire and/or metadata repositories would need to be adapted to facilitate automation.

May start with a demonstrator from one agency having large shares in the Inventory, e.g., NASA.

If a solution is found need to see how to test and implement as ECV Inventory update is continuous.

Should consider CEOS Work Plan action(s) in next update at end of 2019.

Mirko thanked Joerg for the comprehensive information. WGISS can analyse what kind of effort the suggested work with and provide feedback. WGISS analysed inventory 2, and most of the information to register items in the IDN is there. It has been suggested to make some of the optional elements to mandatory, thus enabling automatic registration in the IDN.

Joerg also confirmed that ECV #3 contains confirmation of contents of ECV #2, and about 400 new data records from Asia or from new programs.

Ge Peng asked about the 63 records that were removed in ECV #3; Joerg replied that they were removed because the agencies do not continue the provision of the data. This could be a concern if there are not enough records in a particular element. Mirko noted that WGISS supports a purge alert service aimed to avoid losing data.

Mirko reminded Joerg that the Carbon Portal could be interesting to present this at a WGClimate meeting. Joerg replied that as they add carbon activities to their group this will be very interesting and he will add it to the agenda of their next meeting.

Esther asked if CDRs come into this area, and Joerg replied that they have not been entered yet. Esther offered to provide them and Joerg agreed to put her in contact with Alexandra for this. Esther added that a lot of work on the format has been done; if there is a format that can be identified as best it will be good. Esther mentioned Nigel Fox as a contact.

**Action WGISS-47-29**: CDA SLT to analyse what kind of effort and solution can be accomplished by WGISS to automate future ECV Inventory updates [last slide of presentation].

**Action WGISS-47-30**: CDA SLT to send to WGClimate the list of elements in the inventory that should be made mandatory so they can be registered automatically. By June 30, 2019.

**Action WGISS-47-33**: WGISS to give a demonstration of the Carbon Portal at the WGClimate meeting in September.

**Action WGISS-47-32**: Mirko Albani to send information to Joerg Schultz about the Purge Alert procedure. Due by June 30, 2019.

### [Sustainable Development Goals Ad-hoc Team (SDG AHT)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_09.00_SDG-AHT_Paganini.pptx)

Marc Paganini\* gave a presentation on the Sustainable Development Goals Ad-hoc Team (SDG AHT). He discussed a few of the roadblocks to sustained data for sustainable development. He listed the criteria used to streamline CEOSS engagement on SDGs, highlighting the need to leverage the knowledge and expertise of CEOS bodies for maximum impact and for EO enabling infrastructures.

As a demonstration of the effectiveness of the streamlining measures, the SDG AHT proposes to start with the three SDG indicators that are most ready to integrate EO in their processes. These three indicators are also the primary indicators selected by the IAEG-SDGs WGGI for the Task Stream on satellite Earth Observation data for the SDG indicators. Marc showed examples of this for several of the sustainable goals.

Mirko asked how WGISS can participate, and Marc suggested looking at what is available in platforms making them easily understandable by the SDG holders. Secondly, they need help for easily discovering data.

**Action WGISS-47-34**: Mirko Albani to liaise with Marc Paganini on discovery and access of data and platforms identified in the SDG slides. Due by July 31, 2019.

## [GEO Secretariat Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_10.45_KH_v5-GEOdatatech_GR_HB_PDS_gcamara_.pptx)

Paola de Salvo gave the GEO Secretariat report. She had just returned from the GEO Plenary, where CEOS participation was very important. She thanked everyone for their valuable contribution. Paola discussed the GEO Knowledge Hub (GEO KH). She noted that the GEO Community produces a wealth of incredible scientific knowledge; the knowledge about a specific topic comes when a series of elements are connected together to be able to produce and reproduce outputs and services. This authoritative knowledge is not yet captured in an effective manner, but reside still scattered in various scientific publications. Analysing closely a scientific publication is easy to see many essential elements such input data, models, tools, software and output data mentioned. For technical or non-technical end users, having those elements easily findable and all connected together could bring a great advantage.

The GEO KH will be a single place to discover, access, and reuse the knowledge and results developed by the GEO Work Programme; it will be done in cooperation with the community for knowledge-building via ingestion, curation, and GEOSS Platform connection. The resource types are publications, code, models, and tools, remote sensing data, in-situ data, results and products, and videos and others. When all the components are assembled there will result a collection of curated, linked, and indexed documents. The important message is that the GEO KH is linked together with the GEO Work Programme, GEOSS Evolve, and GEOSS Platform, in coordination with the GEO SEC.

Mirko noted that WGISS is very interested in the GEO KH, and in understanding the relationship with the GEOSS platform, what will happen with the DMP, and what the KH way forward is.

Paola noted that there has been a request by the program board to have a map of the FAIR principles with the DMP and suggested that WGISS participate at the GEO Work Plan Symposium. Iolanda is involved in the FAIR, and WGISS is in line to provide good support.

Richard commented that it may be worthwhile for GEO to contribute with the FDA.

**Action WGISS-47-07:** WGISS-Exec to follow-up GEO Work Plan preparation and identify where WGISS can contribute with: (Due by June 30, 2019)

* Data Management Principles (DSIG)
* WGISS Connected Data Access
* Future Data Architectures

## [Future Meetings](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_11.20_Future%20Meetings.pptx)

Robert Woodcock announced that the WGISS-48 meeting would be held in Hanoi, Vietnam, from October 8 to 11 and will be hosted by VSNC. The meeting will be held at the VNSC Building (A6), VAST Campus. Transportation and accommodation details were given, along with a tentative agenda. Information can be found at [http://CEOS.org/meetings/wgiss-48/](http://ceos.org/meetings/wgiss-48/).

WGISS-49 is tentatively set for April 2020 in Argentina, hosted by CONAE, and WGISS-50 in October 2020 in Moscow, jointly with WGCV.

# Future Data Architectures Elements Demonstration and Interoperability Workshop



## [Introduction and Scope](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_13.50_FDA%20Intro%20and%20Scope.pptx)

Robert Woodcock introduced the workshop on Future Data Architectures Elements Demonstration and Interoperability. He described this as an interactive session, noting that there is a lot of interest in this at the SIT level. The FDA themes are Analysis Ready Data (ARD), Data Cubes, EO Platforms, User Metrics / Resources Inventory, and EO Data Analytics The approach to these is practical implementations and pilot activities, improvements in their conceptual framework, and standardisation initiatives (mainly in WGs and VCs). Rob listed several examples of elements. This is a good time as agencies are experimenting with FDAs; the goal is to see demonstrations, to discuss WGISS implications, and to develop a concrete WGISS project.

## [GEO AquaWatch](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_14.10_GEOAquawatch.pdf)

Steven Greb presented on the GEO AquaWatch Initiative, which covers inland and coastal waters, with the aim to develop and build the global capacity and utility of Earth Observation-derived water quality data, products and information to support water resources management and decision making. He described the AquaWatch organisational model and discussed examples of current data and product sources that are disparate, and with different degrees of validation. The work ahead is to continue expansion of the AquaWatch Knowledge Hub.

Robert asked Steven to elaborate on the discoverability needs. Steven replied that these are diverse depending on the user, and that some portals have products but they are not easy to discover and access; thumbnails of the images would decrease the effort. The community only needs a small number of collections, but at the file level it is time-consuming to identify what is usable. Ideally, it would be helpful if the IDN could be used to describe the files that can be used by the community.

**Action WGISS-47-16:** Robert Woodcock to liaise with AquaWatch representatives to gather their requirements/needs and further define how WGISS could support. Due by June 30, 2019.

## [CEOS Client Interfaces and Open Data Cube](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_1430_OpenDataCube_WGISS47.pptx)

Brian Killough gave a presentation and demonstration of the CEOS Open Data Cube (ODC) and FDA interfaces. He gave details on the ODC, and described its applicability in the Africa Regional Data Cube. He noted that the ODC (based on Open Source Software) promotes the use of CEOS Analysis Ready Data for Land (CARD4L) and supports many implementations that use cloud-based architectures (e.g. Amazon AWS). The ODC shares algorithms through the use of Jupyter Notebooks. Brian described some of the uses of the ODC, and gave examples in the areas of water extent, water quality, coastal change, urbanisation, and land change. He demonstrated the web-based user interface, where, in a matter of a minutes, he obtained a cloud filtered one year mosaic, with ability to download all seven bands for a single scene.

This activity is funded through the SEO.

**Action WGISS-47-13**: Andrea Della Vecchia and Robert Woodcock to define with SEO a roadmap and short term activities on interoperability between ODC / HMDC / DCFS Data Cubes (e.g. Jupyter notebooks with XArray API and WCS / WPS). Due by June 30, 2019.

## [FDA Common Architecture Whitepaper (FDA-08)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_14.50_FDA%20Common%20Architecture%20Whitepaper.pptx)

Robert Woodcock described the activity toward writing the FDA Common Architecture whitepaper of a description of FDA functional blocks and identify interfaces and interoperability approaches (FDA-08). The current emphasis is to promote richer relationships, describe on-demand, user-centric discoverability, addressing the paradox of that more data can be less usable, and that describing the uses of the data. Ultimately, the CEOS role is in the sharing platform.

Robert posed the following questions: Is a whitepaper needed? What must it cover? Who will participate?

Esther said this something she would be quite keen to participate in; she needs to be able to present this to a wide range of users, and the WGISS contribution would be valuable.

**Action WGISS-47-08**: Robert Woodcock to define next steps and schedule for FDA-08 white paper. Due by May 20, 2019.

## [Inventory and Characterise Existing FDAs (FDA-09)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_15.10%20FDA-09-%20Inventory%20and%20characterize%20existing%20FDAs.pptx)

Mirko Albani gave the background of FDA-09: Inventory and characterise existing FDAs operated by both public and private entities including the standards and approaches they use.

Mirko noted having identified 44 FDA elements, and characterised them into four layers: data generation, resource, exploitation, and platform layer, and four services: discovery, processing, download, and view. Next steps are to perform the final round of inputs/review/update by end of July, and formal closure of the CEOS Work Plan action. The generation of metadata for mature services and FDA elements and discovery through WGISS CDA Infrastructure and/or “OSS like” inventory database will follow, along with a pilot interoperability project with selected FDA elements.

Ken asked for the source of the list of elements; Mirko replied that fundamentally through survey of WGISS agencies; it is based on what was collected, but further rounds of review will provide additions/changes.

**Action WGISS-47-09**: Mirko Albani and Iolanda Maggio to complete FDA-09 Inventory deliverable. Due by July 31, 2019.

## [USGS Architecture for Services in the Cloud](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_15.40_USGSCloudArchitecture.pptx)

Kristi Kline gave a presentation on USGS Architecture for Services in the Cloud. She began with the rationale and benefits of moving from web-enabled to cloud-enabled Landsat data. The new paradigm enables opportunity for users to access the data directly allowing execution of algorithms directly on only the data they need, selective data usage (specify bands etc. for use), and reduced need for IT infrastructure.

