

**Notes from the 53rd Meeting of the**

**CEOS Working Group on Information Systems and Services (WGISS)**

**March 22-24, 2022**

**Virtual Meeting**

Table of Contents

[1 WGISS Plenary Session, Part I 5](#_Toc100912981)

[1.1 WGISS Chair Report 5](#_Toc100912982)

[1.2 CEOS Executive Officer (CEO) Report 5](#_Toc100912983)

[2 Data Discovery and Access 6](#_Toc100912984)

[2.1 Introduction 6](#_Toc100912985)

[2.2 General Reports 6](#_Toc100912986)

[2.2.1 IDN, CWIC 6](#_Toc100912987)

[2.2.2 FedEO 6](#_Toc100912988)

[2.3 Data Storage and Format on Cloud 6](#_Toc100912989)

[2.4 STAC for Federated Search 7](#_Toc100912990)

[2.5 GEOSS Platform – Data Brokering 7](#_Toc100912991)

[2.6 Service Metadata and Discovery OpenSearch Best Practices 7](#_Toc100912992)

[3 Data INTEROPERABILITY and USE 8](#_Toc100912993)

[3.1 Explore ARD with the Data Cube for Atmospheric Composition 8](#_Toc100912994)

[3.2 Internet of Things: Monitoring High Quality Geographical Products for Environment and Sustainability 8](#_Toc100912995)

[4 Data Preservation and Stewardship 9](#_Toc100912996)

[4.1 Object Storage Benchmark Report 9](#_Toc100912997)

[4.2 WGISS Data Management and Stewardship Maturity Matrix: Application Use Case 9](#_Toc100912998)

[4.3 Session on Data Integrity and Authenticity on Cloud 9](#_Toc100912999)

[4.3.1 ESA Approach 9](#_Toc100913000)

[4.3.2 NASA Approach 10](#_Toc100913001)

[4.3.3 JAXA Approach 10](#_Toc100913002)

[4.3.4 USGS Approach 10](#_Toc100913003)

[4.3.5 STFC/CEDA Approach 10](#_Toc100913004)

[4.4 CEOS Common Online Dictionary 10](#_Toc100913005)

[4.5 NRSCC’s Methodology for Data Publishing and Repository 11](#_Toc100913006)

[4.6 Topics/Sessions for Future WGISS Meetings 11](#_Toc100913007)

[5 Technology Exploration 12](#_Toc100913008)

[5.1 CEOS Jupyter Notebooks Best Practice and Exemplars 12](#_Toc100913009)

[5.2 Jupyter Notebooks Approach at EUMETSAT 12](#_Toc100913010)

[5.3 Jupyter Notebooks Day and Survey 12](#_Toc100913011)

[5.4 WGCapD and Support for Webinar Plus 12](#_Toc100913012)

[5.5 Topics/Sessions for Future WGISS Meetings 13](#_Toc100913013)

[6 Agency Reports 14](#_Toc100913014)

[6.1 NASA 14](#_Toc100913015)

[6.2 USGS 14](#_Toc100913016)

[6.3 JAXA 14](#_Toc100913017)

[6.4 ISRO 15](#_Toc100913018)

[6.5 NOAA 15](#_Toc100913019)

[7 WGISS Cooperation with other Working Groups 16](#_Toc100913020)

[7.1 Updates from GEO 16](#_Toc100913021)

[7.2 Progress on ISO 19124-1 16](#_Toc100913022)

[7.3 WGDisasters 17](#_Toc100913023)

[8 WGISS Plenary Session, Part 2 18](#_Toc100913024)

[8.1 Future Meetings 18](#_Toc100913025)

[8.2 WGISS Summary and Discussion 18](#_Toc100913026)

[8.3 Review of WGISS Actions 19](#_Toc100913027)

[8.4 Concluding Discussion and Remarks 20](#_Toc100913028)

[9 Glossary of Acronyms 21](#_Toc100913029)

List of Participants

Agencia Espacial Mexicana Adrian Guzman, Julio Castillo

CEOS Marie-Claire Greening

CNES Richard Moreno

CONAE Homero Lozza, Laura Frulla

CSIRO Jonathan Hodge

DLR Jonas Eberle, Karin Molch

ESA Mirko Albani, Puvan Alikana, Paolo Castracane (Rhea), Yves Coene (Spacebel) Sergio Folco, Damiano Guerrucci, Daniele Iozzino (Rhea), Iolanda Maggio (Rhea), Filippo Marchesi, Christophe Noel (Spacebel), Ivan Petiteville, Roberto Roncella (CNR-IIA), Paulo Sacramento, Giuseppe Troina, Joost van Bemmelen