Kristi described the operations concept, and listed key AWS services; she also described the cross-region replication option. Kristi noted that USGS is considering pixel-level metadata using a raster file of 16 bits per pixel. The intention is to process the data from level 0 to 1 to 2. The key thing they have not analysed is cloud-optimised indexes for GEOTiff files; things are moving so fast that standards are not emerging

WGISS would be helpful in developing Best Practices for loading data in the cloud and making it most interoperable.

Robert asked for the timeline; Kristi said that they hope to be operational by the end of the year.

## [ESA PDGS Data Cube](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/1.%20Monday%20April%2029/2019.04.29_16.00_ESA_PDGS_Data_Cube.pptx)

Andrea Della Vecchia gave a presentation on the ESA PDGS Data Cube. He listed the objectives, and described two use cases: UC1-EO Data Discovery, Visualisation and Pixel-based Access, and UC2-EO Service Discovery and Processing. He described ESA’s data storage, and two use cases for EO data management.

For UC1, the user connects to the GUI (<https://eodatacube.eu/>), selects the collection(s), and defines AOI/TOI. The GUI connects to ESA’s data cube and sends the access/processing query to the DAS component. The DAS retrieves the data from the online storage and sends back the GUI.

For UC2, the user connects to the same GUI, selects the collection(s), define AOI/TOI, connects to the central Jupyter hub, and proceeds to use/edit/create notebooks to manage the full data cycle (discover, access, process, visualise).

For in-situ data management, Andrea described two use cases:

UC2.1: In-situ/EO Data Discovery, Visualisation and Pixel-based Access for Atmospheric Correction Inter-Comparison Exercise (ACIX) and for Radiometric Calibration Network (RadCalNet).

UC3.1: In-situ/EO Service Discovery and Processing (on-going) for Atmospheric Correction Inter-Comparison Exercise (ACIX), for Customised UI to support AOT/WV/SR inter-comparison, and for Jupyter notebook to support live coding/sharing.

Chris if there is a possibility of extending to WCVS; Andrea replied that feedback has not been positive because of flexibility. However, with WCS to access data, you could transfer to WCVS or Jupyter. Further discussion is recommended.

ESA PDGS Data Cube software refactoring is currently on-going. Andrea described the adherence to standards/evolution of standards for discovery, access, and processing.

The EO collaborative environment is a virtual working environment providing discovery and access services to EO data, analysis tools, processors, information and communication technology resources required to work with them through one coherent interface. Andrea described the collaborative environment, which aims to change the paradigm from data-to-user to user-to-data. Andrea detailed the next steps for 2019; a demonstration followed.

Robert asked if any of this work should be elevated to a common CEOS approach. Andrea replied that this is a very small activity but ESA has tasks for supporting the activities, and the opportunity exists for discussions and contributions. Robert observed that Brian put forward a method where any component in the OGC can be replaced with any other component; does the same flexibility exist? Andrea replied that it is a matter of implementation; the scalability depends on the implementation. The engine can be adapted to the API.

Cristiano added that this solution has flexibility in that the data access layer is very transparent. Flexibility also exists for the data owner. ARD requires the presence of the in-situ data.

## [The European Data Cube Service](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_09.20_European%20Datacube%20Service%20for%20EO.pptx)

Cristiano Lopes gave a presentation on the European Data Cube Facility Service (DCFS), a service of the EO Platform Ecosystem. DCSF implements a platform service for higher-level data analytics. The DCFS will enable Information Layer Publishing, On-Demand Mapping, and Cross-Mission Analysis. Users will be able to manipulate/work with basic instrument data and level 3/4 data; DCSF will support thematic use cases. Cristiano showed a movie to visualise what a powerful tool the Data Cube is for extensive analysis.

Cristiano described the cost of the DCFS offer of data and services, and described the deployment timeline. He also displayed a diagram of the overall architecture, the interoperability layer and the OGC and the open tools contributions of this project.

Benefits for the users:

On-the-fly processing for generic data access.

Full-fledged data cube solution based on xArray.

Minimising pre-processing makes a flexible system.

Based on already operational components (SH) dealing with millions of daily requests.

Access to global archives.

Interoperability with existing Data Cubes taken to foster federation.

Operational service evolving over project duration from user needs.

Chris asked if the implementation would be Open Source; Cristiano confirmed that the interface part will be.

Robert requested details on the XCube; Cristiano agreed to send them.

Mirko requested an update at WGISS-48, specifically on the federation and interoperability that this addresses, which is what WGISS needs to achieve.

Ken Casey asked about the requirements of the underlying data. Cristiano replied that original mission format is the only requirement.

Mirko asked how WGISS can support this work. Cristiano replied that promotion by CEOS and WGISS and having other agencies on board would be very helpful.

**Action WGISS-47-10:** ESA to present progress of DCFS Data Cube at next WGISS. Due by WGISS-48.

## [GEOhazards Thematic Exploitation Platform](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_09.40_Geohazards.pptx)

Francesco Barchetta\* gave a presentation on the GEOhazards Thematic Exploitation Platform (GEP). He began with an overview of TEPs, discussing examples of coastal, urban, polar, forestry, and hydrology TEPs.

Francesco noted that the GEP provides on-demand processing services for specific user needs, running a systematic processing service to address “common information” community needs. The GEP connects to massive compute power on multi-tenant Cloud Computing resources to address the challenges of monitoring tectonic areas, globally to full Copernicus Sentinels-1/2/3 repositories, and specific data collections from EO missions, such as ALOS-2, Cosmo-Skymed and TerraSAR-X (provided under special arrangements in the framework of the CEOS WGDisasters and GSNL).

The GEP v2 provides new features, such as 20+ on-demand services using both optical and SAR data grouped in thematic applications according to the user profile, including new advanced services available (SAR based SBAS multi-baseline processing chain) and DSM OPT and MPIC OPT chains (CNRS EOST, Optical measurement using Sentinel-2 and VHR data). New basic services providing full resolution GeoTIFF imagery (RASTER on-demand service) and change detection imagery (SNAP based COIN and SNAC tools) for rapid on-line visualisation. There are eight systematic services including the large scale production of Sentinel-1 InSAR browse images at both 100m and 50m resolution over tectonic regions.

GEP v2 is bringing enhanced platform capabilities such as improved processing time and system scalability allowing parallel massive data processing, multi-sourcing, and integration with new Cloud based processing environments. It will also be able to visualise many large images and products on the geo-browser with a low latency and display dynamic histograms of a query result over selected time span and area of interest for any data collection and directly publish the results of a processing chain for community sharing.

The GEP Early Adopter programme is able to on-board additional users at [geohazards-tep@esa.int](mailto:geohazards-tep@esa.int). Francesco demonstrated the system.

Richard asked if it has been made ready on the DIAS. Francesco replied that it is on other DIAS like CREODIAS. What was shown was on internal cloud but some of the services are on different DIAS.

Valerie asked what the services are. It is open to any new development; one can visit platform to see the particular services.

## [Copernicus Data and Exploitation Platform – Deutschland](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_10.00_CODE_DE_Reck%20and%20Strobl_WGISS-47.pptx)

Christoph Reck\* presented the German Copernicus Data and Exploitation Platform (CODE-DE). He began with the background, and described the marketplace featuring discovery of datasets, data services, application projects, processors, and tools with WCS (INSPIRE) and DublinCore metadata display. It is HTTP-based, and includes value-added products.

CODE-DE uses a private Cloud with web user interface, command-line interface, and OGC WPS. Processor selection is from system-provided processors and user-provided processors. Features include fair queue scheduling for processing cluster. CODE-DE runs binaries and docker containers and has direct and high performance access to online data archive. Christoph discussed the architecture diagram, and gave a few usage statistics.

The Current service will exist until the end 2019; improvements of Catalog Client and Integration of Sentinel-5P, and value-adding processors are expected. The CODE-DE II Operational Platform is extended until 2023 and beyond, sharing product online archive with DIAS, having flexible provisioning, elastic cloud processing, and automated deployment. Christoph concluded with a demonstration of CODE-DE.

Mirko inquired about how many users and future plans in term of sustainability. Christoph replied that they have about 1500 users, with 300 active users each month, and the current system is contracted until September; the CODE-DE II will be supported until 2023.

## [Food Security Thematic Exploitation Platform](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_10.20_Food%20Security%20Thematic%20Exploitation%20Platform%20with%20Demonstration.pptx)

Antonio Romeo\* gave a presentation on the Food Security TEP. The TEP provides access to key satellite products and ancillary data, backed up by a scalable processing infrastructure and interacting with a range of users through a dedicated forum. It has the ability to easily develop new services, and to share processors and outputs only with selected user groups. A new business model offer for private companies also exists, with access to tools to derive agricultural and aqua-cultural products. There is also technical support for platform use. Provision on request is high-accuracy, quality checked vegetation parameters, suitable for use in operational scenarios. Access to ready-to-use products or customised services is available. Antonio listed the target user communities, the project team, and the project schedule.

Antonio described the platform portfolio, listed analytical tools, tools for computing, provision of biophysical parameters, and a list of supplemental datasets. He concluded with a demonstration of the platform.

## [Plateforme d'Exploitation des Produits Sentinel (PEPS)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_11.00_PEPS.pptx)

Richard Moreno, gave a presentation on the French Access to Copernicus Data (PEPS). PEPS has three main functions: "Mirror site" of ESA hubs providing access to Sentinel 1, 2 and 3 products; processing workshop allowing "first level" treatments, results visualisation and products downloading; and incubation of relatively "mature" processing chains for "scale" validation. It includes standard tools such as streaming visualisation, NDVI, colour compositions.

After agreement, PEPS hosts companies who want to develop and improve a spatial service. The hosting infrastructure is based on WPS, DOCKER and PROACTIVE meta-scheduler. The data are stored close to the HPC/HPDA. PEPS is linked with Thematic Data Centres. Richard also discussed the data storage.

PEPS has had a stable increase of registered users, and an increase of distributed data volume. Richard noted that Sentinel-1 is the most used data, and there is strong interest in archive data. Richard displayed multiple statistics of PEPS usage.

This year PEPS is running smoothly, experiencing regular upgrades for optimisation of data download functions and for user data access interface improvements. Richard noted that there is very low interest for Sentinel-3 data, but some users ask for S2\_L1B products. Strong and efficient ESA support exists for both for information requests and for backlog retrieving in case of maintenance/crash or data gaps filling.

Mirko commented that it might be interesting to discuss putting an ODC on the PEPS.

## [JASMIN Super Data Cluster](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_11.20_JASMIN.pptx)

Philip Kershaw\* discussed the JASMIN super data cluster, discussing its challenges and growth. JASMIN was developed to bring together diverse datasets. Some of the challenges faced are making the best use of the data held on JASMIN, improving discoverability of data to potential users, providing for the many different user skill levels and domains that want to use the platform, and encouraging collaboration with international partners.

To address these challenges, JASMIN was built in five phases (2012 to present), with increasing growth of infrastructure. The network is non-blocking, moving up to 2 PB per day through the batch cluster at high utilization. Data discovery is a CEDA file-based search, with indexing of 200,000 files per day using 10 different file parsers.

JASMIN’s cluster-as-a-service is a three month project to create ready-made appliances for users to deploy from JASMIN’s cloud, driven with Ansible Playbooks and OpenStack Heat templates. Playbooks are managed by AWX supporting updates and patching to existing clusters.

Philip demonstrated JASMIN, and described various applications: Copernicus projects requiring compute services, ESGF Compute node, NERC DataLab. Activities to contribute to and track are ESA Network of Platforms and OGC Testbeds.

Valerie asked about the relevancy in the elastic search. Philip replied that it has not reached the potential.

Ken asked how much commonality vs. variety exists in the file metadata. Philip replied that the information is in the variables, in the inventory.

## [Big Data Approaches Supporting Market Growth in EO at the Satellite Applications Catapult](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_11.40_Rob%20Fletcher%202.pptx)

Robert Fletcher gave a presentation on Big Data approaches supporting market growth in EO at the Satellite Applications Catapult. He described the Climate, Environment and Monitoring from Space (CEMS) concept for providing access at any point in the value chain depending on the customer requirements. This is cloud computing for the space community, with continuous monitoring of use cases and success stories. Robert demonstrated the Sentinel Data Access Service and described the UKSA International Partnership Programme and the Earth and Sea Observation System.

Chris noted that this is very application-focused. Would an academic person go to these systems? Robert replied that this is commercially focused; academics would go to Jessup.

## Concrete Next Steps for FDA Interoperability and Ecosystem Take-up

Robert Woodcock led a discussion on next steps.

Ken Casey noted that a few years ago the focus was on discovery, but now there is convergence on inventories. Can best practices or standard for these inventories be developed? Valerie noted that the key is to agree upon ways to be more compatible.

Yonsook commented that most of the systems have a discovery aspect with conventional download, and public and custom processing services that can access data at the file level. The agencies do some analysis and the processing helps users get to what they want. The next step is to describe the services so users know what is available. It is a very complex environment but exciting.

Robert Fletcher commented that the demand is increasing because of the synergies that have developed.

Andrea noted that a lot of work has been done in the OGC testbed.

Yonsook reminded that years ago WGISS asked users to identify their pain points, and the main reply was the amount of work they had to do find and obtain the data so they could analyse it. These elements are trying to make the custom services more generalised and portable, massaging the data and providing ARD and services.

Robert and Mirko noted that a lot of the positives and similarities have been highlighted. Despite the fact that there are vast archives of data, many users are still saying that they cannot get access to data. There are slightly different business models being supported.

Robert Woodcock made the following points:

1. There are catalogues everywhere with services and tools; the task is making them discoverable. Valerie suggested sharing various architectures. Chris noted the need to expand knowledge. Mirko commented that this is enlarging on the work that has already been done by the CDA. Chris added that the risk that the more concepts that are discoverable the less they are useful because of determining relevancy. Robert noted that the relevancy will be improved as data providers feed upstream how the data is being used; this is very doable at the present time. Richard said that a few experts can be assigned/help desk access; Cristiano observed that this is available with the thematic communities.
2. Pursuing interoperability between FDA elements: it does appear to be around DCs Jupyter notebooks, etc. but two or three options that CEOS could promote could be documented; more engagement in the FDA team is needed.
3. In terms of engagement:
   1. It would be useful to do this collaboratively the (OGC, work done by SEO).
   2. WGISS showcasing FDA element that is supporting another working group.
   3. WGISS assisting in AquaWatch and ODC.
   4. SDG ad-hoc team have similar issues and needs.

# Data PRESERVATION and STEWARDSHIP



## [ESA EO Data Preservation System](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_13.50_Solution%20for%20Long%20Term%20Preservation%20-%20Workshop%20outcomes.pptx)

Daniele Iozzino\* gave a presentation on the ESA EO Data Preservation System: CBA Infrastructure and its future. He described the data preservation system and the preservation element front-end. The Data Archive Service (DAS) implements the ESA Master Archive through a dedicated service awarded to industry through an open ITT. He described the DAS infrastructure, data flow, and cold back-up archive (CBA). Daniele described the CBA ingestion tool, the Preservation Element Web Portal and hardware, and outlined the processes for business continuity. Daniele concluded with the Large Tape User Group outcomes.

## [Data Preservation Activities at NOAA](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_14.00_Data%20Stewardship%20Activities.pptx)

Nancy Ritchey gave a presentation on data preservation activities at NOAA. The data stewardship goal is to have a trusted archive, strengthening the trust of user communities, partners and the public by quantifiably demonstrating capabilities for ensuring data integrity and reliability over the long term. NCEI is working toward Core Trustworthy Data Repository.

Challenges include location-based processes and guidance, varied roles and responsibilities for data preservation and stewardship, uneven distribution of DSD staff across locations, and geographically dispersed NCEI staff making it difficult to communicate and to build trust. To handle this, NOAA initiated the NCEI Data Stewardship Council to address data preservation and stewardship topics and developed a staffing plan to support data stewardship throughout the organization.

Nancy discussed NOAA’s OneStop data stewardship achievements, noting the ever-increasing data volumes and diversity, and the need for adopting cloud technologies thoughtfully.

Mirko thanked NOAA for their participation in the WGISS Maturity Matrix work.

In response to a question, Nancy acknowledge that the communication strategy is internal because NCEI is very diverse; they are really emphasising that input and feedback is important.

Richard asked if they implement the ISO standard; they do.

## EO Ontologies

Iolanda Maggio introduced the topic of ontologies, which is something that needs to reach interoperability. The ontologies of the different agencies will be presented, with a goal of analysing the commonalities between agencies and develop BP or guidelines for an integrated ontology.

### [NASA](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_14.00_NASA_Ontologies.pptx)

Michael Morahan presented NASA’s ontology: the GCMD keywords. The GCMD is a hierarchical set of controlled vocabulary covering Earth science disciplines that have been evolving for over 25 years. It contains 12 published keyword types, over 11,000 unique keywords, and more than 100 defined keyword relationships. Each keyword created has a universally unique identifier (UUID) that does not change.

Michael displayed a visualisation of keyword relationships and the Global Change Information System (GCIS) implementation process from review to publication process. On the GCIS page you can click on a keyword and it returns everything associated with that keyword (including related publications). Michael made the following points:

GCMD keywords continue to evolve based on feedback from agencies, research universities, and scientific institutions.

Implementing keyword relationships can be used to improve search and discovery of Earth science data and information.

The defined process could be reusable for other providers who want to implement the GCMD keywords and see associated relationships.

Anyone can contribute keywords by contacting the GCMD staff or sending an email [support@earthdata.nasa.gov](mailto:support@earthdata.nasa.gov).

Michael added that they offer RDS, JSON, atom and are using SKOS for setting up the ontology.

### [ESA](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_14.30_ESA_EO_Ontology.pptx)

Andrea Della Vecchia gave a presentation on the ESA EO Thesauri. Its objectives are a common terminology to avoid naming inconsistencies about platform, instrument, instrument type and keyword. FedEO metadata entries are natively ready for being exported into DIF10, according to both IDN metadata guidelines and encoding. This will permit metadata discovery and browse via semantic search.

The platform thesauri provide relationships among platform, instrument, instrument type, scientific keywords and collections. Andrea described the instrument thesaurus and the scientific terms thesaurus. The Platform Thesauri provides relationships among platform, instrument, instrument type, scientific keywords and collections. The Instrument Thesauri links between instrument and instrument type is part of the hierarchy as well (narrower/broader).  The Scientific Terms are linked to GCMD terms and instrument type concepts. Semantic Annotation of EO Resource Metadata aims that every user interface is able to interact.

Activities this year are to provide online access to Thesauri with similar RESTful API capabilities as NASA GCMD. This may be realised selecting an open-source solution. One candidate is SKOSMOS tool (<http://skosmos.org/>), used by UNESCO, MIT.

### [CNES](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_15.10_Ontology%20CNES.pptx)

Richard Moreno gave a presentation on the CNES ontology. The objective is to link/open CNES data to other domains in a transdisciplinary manner, fostering data usage allowing development of evolved search engines, compliant to FAIR principles.

The principles of the ontology are a thesaurus from GCMD that will be enriched to take into account the variety of in-situ data linking granules to other domains, enriching granules description with other information, and accepting that will be used for objective not foreseen.

In response to a question from Cristiano, Richard said that they are using natural language search, but are very focused on the CNES domain.

### Discussion

Chris noted that the GCMD has been built, is maintained, can be mapped with others, is designed to evolve with the community, and is continuously enriched. The concepts that the agencies are trying to represent are in the GCMD, and is working well. It has a good interface that is regularly updated.

Richard observed that having a different ontology (by a given entity) is complicated – it becomes out of date and is difficult to maintain.

Robert emphasised that the GEO KH should use the GCMD, and should be made aware of it.

The topic of ontology is different from thesaurus. The thesaurus would map different names in other communities.

Robert suggested that a session at WGISS-48 on SKOS (with some examples of its capabilities) would be helpful.

**Action WGISS-47-25**: Andrea Della Vecchia to put together a guide (e.g. video, short doc, demo) on the use of SKOS interface and tool by July 31, 2019.

## [RDA Possible Collaboration](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_16.30_RDA%20Possible%20Collaboration_March.pptx)

Iolanda Maggio gave a presentation on the Research Data Alliance. Its vison and mission is that researchers and innovators openly share data across technologies, disciplines, and countries to address the grand challenges of society. RDA builds the social and technical bridges that enable open sharing of data. Any individual or organization can join RDA that has an interest in reducing the barriers to data sharing and re-use and who agrees to RDA’s guiding principles of Openness, Consensus, Balance, Harmonization, Community-driven, Non-profit and technology-neutral.

Iolanda described the various interest groups and working groups of RDA, and its organizational members. She also described a number of RDA recommendations. Working groups of interest to WGIS include:

Data Citation WG

Research Data Repository Interoperability WG

Data Usage Metrics WG

Data Discovery Paradigms IG

RDA/CODATA Legal Interoperability IG

FAIR Data Maturity Model WG

Preservation e-Infrastructure IG

Preservation Tools, Techniques, and Policies IG

RDA/WDS Certification of Digital Repositories IG

Metadata Standards Catalog WG

Disciplinary Interoperability Framework IG

Big Data IG

PID IG

Federated Identity Management IG

## [RDA FAIR Data Maturity Model Working Group](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_16.10_Peng_RDA-FAIR_Data_MMWG_r1.pptx)

Ge Peng\* gave a brief update on the activity of the RDA FAIR Data Maturity Model Working Group. She listed the FAIR Data Principles: Findable, Accessible, Interoperable, and Reusable. However, there is a wide range of interpretations and implementations of FAIRness.

The RDA FAIR Data Maturity Matrix Working Group aims to bring together stakeholders and build on existing approaches and expertise. The WG hopes to develop recommendation of core assessment criteria, self-assessment model and toolset, and FAIR data checklist.