EUMETSAT Federico Fierli, Julia Wagemann

European Commission Fabienne Jacq

GEO Paola de Salvo

GISTDA Pisut Nakmuenwai

ISRO Nitant Dube, Ankitha E, Uzair Mujeeb, Shashikant Sharma, Sai Kalpana Tanguturu, Utkarsh T

JAXA Makoto Natsuisaka (WGISS Vice-chair), Yosuke Ikehata, Michelle Piepgrass (WGISS Secretary)

NASA Andrew Mitchell, David Borges (SEO), Lauren Childs-Gleason, Liping Di, Diane Davies, David Green, Brian Killough, Dawn Lowe, Manil Maskey, Michael Morahan, Doug Newman, Hampapuram Ramapriyan, Kenton Ross, Jennifer Wei, Min Wong

NOAA Kenneth Casey, Prasanjit Dash, Nancy Ritchey

NRSCC Chuang Liu, Yufu Liu, Yuqi Bai

AIR, CAS Xuesong Li, Ziyang Li

Polish Space Agency Oskar Zdunek

UKSA Esther Conway, Katie Awty-Carroll, Sam Pepler, Ag Stephens, Christopher Torbert

UNOOSA Jorge del Rio Vera

USGS Tom Sohre, Christopher Barnes, Kevin Costinett, Ryan Longhenry, Tim Mentele

**WGISS 53 Summary**

Plenary 1

* CEOS Chair CNES initiative “Paths to Sustainability: from strategy to practical measures” was introduced. The initiative defined three priorities in response to the Paris Climate Agreement, the Sendai Framework for Disaster Risk Reduction, and the 2030 Agenda for Sustainable Development.
* CEOS Work Plan 2022-2024 to be approved at SIT-37 and associated WGISS part (Chapter 3.9) were reported. Near-term plan, joint collaborations and information exchanges with other groups were summarized.

Data Discovery and Access

* CEOS/IDN collaborations, IDN upgrades, and IDN metrics were reported. FedEO migration to CEOS domain and DIF-10 export status were also reported. Both systems exchange registered metadata on daily basis.
* ESA GSTP project aiming to provide a cloud-native access to hyperspectral data was introduced. The formats trade-off resulted in the selection of Zarr format and GeoZarr format was proposed. Demonstrations on Jupyter Notebooks, OpenLayers, FedEO + STAC catalogue, and advanced applications were also made.
* A parallel STAC implementation in CWIC was proposed. STAC is a great candidate for a federated search solution and has a lot of traction in web- and cloud-based user communities.
* The current status on GEOSS Platform was reported. The platform consists of four main components, GEOSS Portal, GEO DAB, GEOSS Status Checker, and GEOSS Yellow Pages. The platform is brokering CEOS catalogues from IDN, CWIC and FedEO.
* The current status on service metadata and discovery best practice was reported. Based on the draft version, the contents were introduced. The comments and ideas were solicited.

Data Interoperability and Use

* Explore ARD with the Data Cube for Atmospheric Composition developed by EUMETSAT was introduced. The system was one of the perfect implementations enabling interoperability and use in the cloud.
* Internet of Things: Monitoring High Quality Geographical Products for Environment and Sustainability was introduced.

Data Preservation and Stewardship

* Object Storage Benchmark in ESA was reported. The benchmarks showed benefit of usability (Scalability, Manageability, etc.) and costs against small lack of performance.
* WGISS Data Management and Stewardship Maturity Matrix Use Case was reported. Main concept, use case for ENVISAT, etc. were introduced.
* Considering increase of data replication on cloud, a session on “Data Integrity and Authenticity on Cloud” was organized. To keep data integrity, data authenticity, provenance, traced modification, PID in the metadata, Hash Code in the metadata, watermark, KSI block chain were discussed in the previous meetings. Member agencies reported their approaches which include as the possibility to use Filename Lookup, Hash Lookup, Provable Data, Possession Signed Hashes, Checksum, and Non-Fungible Token (NFT). The necessity of further exchange of information on ongoing research and prototyping activities was pointed out.
* Common Online Dictionary Initiative, which is a joint activity with WGISS, WGCV, LSI-VC, and ISO/OGC, was introduced by the CEOS Common Terminology Group and called for more WGISS participation. Participant’s vocabularies and glossaries were provided by WGISS as input to the group.
* NRSCC’s Methodology for Data Publishing and Repository was reported.
* Topics for future DSIG sessions were agreed: Archiving technologies and approaches, preservation of space data associated content (information, software, tools), recovery of heritage data not accessible to users, data quality assessment and indicators (in cooperation with WGCV).

Technology Exploration

* Concept of CEOS Jupyter Notebooks Best Practice and exemplars was introduced.
* Jupyter Notebooks approach at EUMETSAT was introduced.
* “Jupyter Notebooks Day” at the end of June, a kind of hands-on event including CapD training, was proposed.
* Expectation for further webinars was expressed by the WGCapD representative.
* A co-lead of Technology Exploration Interest Group solicited future topics.

WGISS Cooperation with other Working Groups

* Latest information on GEO Data Working Group was updated. DWG consist of four subgroups; Data Management Principles, in-situ Data, Law and Policy, and Data Ethics. The status on developments of GEO Knowledge Hub and GEOSS Plus were shared.
* Progress on ISO 19124-1 and the New Joint ISO-OGC Standard Initiative on Analysis Ready Data was reported. The latter initiative is aiming at standardizing CEOS ARD. WGISS activities will be input to the ISO TC 211 54th Plenary will be held in Vienna, Austria on May 12-13, 2022.
* The activities of WGDisasters were introduced. WGDisasters consist of 9 subgroups, Landslide Demonstrator, Volcano Demonstrator, Seismic Hazards Demonstrator, GeoHazards Lab, GEO-DARMA, Flood Pilot, GSNL, Recovery Observatory Demonstrator, and Wildfire Pilot.

Plenary 2

* Member agencies reported the current activities.
* WGISS-54 will be held in-person in Tokyo in the week of Oct. 3rd, 2022. The logistics will be delivered later.
* Special Sessions in WGISS-54:

Joint symposium with WGCV

Data Integrity and Authenticity on Cloud (TBD)

AI/ML (TBD)

EAIL (TBD)

Conclusions

* WGISS-53 was held virtually on 22, 23, and 24 March 2022.
* Forty-four people from WGISS, CEO, CEO, WGCapD, WGDisasters, EUMETSAT, etc. participated in the meeting.
* Key issues, which are necessary for “data interoperability and co-use”, were identified. “Cloud” is the inevitable element and associating technologies such as data archiving in the cloud (including ARD and data cube), data analysis in place in the cloud (including AI/ML), data integrity and authenticity, knowledge share with Jupyter Notebooks should be continuously pursued.
* Further discussions in Living on Planet Symposium would be productive in addition to the WGISS regular meeting.

# WGISS Plenary Session, Part I

## [WGISS Chair Report](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_10.05_WGISS_Chair_Report.pptx)

Makoto Natsuisaka (JAXA) chaired the WGISS-53 meeting, and gave the following report.

* CEOS Chair Transition - Goals of the new leadership of CEOS “Paths to Sustainability: from strategy to practical measures Version 1.0 – 25 October 2021”.
* CEOS Work Plan 2022-2024 is to be approved in CEOS SIT-37.
* WGISS contributes to the EO communities including CEOS by providing WGISS Connected Data Assets (IDN, CWIC and FedEO), Carbon and Water Portals, etc.
* WGISS developed and maintains guidelines in line with EO data life cycles.

Data Stewardship Best Practice

Persistent Identifier Best Practice

OpenSearch Best Practice etc.

* WGISS organized seven WGISS Webinars.
* Joint collaborations with other CEOS entities, EAIL initiative, Jupyter Notebook initiative, Common Maturity Metrics, Common On-line Dictionary, ARD Oversight Group, etc. are on-going.
* Information exchange with EUMETSAT and WGDisasters were carried out. One with CGMS (Working Group-1V) is planned.

## [CEOS Executive Officer (CEO) Report](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_10.20_CEO%20Report.pptx)

Marie-Claire Greening (CEOS) presented CEOS mission and objectives, long-term priorities, CEOS governance documents, and the CEOS Work Plan. She also discussed the 2022 CEOS Chair (CNES) priorities. She continued with the WGISS mission and objectives:

* To enable Earth observation data and information to be more accessible and usable to both data providers and data users world-wide through international coordination.
* To foster easier exchange of Earth observation and related data and information to meet the requirements of users and data providers.
* To foster the development of best practices and encourage the development of interoperable services that exploit space-borne Earth observation data.
* To enhance the complementarity, interoperability and standardization of Earth observation data and information management and services with other types of geospatial data such as *in situ* data.