In response to a question from Mingrui Huang, Ge Peng stated that it is intended for individual datasets.

## [WMO Expert Group Possible Collaboration](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_15.50_Peng_WMO_SMM-CD_r1.pptx)

Ge Peng\* gave a presentation on the WMO Stewardship Maturity Matrix for Climate Data (SMM-CD). She began with background information on the World Meteorological Organization (WMO), adding that the WMO established the High-Quality Global Data Management Framework for Climate (HQ-GDMFC) initiative under the technical guidance from CCl and Commission for Basic Systems (CBS) to help its members address data stewardship challenges.

The SMM-CD is working on a stewardship maturity assessment model in the form of a matrix, assessing how well climate datasets are managed and governed. She described its structure and various definitions. The outcomes are the Guidance Booklet, which is under extensive review, and a self-evaluation template. Eighteen global datasets have been assessed by dataset SMEs.

As a specialised agency in weather, water and climate, WMO needs to enhance collaboration for promoting the use of trustworthy data for decision making. SMM-CD provides structure and guidance to measuring and improving quality of climate data management and stewardship. Together with the Manual and Catalogue, the SMM-CD aims to help WMO and its members establish trustworthiness and improve sharing of high-quality climate datasets and associated information.

Final approval of the documents will occur in early June. The documentation be publicly available.

**Action WGISS-47-26a**: Iolanda Maggio to liaise with Ge Peng to find out when the WMO documents on stewardship are approved by October 31, 2019.

**Action WGISS-47-26b**: Iolanda Maggio to liaise with WGCV to collect comments on WGISS and WMO maturity matrices (once the WMO documents are approved) by December 31, 2019.

## [GEOSS DMP IG Proposal](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_16.40%20GEOSS%20DMP%20IG%20Proposal%20-%20Discussion.pptx)

Iolanda Maggio gave a presentation on the GEOSS DMP IG Proposal. She gave background on the DMP and on the Maturity Matrix for Long-Term Scientific Data Stewardship. She described the two exercises that have been performed:

1. DMP IG content completeness measuring through the Data Stewardship Maturity Matrix. She displayed a sample assessment with resulting scores and rating.
2. Created a WIGISS Maturity Matrix with DMP IG simplifying the Data Stewardship Maturity Matrix

The proposal for improvement of the GEOSS DMP IG using the Maturity Matrix involve DMP-1 (Discoverability), DMP-6 (Quality), and DMP-9 (Review and Reprocessing).

Open points include: GEOSS Evolve initiative could no longer be supported but GEO indicated interest on FAIR DMP. RDA is fully involved in FAIR Data Maturity Model with a Working Group already fully active on this topic.

Next steps are the verification of the last version of DMP IG, and finalisation and circulation the GEOSS DMP IG improvement proposal.

**Action WGISS-47-27**: Iolanda Maggio to define way forward for finalization of WGISS Maturity Matrix using WMO matrix and RDA-FAIR WG results; finalise input for DMP by June 30, 2019.

## [PV2020 and Other Relevant Conferences](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_16.50_PV2020.ppt)

Iolanda Maggio gave a summary of recent conferences related to data stewardship. She began with the planned PV2020@CERN. The goals of the PV2020 are to attract more scientific communities, broaden information exchange, sharing of experiences, tools and services, and keep in step with (or ahead of) funding agencies/policy makers in their push for LTDP and OD. She listed the tentative schedule, plans and logistics. The conference will be 12 to 14 May 2020.

Iolanda discussed the BiDS 2019 Conference, the EGU 2019 Conference, the Living Planet 2019 Conference, the IAC-19 Conference, the ISDE11 Symposium, the IDCC 2020, and the EO Phi-Week Conference. Mirko will chair the session on long-term data series at the Living Planet Conference.

# Data DISCOVERY and ACCESS



## [WGISS Connected Data Assets Status Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_13.40_WGISS%20CDA2.pptx)

Yonsook Enloe gave a status report on the WGISS Connected Data Assets (CDA). She reviewed the background and architecture, and the guidelines established toward a federation of connected agency EO data assets. She gave the link for IDN Metadata Registration set of documents and tools, and listed the documents completed since October 2018:

WCDA (WGISS Connected Data Assets) Client Guide: A new document written by the System Level Team with details how to search for collection and granule data at IDN, CWIC, and FedEO

CWIC Data Partner Guide (CSW, OpenSearch)

FedEO Data Partner Guide

CWIC Synchronization Document

WCDA Error Handling

The WGISS search standard documents completed are:

CEOS OpenSearch Best Practices v1.2

CEOS OpenSearch Best Practices Conformance Test Plan

CEOS OpenSearch Developer’s Guide

Yonsook reported that the WGISS CDA status page contains automated (updated daily) metrics for WGISS CDA searchable collections/granules. This can be presented again to the GEO SEC since they often report how many resources they access. A single document, the WDCA Client Guide, gives details on how to search for all the data collections/granules in the WCDA. The System Level Team (SLT) has also determined an error handling approach agreement.

Challenges ahead that the team will tackle are developing FDA capabilities to give users a specialised “chunk” of data, pre-processed and ready to analyse, and adding processing, analytic services/tools to WCDA.

The SLT continues to work with GEOSS as follows:

IDN, FedEO and CWIC accessible from Geo Web Portal and integrated via GEO DAB

Continued participation in the GEO Data Providers-GEO Data Technology workshop

WGISS Connected Data Assets webpage on the WGISS webpage display metrics; this information is given to the GEO Sec; the completely revamped automated page has been given to CEOS CEO and GEO Sec.

Suggestion from FedEO team to share the CWIC Synchronization document with the GEO Sec.

The WDCA Client Guide document will be shared with the GEO data providers

Yonsook concluded with a reminder of the one-page flyer that was developed as outreach to potential tool developers.

## [IDN Update](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_14.00_IDN_Update.pptx)

Michael Morahan gave an update on the International Directory Network (IDN). He noted that they have added FedEO Platform/Instrument relations in KMS and Platform/Instrument relation using KMS API (JSON), and have been working with NOAA NCEI to ingest their metadata. The team is also working with JAXA to ingest GCOM-C collection records, and is continuing the transition from DIF-9 to DIF-10.

Michael noted that in future CEOS providers will be able to register services using the Draft Metadata Management Tool (MMT). Until then, providers can send to the IDN the UMM-S Service JSON file or fill out a template of the fields’ values. Michael displayed an example of the draft Metadata Management Tool (MMT) and listed the required fields in the Unified Metadata Model System (UMM-S). He added that service records can be searched using the CMR SEARCH API.

Michael described two methods for setting up CEOS 2-Step OpenSearch in the IDN and discussed search for tags using CMR API, and how to register services in the CMR (for the IDN).

**Action WGISS-47-31**: Mirko Albani and Yonsook Enloe to contact the KMA, JAXA and JMA to get their assets connected. Due by September 30, 2019.

## [CWIC Update](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_13.40_WGISS%20CDA2.pptx)

Yonsook Enloe gave an update on CWIC. She reported that the CWIC Synchronization document has been completed; this document gives description and details (test scripts) of the daily dataset accessibility testing that is performed automatically. She added that FedEO team implemented the test scripts also. The CWIC Data Partner Guides (CSW, OpenSearch) have been updated and completed. Testing with NOAA One-Stop continues; the system will contain NOAA’s current and heritage EO data collections. Support for China GEOSS system is on hold until the China GEOSS system is ready for continuous access.

The team has discontinued support for the prototype clients: CWIC Start (CSW) client and the CWIC Smart (OpenSearch) client; however, software for both clients is open source for any agency that wants to operate their own client. Support will continue for the CEOS OpenSearch Validator.

## [FedEO Evolution](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_14.30_FedEO_Evolution%20(2).pptx)

Andrea Della Vecchia gave a presentation on the Federated Earth Observation (FedEO) Evolution. He began with a discussion of the software refactoring objectives: porting all FedEO components to Docker and Kubernetes, optimisation of the gateway, the catalogue and the dataset metadata ingestion job, and preserving all functional/interoperability requirements.

Andrea reported on the IDN metadata population using the FedEO Metadata Mediator. A new FedEO Metric web page provides information about collections, granules, reporting about DIF-10 metadata conversion and DIF-10 metadata ready for ingestion into IDN.

Open issues:

FedEO shall be connected to ESA EO Catalogue after its TTO, by June 2019. All ~200 collections shall be visible and >35M products shall be accessible (further products shall be integrated all along 2019).

FedEO currently adopted as translator between DLR and IDN. DLR is preparing all its collection for being exported into IDN, asking to update ESA and GCMD thesauri. DLR shall ask NASA to remove all the old/deprecated collections currently on IDN.

A subset of EUMETSAT collections can be ingested passing through FedEO. Misalignment between EUMETSAT platform values and GCMD thesauri (e.g., “METOP” assuming both “METOP A” and “METOP B”).

All available collection correctly of VITO/JAXA encoded in DIF-10 .

ROSCOSMOS: Waiting for formal agreement.

CCI / CMEMS: Waiting for feedback.

Regarding ECVs, the following tasks are planned

To export ESA-ECV metadata from <http://climatemonitoring.info/ecvinventory/> to IDN passing through FedEO.

The ECV 2.0 is currently under review, under migration toward release 3.0.

Report to ECV team.

Manage mappings for platforms/instruments/organization names. Names used in the ECV Excel (e.g., NASA, CPOM, NASA JPL) are not identical or missing in the GCMD thesauri.

Manage missing platforms/instrument relations in the thesauri.

Manage abstract info, concatenating some columns to have some content in the “ISO abstract”, and prefix it with the title of the column, as ECV Excel does not contain a column with a proper “abstract” describing the collection.

Andrea discussed ESA PDGS Collaborative Environment, showing the conceptual model, and also the ESA Catalogue TTO results. The team is involved in various OGC initiatives, and with the ESA GSTP Project “ESE-ERGO”.

Yonsook noted that the additional 200 collections will be accessible in June.

Richard asked about in-situ data. Andrea replied that the FedEO Evolution has a requirement for including in-situ data.

## [NOAA NCEI](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/2.%20Tuesday%20April%2030/2019.04.30_14.45_NCEI-DataAccess.pptx)

Rich Baldwin\* gave a presentation on NCEI data access. He listed the core objectives and discussed interoperability with CWIC. He reported that NOAA is reorganising support for CWIC and listed the details of what that support entails. He also amplified on NOAAs data diversity and extent.

Rich mentioned the significant challenges that prevent growth and collaboration:

Funding and the timing associated with those funds

Available time and attention

“Good Neighbour” Data Access

Volume of Data

Volume of Metadata

Transitioning Users to Cloud

## IDN Front End Portal, Branding and User Metrics

The SLT led discussion on two actions from WGISS-46:

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

The Best Practice document already contains all the metrics recommended explicitly by the FDA ad-hoc team. With the CDA infrastructure are there data usage metrics that should be collected? Yonsook replied that there are many different clients, the servers can identify what clients are coming in, but most OS clients do not identify themselves. CDA is just an entry point, it is up to the agencies to collect metrics. The searches can be measured within the scope of the CDA, but not how much data is accessed. There should be agreement on which metrics have value, and how to collect them, otherwise the metric is not useful.

Andrew asked if prototyping could be done with one client, and Chris added that even a partial metric might be useful. CWIC and FedEO can discuss the metrics that they already capture; interesting metrics include identification of the more popular datasets, and the number of searches and users.

WGISS-46-19: To discuss way forward to implement a single front-end/portal within the IDN to access the WGISS CDA for data discovery and access

Mirko emphasised that searches should have CEOS branding at the IDN portal to motivate IDN registration by data providers, and that the portal should share a look-and-feel with the CEOS website, while at the same time displaying the identity of the data partner.

It would be ideal to be able to provide users with a link that is coherent and with CEOS branding: A CEOS Flagship portal for discovery.

Andrew and Yonsook agreed that the user should not even know the terms IDN, FedEO, or CWIC. Andrew suggested prototyping, and added that this seems to be something the SEO should do. These activities will be discussed within the SLT.

**Action WGISS-47-17**: WGISS CDA System Level Team (SLT) to:

Identify the best approach to implement a single map-based front-end/portal within the IDN to discover and access CEOS agencies data through WGISS CDA back-end. Agree on way-forward with WGISS-Exec. Due by June 30, 2019.

Start Actions as agreed (e.g. implement CEOS branding on IDN portal and underlying FedEO and CWIC when appearing for second step searches, etc.) Due by December 31, 2019.

Assess advantages and feasibility to implement (part of) CEOS Data Usage Metrics in the IDN portal / WGISS Connected Data Assets. Due by December 31, 2019.

# Data INTEROPERABILITY and USE



## Earth Observation Services Metadata Use Cases

Valerie Dixon introduced the session on EO services metadata use cases that are applicable to the community.

### [ISRO Use Cases](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_09.20_EO_Service_Metadata_ISRO_Usecases.pptx)

Nitant Dube\* presented EO services metadata use cases at ISRO. He introduced the Indian storehouse for space-based weather data at MOSDAC. The storehouse includes the Multi Mission Satellite Data Repository, calibration and validation in-situ data, forecast, tools and utilities, and research and training. Integrating services (WMS, THREDDS, HTTPS, email, RSS Feeds and microservices) across the platforms is a very difficult task. Nitant presented the use case of MOSDAC to disseminate meteorological and oceanic satellite data. Nitant described te services:

WMS is self-describing, GetCapablities response returns a defined XML document providing information about services. THREDDS: catalogue/enhance catalogue files provide metadata about services. HTTPS: Html <meta> tag provides metadata about the html document that is machine readable. Email: User can request the metadata about service by sending blank mail to defined email address. RSS-Feeds: metadata tags in RSS-Feeds can be used to access forecast and nowcast to send alerts.

Micro services: service registry with core services dissemination. Functionality can be exposed as a service; supports design patterns. Microservices at MOSDAC: Safe Beach Microservice, Alert Dashboard Micro-service, and Weather Forecast Micro-service. ISRO is in process of releasing the Microservice service registry URL for operational use.

Challenges: There is very little work and standardisation for services metadata. As services are assumed to be self-describing, the metadata becomes part of the service; this results in non-standardised metadata tags across different services. There is a lack of generalised tools for reading/discovery of service metadata.

Valerie asked Nitant to elaborate on the design patters in microservices service registry API gateway. When providing multiple services, they should be designed in such a way that they are independent of each other. Rather than providing direct access to the gateway, the API gateway provides a single access and provides capability to a single registry. Service registry and discovery is a database but exposed to the user as an XML response.

### [NASA EOSDIS](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_09.20_NASA%20EOSDIS%20Use%20Cases.pptx)

Valerie Dixon gave a presentation on services metadata use cases. She began with definitions of tools and services, how they should be used, and how to structure the metadata. A possible solution is to break out metadata models for each distinct use case; this allows for more agile approach, and distinct metadata models for distinct use cases reduces ambiguity and complexity. But there is more to integrate and maintain, and there is potential for repetitious model entries.

A key use case to consider is End-to-End Services (E2E), which enables a data transformation service to be applied to datasets behind-the-scenes: user finds and selects a data set; user selects the desired service option. Metadata linking dataset and applicable services calls the selected service to perform upon the selected dataset and user receives transformed data set.

Another key use case to consider is Smart Handoffs. This enables transferal of search results and/or context to another tool. User finds and selects a data set; user wants to view or manipulate the dataset in another tool so selects option to transfer search context to associated, enabled tool. User is redirected to the desired tool with previous search context pre-populated and carries on with their analysis in their desired tool.

A third use case to consider is Downloadable Tools. This associates datasets to applicable tools which must be downloaded to the user’s processing platform to use them locally. User finds and selects a data set; user wants to see what Downloadable Tools are applicable to the selected data set so selects desired Local Tool and is redirected to a primary splash page; user downloads the dataset to perform local processing.

Another key use case to consider is Service Entry Resource Formats (SERFs). Legacy records from the GCMD describe tools and services, primarily from the IDN community; migration effort recently concluded.

Yonsook noted that the use cases would be useful if they can be described from the FDA perspective.

Andrea suggested discussing the MAAP joint ESA NASA venture; Chris is participating in the MAAP that has simplified some of the use cases. MAAP will have more on-demand data production; the input datasets and the output datasets need to be understood. A lot is being done with Jupyter notebooks and it is not clear what kind of service that is.

Cristiano noted the link of data to service and service to data. The relationships is just as important.

Valerie remarked that the goal is not to become the next tool repository, but rather to have the tools and services that are usable with the agency’s data.

Robert pointed out that the ODC are working on an applications library, and this has a lot of parallels to the services metadata. What it does not have is the dimension of the cross-section of the data to the service: the toolboxes. Chris added that it becomes even more complicated with the different forms of the toolboxes; there are commonalities, but also uniqueness.

Mirko summarised that the first step is to understand the use cases and attempt to capture all of them. The next step is to define the metadata model. WGISS needs to have a follow-up discussion of what is coming out of the MAAP. How can WGISS address this first step? Chris replied that a superset of all the attributes should be identified, and then be broken up into metadata models; perhaps a remote workshop would help to decide if multiple models or a single model.

**Action WGISS-47-11**: Ad-hoc team led by Valerie Dixon (Michael Morahan, Andrea Della Vecchia, Chris Lynnes, Cristiano Lopes, Richard Moreno) to define proposal for metadata model for services addressing a set of high priority use cases. Proposed model to be circulated to WGISS-all for review and approval. Due by July 31, 2019.

## [Inventory of Software and Tools (Open Source) (FDA-10)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_09.40%20FDA-9%20Inventory%20of%20OSS2.pptx)

Iolanda Maggio gave a presentation on FDA-10 Inventory of software and tools. The OSS Inventory was triggered by the CEOS and WGISS work plans. A survey of CEOS agencies was done on OSS, compiling the reports with relevant metadata. Iolanda displayed tables with the results, as well as some statistics.

Andrew Cherry\* demonstrated the CEOS Software and Tools Inventory. The data was loaded into the database from an existing spreadsheet and is still a work in progress. The plan is that the front page promotes the available software. The actual content is subject to change; input is welcome. The browse capability is mocked up, and contains live links pulled from the data. The site is responsive and includes a mobile-friendly view.

The requirements for the action have been completed, but there is still a lot of work to be done. The normalisation of the table needs to be finalised, and a statement for the OSS inventory management needs to be provided. A simple dashboard with new filters needs to be developed, and the inventory needs to be advertised. In future the ingestion of this information (owner, contact point, licensee) can be automated.

Robert commented that there is still a lot to learn about discovery of tools. The SEO library and the services metadata group need to be included.

**Action WGISS-47-12**: Michael Morahan to register services/tools from inventory tables (FDA elements and SW/tools) into IDN starting from the ones presented at WGISS-47. Due by WGISS-48.

**Action WGISS-47-19**: Iolanda Maggio and Andrew Cherry to finalise SW Inventory database tool, present to WGISS-Exec and publish on CEOS website. Due by July 31, 2019.

## [Carbon Data Portal Prototype (CARB-15)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_10.10_CARB%2015%20Status.pptx)

Liping Di gave a presentation on status and way forward for the Carbon Data Portal prototype. He listed the goal and objectives, the composition of the team and other details. The portal was released to CEOS WGClimate and GCP for evaluation and feedback, which require a set of new capabilities and improvements. Since WGISS-46, the major activity has been implementing the new capabilities and improvements: Enhanced keyword autocomplete and improved carbon categories, performance improvement, search-by-polygon, geocoding, information pages, keyword-based search enhancements, multiple keyword combination, visibility of key carbon datasets, filtering, and clipping-by-polygon data access.

Work in progress includes implementing more filters and extending support of more polygon data format. The new version of portal will be released to the WGClimate, GCP, and any other carbon community for further evaluation and test. The team is also involved in supporting the CEOS Chair initiative by customising the portal to support the forest carbon project at Mekong River basin. This will be an especially customised portal that is pre-populated products and with specific datasets for Mekong.

Mirko thanked NOAA and NASA for supporting this activity. Ken reported that Steve Volz has directed them to maintain and support the portal; over the next five years the long-term sustainability will be evaluated. What has been done already satisfies the task, and future efforts will depend on the VNSC initiative. Liping noted that the GEOSS community needs a portal, so it would be a good idea to review the GEO Work Plan from this perspective.

**Action WGISS-47-20**: Liping Di and Kenneth Casey to confirm EXEC of closure of Action CARB-15 (i.e. when next version of Carbon portal is available online). Mirko will then inform CEOS SEC about closure and that activities will in any case continue. Due by June 30, 2019.

## [White Paper on Single Sign-On (SSO) Authentication (DATA-14)](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_10.30_SSO-WhitePaper.pptx)

Damiano Guerrucci gave a presentation (on behalf of Marco Leonardi\*) on the white paper on SSO. He observed that relevant concepts in the identity and access management domain are (federated) SSO, federated access, authorization, and accounting. Each term identifies a possible use case for data accessibility and data discoverability scenario. All the different entities should consider their access requirements in order to contribute to the overall use cases definition. The white paper on the federated single-sign-on aims at supporting the identification and definition of the possible data access use cases.

Chris asked if they are involved in the MAAP project (they are). He added that there is another access beyond accessing the data: there is also the access the agency’s services.

**Action WGISS-47-21**: Marco Leonardi to circulate first version of SSO White Paper to WGISS-all for review and comments. Due date by May 31, 2019.

## [Recovery Observatory](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_11.00_Recovery%20Observatory.pptx)

Richard Moreno discussed the Recovery Observatory (RO), a former joint work between WGISS and WGDisasters. Space agencies are well organised for support immediately after a disaster, but there was nothing for the long-term reconstruction, and the RO grew from this.

Dotcloud (a generic tool) displays full resolution data for authorised users, and moderate resolution for other users. The catalogue of data is acquired in the framework of Projects and its value-added products, and it includes a community of users: community animation, exchanges between actors, and user feedback.

Richard continued with a status overview of RO Haiti; he listed thematic products, user workshops and results. He concluded saying that there is interest in a generic RO.

Mirko asked who is operating the RO; CNES is doing so, but an agency can ask for and operate the system. A white paper is being written right now by the WGDisasters that identifies lessons learned.

**Action WGISS-47-22**: David Borges to share with WGISS the WGDisasters Generic RO white paper when it becomes available, to identify any integration opportunities where WGISS could support WGDisasters Generic RO technical requirements.

## [GEO Knowledge Hub and WGISS Contribution](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/3.%20Wednesday%20May%201/2019.05.01_11.20_GEO_KNOWLEDGE_HUB_AND_WGISS.pdf)

Chris Lynnes gave a presentation on the GEO Knowledge Hub and the WGISS contribution. The proposed EO Knowledge Graph is meant to connect data, applications and people: connect the main elements of EO knowledge and context in a way that is machine-readable, human-usable, and curatable. Important entities include publication, result, exploitation system, instrument, satellite, person, workflow, dataset, application, discipline, software, model, essential variable, project, organization, and training material. The connections among these entities were discussed.

Chris listed several examples, and displayed potential use cases that answer questions for developers and users and connect systems. Suggested pilots:

Knowledge graph for MAAP project: Multi-Mission Algorithm and Analysis Platform

Knowledge Graph for GEOGLAM Crop Monitor

Demonstration queries

The key is interest in all the elements that contribute to the knowledge and the relationships between elements. Elements that are leveraging and learning from current efforts are:

Global Change Information System (GCIS)

Global Change Master Directory (GCMD)

Committee on Earth Observing Satellites (CEOS)

Tool Inventory

Unified Metadata Model (Services ↔ Collections)

JavaScript Object Notation-Linked Data (JSON-LD)

OGC Testbed 15 EOPAD thread (JSON + Services)

Every entity that may be added should have a query for the relationships with other elements, and a use case for it. For some elements a unique identifier does not quite exist. The KH is targeted at the flagships: scaling up to CEOS user base is a different question.

Chris suggested that GEO handle the requirements design implementation and operations, and WGISS handle requirements and design, graph collection mechanisms, graph population, and authority engineering.

Liping inquired about the difference between this and the ontology; Chris replied that the emphasis is on the relationships between the entities. The weak point is that it is easy to track additions and changes, but very difficult to track deletions.

WGISS has action in place to better understand how it fits into the GEO Work Plan, and this might be one option. Robert recommended a more nuanced version of the table.

**Action WGISS-47-23**: Chris Lynnes to provide a more nuanced version of the “Entities and Connections” for the GEO Knowledge Hub by July 31, 2019.

**Action WGISS-47-24**: Mirko Albani to distribute to WGISS the draft version of the GEO Work Plan by May 20, 2019.

# TECHNOLOGY EXPLORATION



## Data Services in the Cloud

### [Advanced Geospatial Techniques: Aiding Earth Observation Applications](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_09.25_Advanced%20Geospatial%20Techniques.pptx)

Nitant Dube\* gave a presentation on advanced geospatial techniques for aiding Earth Observation applications. He discussed the geospatial data challenges: More data are becoming unstructured and volumes are growing exponentially, requirements of data availability/access on demand are expanding, and data security and interoperability to support new technologies/applications.

Nitant described features of geospatial databases such as geo-relational, spatially-enabled, and NoSQL. He also described the features of in-memory and cloud databases, adding that a hybrid is also an option. Geospatial data processing include big data, cloud computing and real time data processing.

Nitant discussed a new paradigm of augmented data analytics where the emphasis is on automation of data discovery, and preparation and insight discovery using machine learning techniques that automatically generate actions from insights. EO applications include

Automated feature extraction using deep learning techniques (Applications in Urban management)

Automated event tracking using machine learning techniques (Oil Spill, forest fires, cyclone, Heavy rains)

Natural Resource and Crop Monitoring (Monitoring of Forest cover and Crop yields.

High Definition Maps: Autonomous Navigation

Earth-Digital Twin

Data cubes: Provides a portable environment of analysis ready products for quick analysis and interpretation by end users.

Dwarf data cubes: Dwarf is a highly compressed structure for computing, storing and querying data cubes. Dwarf identifies prefix and suffix structural redundancies and factors them out.

Nitant recommended frequent interactions on emerging technologies that can impact EO agencies.

Chris asked if anyone has started looking at the dwarf DCs for EO to see how applicable they are. Nitant said ISRO is working on converting a DC to a dwarf DC.

### [OGC Testbed 14 and 15](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_09.45_OGC%20Testbed%2014%20and%2015.1.pptx)

Cristiano Lopes began his presentation with a reminder of OGC testbeds, initiatives and programs. He continued with a detailed description of OGC T14 EOC. The achievements of this testbed are that there are five different companies working together in the OGC manner (consensus-based) to:

Propose and write-down a standard Application Package format (“Application Package” ER)

Propose a WPS-T JSON/REST API - partly developed in coordination with the WPS 2.0 SWG - and write-down best practices and experiences about the EMS and ADES, two essential components of a EP ( “EMS & ADES Best Practices and Results” ER)

In addition, one client, three different EMS and three different ADES implementations are able to interact in a standardised and interoperable way using the proposed AP format and the proposed WPS-T REST/JSON API, and integrating several different TEPs and MEPs, authN/authZ, catalogue in flow.

Cristiano continued with a detailed description of OGC T15 EOPAD with NASA/ESA sponsorship. He listed the sponsored activities, use case and deliverables. Cristiano noted that Testbed 16 has just been announced with a call for sponsors. Cristiano concluded that:

ESA is actively working within OGC to experiment, prototype and define standards for EO Exploitation/Analytics, following the OGC Processes (IP to SP)

These activities are highly visible in Industry and several Organisations (from different domains) have demonstrated interest in cooperation.

OGC is right now performing the ground work for its next generation of Web Services, and CEOS Agencies have the unique opportunity to influence this according to our needs.

CEOS agencies should also be aware of the work being done in the OGC IP for ML/AI, Federations/Cloud Use, etc.

Mirko asked if there is an intention to use this API for interoperability in Testbed 14; Cristiano replied that it is the next experiment. Mirko suggested that WGISS exploit the work that has been done Testbed 14 to the same WGISS activities. Cristiano said this is experimental, not a fixed solution. The outcomes of T14 are now stable. The reports are long and detailed. The key is to have people reading those reports, and feed information to OGC for future.

**Action WGISS-47-14:** Cristiano Lopes to inform WGISS-Exec about outcomes of ongoing OGC standardisation activities (including OGC test-beds) and ask for any additional support as needed. Due by WGISS-48.

**Action WGISS-47-35**: CDA SLT and Ad-hoc team led by Valerie Dixon (Michael Morahan, Andrea Della Vecchia, Chris Lynnes, Cristiano Lopes, Richard Moreno) to look into OGC Testbed 14 results and assess possible use and application within WGISS FDA interoperability activities. Due by July 31, 2019.

### [CNES Initiatives on Big Data and Cloud](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_10.05_CNES%20Services%20in%20the%20Cloud.pdf)

Pierre Lassalle gave a presentation on CNES initiatives on big data and cloud. He began with the context and motivation for CNES. Recent R&D studies on big data processing include image processing using Cloud and Big Data technologies, and 3D Image processing using Big Data technologies on the cloud. Richard also discussed a new generation of image processing chains to adapt algorithms to new logic paradigm of recent Big Data Frameworks. Tasks include selection of a Big Data framework for maturity, separation of data management and core algorithms, handling node failure and preserving data integrity, and need for simple work orchestration. The current decision is the selection of Apache Spark framework but Dask is getting more and more our attention.

Richard discussed the DAG adaptation of image processing algorithms, giving the execution strategy and functional approach. He also discussed 3D image processing using Big Data technologies on Cloud. The main goal is to validate the ability of massive image production using Spark in a Cloud production-like environment and the flexibility and scalability of such technologies, and to develop image processing chain for Big Data and Cloud environment and introduction of DevOps tools and methods in a production-like environment. This will establish a new reference image processing framework for new Earth Observation missions.

### [EOSDIS Approach to Data Services in the Cloud](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_10.25_EOSDIS%20Cloud.pptx)

Chris Lynnes gave a presentation on EOSDIS approach to data services in the cloud. Data transformation services in the cloud include subsetting (variable, spatial, temporal), reformatting (shapefile, etc.), re-gridding/re-projection/ortho-rectification, stitching/mosaicking, and dataset-specific pre-processing.

Chris discussed differences with cloud storage and processing, and user interaction patterns. He also discussed the benefits of reuse and the reuse targets, and the considerations of managing the cost of using the cloud.

Chris concluded with discussion of user-application interface convergence in Jupyter, and search analysis convergence.

## Artificial Intelligence and Machine Learning

### [Workshop on Leveraging AI in the Exploitation of Satellite EO and Numerical Weather Prediction](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_11.10_Workshop%20on%20Leveraging%20AI%20in%20EO%20and%20NWP.pptx)

Sid Boukabara discussed the motivation and main takeaways of the 1st Workshop on leveraging AI in the exploitation of satellite EO and numerical weather prediction. He discussed the main motivations and highlights for the workshop, noting that evidence was presented that AI has a significant potential for being a positively disruptive technology to support the NOAA mission (more skills, more efficiency). Applications of AI for NOAA-related issues include detection, translation, prediction, enhancement, emulation, and parametrization.

Sid observed that two major drivers will continue to make AI attractive in EO: Significantly increased efficiency (therefore driving down cost) and enhancement of skills (accounting for unknown or difficult to model phenomena, etc. present in the data).

Chris asked about capturing semantics. Sid replied that he learned during the interactions at the meeting of a tool that could be used.

Paul Briand commented on the need for bridging the knowledge gap, which can be done by developing partnerships, interacting with AI scientists and transferring the knowledge.

### [AI for EO and EO for AI](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_11.30_AI4EO4AI_4.pptx)

Sveinung Loekken\* gave a presentation on AI4EO and EO4AI and what has been learned from Climate Informatics. With the massive increase in observational capability and data availability, the 4th paradigm applied to EO (and related) is still a promise. AI is the enabler to increase uptake and impact of EO data: AI4EO.

By developing collaboration among European AI and EO the socioeconomic benefits of EO can be maximised. By applying the unique EO assets, the development of AI can be accelerated, resulting in an AI4EO4AI synergy.

Data availability is a major enabler for second wave AI; AI4EO and X-AI potential in 3rd Wave AI. EO is AI’s next big data source. The EO/RS community has definitely picked up on AI, but EO data is not yet driving research in AI.

To engage the AI community, Sid suggested looking for impact: Less Big Data, more ‘good numbers’, mission driven, addressing grand challenges and preparing for 3rd wave AI. Invent, discover, bridge Earth sciences and social sciences.

### [Machine Intelligence towards Tomorrow’s Earth Science Data Systems](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_11.50_Machine-Intelligence-towards-Tomorrows-Earth-Science-Data-Systems.pptx)

Manil Maskey\* discussed machine intelligence toward tomorrow’s Earth science data systems. He emphasised that applying ML is critical to fully exploiting the large archives of Earth science data. Novel machine learning models show promise across disciplines by outperforming established benchmarks in prediction, forecasting, classification, and recommendations. Open and large datasets from multiple measurement systems now provide a unique opportunity to innovate new ways of analysis and obtain valuable insights. Mature ML systems can be applied to augment data systems operations, search/discovery, access, and use. The cloud platform enables scaling ML applications with collocated data.

Challenges include issues with training dataset (Scarcity, Ensuring data access at a granular level, Inconsistency, Requires SMEs, disparate systems/sensors). Other challenges are integration of heterogeneous data sources, use of suitable ML architectures and algorithms for a problem, high barriers to entry (requires team science), expertise in problem-specific optimisation, and lack of human-understandable solutions for interpreting ML model.

Manil displayed diagrammatically a vision for ML, and listed near-term priorities and ongoing efforts. He showed several examples of the phenomena detection portal.

### [A New Large-Scale Sentinel-2 Benchmark Archive to Drive Deep Learning Studies in Remote Sensing](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_12.10_A%20New%20Large-Scale%20Sentinel-2%20Benchmark%20Archive%20to%20Drive%20Deep%20Learning%20Studies.pptx)

Begum Demir\* gave presentation on a new large-scale Sentinel-2 benchmark archive to drive deep learning studies in remote sensing. She listed existing benchmark archives in remote sensing, and examples of state-of-the-art solutions. She also discussed limitations on existing archives in RS, and discussed BigEarthNet, a large-scale Sentinel-2 benchmark archive. Begum introduced Three-Branch CNN TB-CNN which includes three different convolutional branches specifically designed for different spatial resolutions of Sentinel-2 bands; each branch acts as a feature extractor for different resolutions. She also presented experimental results.

Conclusions and future developments:

Introduced a large-scale benchmark archive that consists of 590,326 Sentinel-2 image patches annotated by multi-labels, for RS image understanding.

BigEarthNet will make a significant advancement for the use of deep learning in RS by overcoming current limitations of the existing archives.

Regularly enrich the BigEarthNet by increasing the number of annotated Sentinel-2 images.

Working on designing and implementing a scalable architecture for massive processing and analysis of images in the BigEarthNet.

The BigEarthNet archive is open.

### [Multi-task Deep Learning from Sentinel-1 SAR: Ship Detection, Classification and Length Estimation](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_12.30_ShipDetectionCharacterization.pdf)

Clement Dechesne\* gave a presentation on multi-task deep learning from Sentinel-1 SAR: Ship detection, classification and length estimation.

Chris asked about sub-pixel error. Clement replied that in SAR images it is not possible to count the pixels in a footprint to get the size. Surrounding pixels contain information of the central pixel.

Paul observed that there are a lot of issues to match the SAR and the IES; SAR and IES are not captured at the same moment.

### [CNES Initiatives on AI](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_12.50_CNES%20AI.pdf)

Pierre Lassalle gave a presentation on CNES initiatives on AI. He discussed an automatic car counting project. The challenge is to build a predictive model able to detect cars on VHR Pleiades satellite images to cross the car number over interest areas with other indicators to predict economic trends in field of interest. They use labelling tools based on segmentation algorithms to help building training and test datasets, and post-processing to filter false positive detections by looking at local NDVI mean.

Pierre also discussed an automatic building counting project. The context is preparing the applications of the 3D data provided by the future mission CO3D. Training is done using Ortho + DEM on Toulouse city area with14157 tiles for training dataset and 6067 tiles for test dataset. There are only two classes: no-building (0) and building (1); labelling is made with OpenStreetMap, and a 3-Fold stratified splitting strategy is used. Perspectives to building detection from 3-D data: Improve the correlation methods to find corresponding points between stereo images; noisy DEM products; extrusion of building from classification results.

Pierre concluded with a discussion of transfer learning to build predictive model on unlabelled dataset, taking full advantage of trained models on a specific dataset. Pierre described the project, and listed these perspectives: Solve the scalability issue when using Gromov Wassertein since it is a quadratic algorithm. Currently, this method is not relevant for Earth Observation data; exploring alternative solutions for optimal.

## Discussion and Way Forward

Chris Lynnes led a discussion on way forward for data services in the cloud and AI and machine learning.

Mirko reminded that there is a CEOS action to WGISS to explore emergent technologies and summarise analysis in a white paper, identifying the most useful use cases.

Chris commented that this is technical territory that might be difficult to explain in a white paper that adds value. Another approach would be to describe where there are implementations, and the problems that need to be solved.

Mirko reminded that this action originated during the technical workshop, so the goal is a generic white paper that discusses how AI could help understand EO data topics, or to summarise what agencies are doing in these areas in terms of benefits and limitations. Another topic could be labelled datasets and how the agencies are using them.

Andrew noted that in the past WGISS has queried the community to obtain the topics for the white paper, and this approach might work. The next step would be developing an archive or something that could be shared among agencies. Paul suggested writing about leveraging and using a common standard.

Chris suggested beginning with a use case white paper although use cases that are transformative have not yet been identified; it would not be appropriate to give the impression to the principals that WGISS is expert and has all the solutions.

In conclusion, the goal is for the Technology Exploration team to produce a white paper describing a high level summary of AI and ML.

**Action WGISS-47-36**: Technology Exploration IG to define a template suitable for gathering inputs from agencies (domain, use case, datasets, labels, dataset accessible accessibility) in the areas of Machine Learning and Artificial Intelligence. Due by August 31, 2019.

# Agency and Liaison Reports



## [Hungarian Space Office Liaison Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_14.40_HSOv2.pptx)

Gabor Remetey gave a report on the Hungarian Space Office (HSO), with insight into MFTTT WG4SDG, EO industry and a DIAS developed for the benefit of the Public Administration. The Hungarian Space Office is re-setting the Scientific Council on Space Research, re-setting the Hungarian Space Board, and establishing a new, comprehensive National Space Strategy. The MFTTT WG4SDG mission is to strengthen stakeholders’ engagement, collaboration, and capacity building in use of GEO/GI for SDGs.

Gabor gave an update on recent activities, including involvement with GEO and CEOS. He also gave general information on the Hungarian Space Industry cluster HUNSPACE. Three major stakeholders are Airbus DS GEO Hungary, GEODATA Services Ltd, and Lechner Knowledge Center.

Gabor discussed the conceptual model of operation of a DIAS for the benefit of Public Administration: EO Information System (FIR) in which HSO is involved in the development. He also listed other news and activities.

Gabor reported that there is a serious consideration by HSO to join CEOS; it is anticipated that the newly appointed advisory boards of the Ministerial Commissioner will discuss this issue soon, and the ministerial commissioner responsible for the Hungarian space research and activities will act based on the advice of the boards.

After WGISS-47, CEOS WGCapD will be contacted in order to discuss engagement opportunities or a stakeholder entity named MFTTT WG4SDG via HSO.

## [JAXA Agency Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_14.50_JAXAReport.pdf)

Makoto Natsuisaka gave a report on the Japanese Aerospace Exploration Agency (JAXA). He listed JAXA’s past, current and future EO satellites and sensors, presented JAXA’s EO strategy, and described JAXA’s cooperation with CEOS.

Makoto described the GOSAT-2 (Greenhouse gases Observing SATellite-2) joint mission with Ministry of the Environment (MOE) and National Institute for Environmental Studies (NIES). It is jointly launched with KhalifaSat of Mohammed bin Rashid Space Centre.

Makoto presented the ground segments for the JAXA EO Satellites G-Portal and JAXA EO portals connections with GEO/CEOS Portals. The GCOM-C product release will contribute to global monitoring for the climate change.

Makoto reported that an open and free platform for EO data “Tellus” developed by METI (Ministry of Economy, Trade and Industry) has started to be operated since February 2019. JAXA is supporting the activities by providing ALOS/AVNIR-2, ALOS/PALSAR, AW3D30, and GSMaP data to the platform.

Makoto concluded with JAXA’s space mission long-term plan.

## [NASA Agency Report](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_14.20_WGISS47NASAReport.pptx)

Andrew Mitchell gave a report on the NASA’s Earth Observing System Data and Information System (EOSDIS).He described EOSDIS’ role, core services, and components, and listed its data access role centralised reusable capabilities.

Andrew described the plans and status for transition to the cloud. The Earthdata Cloud project will improve the efficiency of NASA’s data systems operations, increase opportunity for researchers and commercial users to access/process PBs of data quickly without the need for data management and will allow for a transparent/extendable open source processing framework.

Andrew discussed the prioritising of EOSDIS data products for cloud migration. He concluded with key metrics of EOSDIS data and products.

Mingrui asked if EOSDIS provides regular assessments, annual reports for each data centre. Andrew replied that they have comprehensive reports and also conduct regular surveys.

Andrew asked if ALOS3 will go into METI; Makoto replied that it will.

## [ISO TC211](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_15.00_ISOTC211Report.ppt)

Liping Di, WGISS liaison with ISO TC211, gave a report on its status. He began with background and scope of the ISO TC 211 standard, listing the five current working groups. He listed information on each of the following standards:

ISO 19115:2003- geographic information-metadata: old version deprecated.

ISO 19115-2:2019 - Metadata - Part 2: Extensions for imagery and gridded data: old version deprecated.

ISO 19115-1:2014 – Geographic Information-Metadata-Part 1: Fundamentals: New version of metadata standards have been published.

ISO 19115-2:2019: Geographic information -- Metadata -- Part 2: Extensions for acquisition and processing

ISO 19115-3:2016 is the XML implementation schema of ISO 19115-1:2014. New version of metadata standards have been published

ISO 19157:2013 – Geographic Information – Data Quality; combined all data quality related metadata standards in ISO TC 211. New version of metadata standards have been published

ISO 19157-2:2016 Geographic information -- Data quality -- Part 2: XML schema implementation. New version of metadata standards have been published.

Liping noted the published remote sensing imagery standards and ongoing projects that WGISS can contribute to:

ISO 19123-1 Coverage – Part 1: Fundamentals

ISO 19130-3 Geographic information – Imagery sensor models for geopositioning - Part 3: XML schema implementation

ISO 19150-4 Geographic information - Ontology, Part 4: Service ontology

ISO /TS 19163-2 Content components and encoding rules for imagery and gridded data -- Part 2: Implementation schema

ISO 19165-2 Preservation of digital data and metadata -- Part 2: Content specifications for earth observation data and derived digital products. Currently at DIS stage

Liping reported the standard projects being proposed:

ISO 19159-4 Calibration and validation of remote sensing imagery sensors and data -- Part 4: Passive Microwave Sensors

ISO XXX Calibration and validation of remote sensing data and derived products – multiple parts

Liping concluded noting that ISO TC 211 holds its plenary and working group meeting twice per year. The 47th plenary was held in November 2018. The 48th plenary will be held in Maribor, Slovenia in the first week of June 2019. CEOS WGISS can contribute significantly to ISO standard development, especially in remote sensing and data systems.

Liping will need WGISS input for the liaison report to the 48th plenary. He asked for slides of background on WGISS the first week of June.

**Action WGISS-47-28:** Mirko Albani to send to Liping Di a packet of slides describing CEOS/WGISS for presentation at the ISO 48th Plenary by May 31, 2019.

# WGISS Plenary, Part II



## [WGISS Summary](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_15.30_WGISS%20Chair%20Summary.pptx)

Mirko Albani summarised WGISS-47 with the following points:

WGISS successfully promotes collaboration in the development of systems and services that manage and supply EO data.

WGISS is seeking support for WISP.

WGISS will support the CEOS Study Team for a working group on information provision.

WGISS is seeking a vice-chair for the period 2019-2021.

WGISS actively participated in the 2019 SIT Technical Workshop.

WGISS-48 meeting will be held October 8-11 in Hanoi, hosted by VNSC.

WGISS discussed the 2019-21 Work Plan.

WGISS discussed collaboration with WGCV, WGCapD, WGClimate, ISO TC211, SEO, CEOS SDG AHT, and GEO.

Eleven FDA elements were discussed and demonstrated, and FDA-08 and FDA-09 actions addressed. FDA activities in WGISS proceeding along three lines:

* + Making FDA services/tools/elements discoverable/accessible through WGISS CDA (FDA-9, FDA-10, FDA-14) – and agency services catalogue harvesting/registration.
  + Pursuing interoperability between FDA elements through a manageable number of possibly standardised interfaces (FDA-8). Data Cubes interoperability (ODC / HMDC / DCFS), and follow-up standardization activities in OGC.
  + Showcases of FDA Elements supporting CEOS/GEO Initiatives: Data Cubes for WGCV – two versions, ESA and ODC DEA; GEOhazards TEP in support to WGDisasters; WGISS to support AquaWatch and CEOS SEO ODC with data pipelines from existing FDA sources

WGISS Connected Data Assets reported status on IDN, FedEO, CWIC, and NOAA NCEI.

WGISS CDA GUI Interface and Metrics Discussion resulted in projected activities.

EO services metadata use cases were discussed; next steps are to define services metadata model for an initial set of use cases; register services in the IDN; linking services to data.

Activity is ongoing on the CEOS Carbon Portal, CEOS Software and Tools Inventory, and CEOS federated access.

WGISS potential involvement in generic RO.

Data stewardship investigation on EO ontologies, and maturity matrices.

Technology exploration of cloud services, artificial intelligence and machine learning.

## Review WGISS-46 Actions

The following WGISS-46 actions are still ongoing:

**Action 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. (On Hold)

**Action 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 by end of March 2019. (In Progress)

**Action 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 by end February 2019. (In Progress)

## [WGISS-47 Actions](http://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-47/4.%20Thursday%20May%202/2019.05.02_15.50_WGISS-46%20Actions.xlsx)

**Action WGISS-47-01**: Michelle Piepgrass to finalise and post the WGISS organizational structure chart. Due by May 20, 2019.

**Action WGISS-47-02**: Iolanda Maggio to update the WGISS Work Plan based on the CEOS Work Plan 2019-2021. Due by June 15, 2019.

**Action WGISS-47-03**: Michelle Piepgrass to contact Cindy Ong to obtain the BP and Guidelines for WGCV. Due by May 31, 2019.

**Action WGISS-47-04**: Mirko Albani to request SIT Chair Team to include a WGISS representative to the CEOS Study Team on WGs. Due by May 20, 2019.

**Action WGISS-47-05**: Mirko Albani and Robert Woodcock to provide SIT Chair Team with the name of WGISS representative for the GEO Implementation Task Force to be proposed to GEO Secretariat. Due by May 20, 2019.

**Action WGISS-47-06**: Iolanda Maggio to complete the Google Dataset Search Tool CEOS report using input from NASA/ESA/DLR/NOAA. Due by June 30, 2019.

**Action WGISS-47-07:** WGISS-Exec to follow-up GEO Work Plan preparation and identify where WGISS can contribute with: (Due by June 30, 2019)

* Data Management Principles (DSIG)
* WGISS Connected Data Access
* Future Data Architectures

**Action WGISS-47-08**: Robert Woodcock to define next steps and schedule for FDA-08 White paper. Due by May 20, 2019.

**Action WGISS-47-09**: Mirko Albani and Iolanda Maggio to complete FDA-09 Inventory deliverable. Due by July 31, 2019.

**Action WGISS-47-10:** ESA to present progress of DCFS Data Cube at next WGISS. Due by WGISS-48.

**Action WGISS-47-11**: Ad-hoc team led by Valerie Dixon (Michael Morahan, Andrea Della Vecchia, Chris Lynnes, Cristiano Lopes, Richard Moreno) to define proposal for metadata model for services addressing a set of high priority use cases. Proposed model to be circulated to WGISS-all for review and approval. Due by July 31, 2019.

**Action WGISS-47-12**: Michael Morahan to register services/tools from inventory tables (FDA elements and SW/tools) into IDN starting from the ones presented at WGISS-47. Due by WGISS-48.

**Action WGISS-47-13**: Andrea Della Vecchia and Robert Woodcock to define with SEO a roadmap and short term activities on interoperability between ODC / HMDC / DCFS Data Cubes (e.g. Jupyter notebooks with XArray API and WCS / WPS). Due by June 30, 2019.

**Action WGISS-47-14:** Cristiano Lopes to inform WGISS-Exec about outcomes of ongoing OGC standardisation activities (including OGC test-beds) and ask for any additional support as needed. Due by WGISS-48.

**Action WGISS-47-15**: Andrea Della Vecchia and Robert Woodcock to define with WGCV colleagues (Medhavy Thankappan, Phillipe Goryl, and Cindy Ong) coordinated roadmap and short term activities on WGISS-WGCV Data Cubes (ESA and ODC DEA) and possibly demo to be shown at SIT TW in September 2019. Due by June 30, 2019.

**Action WGISS-47-16:** Robert Woodcock to liaise with AquaWatch representatives to gather their requirements/needs and further define how WGISS could support. Due by June 30, 2019.

**Action WGISS-47-17**: WGISS CDA System Level Team (SLT) to:

* Identify the best approach to implement a single map-based front-end/portal within the IDN to discover and access CEOS agencies data through WGISS CDA back-end. Agree on way-forward with WGISS-Exec. Due by June 30, 2019.
* Start Actions as agreed (e.g. implement CEOS branding on IDN portal and underlying FedEO and CWIC when appearing for second step searches, etc.) Due by December 31, 2019.
* Assess advantages and feasibility to implement (part of) CEOS Data Usage Metrics in the IDN portal / WGISS Connected Data Assets. Due by December 31, 2019.

**Action WGISS-47-18**: Chris Lynnes, Yousuke Ikehata and Kent Ross to organise a webinar in September focusing on existing (pre)operational Exploitation and Application Platforms targeting the academic community. Following webinars should target the new developing countries on the same topics. Due by September 30, 2019.

**Action WGISS-47-19**: Iolanda Maggio and Andrew Cherry finalise SW Inventory database tool, present to WGISS-Exec and publish on CEOS website. Due by July 31, 2019.

**Action WGISS-47-20**: Liping Di and Kenneth Casey to confirm EXEC of closure of Action CARB-15 (i.e. when next version of Carbon portal is available online). Mirko will then inform CEOS SEC about closure and that activities will in any case continue. Due by June 30, 2019.

**Action WGISS-47-21**: Marco Leonardi to circulate first version of SSO White Paper to WGISS-all for review and comments. Due date by May 31, 2019.

**Action WGISS-47-22**: David Borges to share with WGISS the WGDisasters Generic RO white paper when it becomes available, to identify any integration opportunities where WGISS could support WGDisasters Generic RO technical requirements.

**Action WGISS-47-23**: Chris Lynnes to provide a more nuanced version of the “Entities and Connections” for the GEO Knowledge Hub by July 31, 2019.

**Action WGISS-47-24**: Mirko Albani to distribute to WGISS the draft version of the GEO Work Plan by May 20, 2019.

**Action WGISS-47-25**: Andrea Della Vecchia to put together a guide (e.g. video, short doc, demo) on the use of SKOS interface and tool by July 31, 2019.

**Action WGISS-47-26a**: Iolanda Maggio to liaise with Ge Peng to find out when the WMO documents on stewardship are approved by October 31, 2019.

**Action WGISS-47-26b**: Iolanda Maggio to liaise with WGCV to collect comments on WGISS and WMO maturity matrices (once the WMO documents are approved) by December 31, 2019.

**Action WGISS-47-27**: Iolanda Maggio to define way forward for finalization of WGISS Maturity Matrix using WMO matrix and RDA-FAIR WG results; finalise input for DMP by June 30, 2019.

**Action WGISS-47-28:** Mirko Albani to send to Liping Di a packet of slides describing CEOS/WGISS for presentation at the ISO 48th Plenary by May 31, 2019.

**Action WGISS-47-29**: CDA SLT to analyse what kind of effort and solution can be accomplished by WGISS to automate future ECV Inventory updates [last slide of presentation].

**Action WGISS-47-30**: CDA SLT to send to WGClimate the list of elements in the inventory that should be made mandatory so they can be registered automatically. By June 30, 2019.

**Action WGISS-47-31**: Mirko Albani and Yonsook Enloe to contact the KMA, JAXA and JMA to get their assets connected. Due by September 30, 2019.

**Action WGISS-47-32**: Mirko Albani to send information to Joerg Schultz about the Purge Alert procedure. Due by June 30, 2019.

**Action WGISS-47-33**: WGISS to give a demonstration of the Carbon Portal at the WGClimate meeting in September.

**Action WGISS-47-34**: Mirko Albani to liaise with Marc Paganini on discovery and access of data and platforms identified in the SDG slides. Due by July 31, 2019.

**Action WGISS-47-35**: CDA SLT and Ad-hoc team led by Valerie Dixon (Michael Morahan, Andrea Della Vecchia, Chris Lynnes, Cristiano Lopes, Richard Moreno) to look into OGC Testbed 14 results and assess possible use and application within WGISS FDA interoperability activities. Due by July 31, 2019.

**Action WGISS-47-36**: Technology Exploration IG to define a template suitable for gathering inputs from agencies (domain, use case, datasets, labels, dataset accessible accessibility) in the areas of Machine Learning and Artificial Intelligence. Due by August 31, 2019.

## Concluding Remarks

Mirko Albani concluded the meeting thanking NOAA 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 Contac

RS Remote Sensing

SEO Systems Engineering Office

SDCG Space Data Coordination Group

SIT Strategic Implementation Team

SLT System Level Team

SWG Standards Working Group.

TEP Thematic Exploitation Platform

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