Following are the current WGISS deliverables:



Makoto asked about the coordination meeting between CEOS and GEO in February. Marie-Claire replied that a summary paper was distributed, adding that Brian Killough gave an excellent presentation.

# Data Discovery and Access

## Introduction

Damiano Guerrucci (ESA) introduced the session and described the upcoming presentations.

## General Reports

### [IDN, CWIC](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_10.40_IDN_Report.pdf)

Michael Morahan (NASA) reported on the International Directory Network (IDN), specifically:

* CEOS/IDN Collaborations
	+ ESA/FedEO Updates: Added FedEO CMR Provider ID for FedEO consortium
	+ NOAA TPIO and GCMD keywords: Added new Ocean keywords
	+ NOAA GHRSST Datasets: Added CMR NOAA\_NCEI GHRSST Provider ID for CWIC consortium
	+ ChinaGEOSS, NRSCC, and CCMEO Datasets: Added NRSCC and New CMMEO CMR Provider IDs CWIC consortium
* IDN Upgrades
	+ IDN Search Portal: granule discovery and download
	+ CMR Tagging replaced by Consortiums: to identify CWIC, FedEO, CEOS, and GEOSS.
	+ CMR/DraftMMT Progressive Update Validation
	+ Platform Facets: Added New Platform Top-level “BASIS” group.
	+ UMM-C OrbitParameter and MetadataSpecification: Added orbit Unit and Footprint sub-fields to the schema.
	+ IDN Search Portal Cloud Datasets refinement
* IDN Metric
	+ IDN Homepage Usage
	+ IDN Search Portal
	+ Draft Metadata Management Tool (MMT)

### [FedEO](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_10.50_FedEO_Report.pptx)

Yves Coene (Spacebel/ESA) gave a report on progress and status of the Federated Earth Observation (FedEO):

* FedEO in CEOS domain
	+ Client and Server available under CEOS domain
	+ WCDA Client Partner Guide update: 26/11/2021
* FedEO current metrics and DIF-10 export status
	+ Evolution: 107 to 112 million granules
	+ Collections: 1774
	+ DIF10 export improvements (DOI, Use Constraints)
	+ Completed: ESA CCI collections (180) ready for IDN
	+ Ongoing: JAXA G-PORTAL collections (1131) and granules (19 million)

Damiano commented how well services can be augmented.

## [Data Storage and Format on Cloud](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_11.00_Data%20Storage%20and%20Format%20on%20Cloud.pptx)

Christophe Noel (Spacebel/ESA) gave a presentation on data storage and format on cloud. He summarized as follows:

Hyperspectral Data Store and Access is an ESA GSTP project aiming to provide a cloud-native access to multidimensional data.

* Formats trade-off resulted in the selection of Zarr format.
* Implementation of Zarr-based Data Store prototype (OVH + AWS).
* Converted PRISMA, Sentinel-2, ENVISAT, Copernicus products.
* Proposed GeoZarr format supporting geospatial data and visuals.
* Demonstrators including Jupyter Notebooks, OpenLayers, FedEO + STAC catalogue, advanced applications.

Results:

* Outstanding performances based on independent portions (chunks) that are parallel-friendly, big-data friendly, and can be optimised (e.g., time).
* Affordable and scalable object storage (store is cloud-agnostic).
* Multidimensionality is an asset, facilitates time series/spectrum analysis.
* Serverless: efficient and facilitate applications development with many libraries.

## [STAC for Federated Search](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_11.15_Using%20STAC%20for%20Federated%20Search.pptx)

Doug Newman (NASA) gave a presentation on STAC for Federated Search.

* STAC is a great candidate for a federated search solution
* It has a lot of traction in the community beyond the use case of federated search
* OpenSearch development is somewhat static
* STAC can be used for federated search using the same CWIC architecture
* Proposal: develop a parallel STAC implementation in CWIC

Damiano included a link to the [STAC Primer](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_11.15_STAC%20Primer.pdf).

Damiano added that extending the interface is needed, and Yves noted that everyone is using the same search parameters. Damiano invited others to collaborate; contact Damiano to be involved in the discussion.

Y. Coene asked regarding slide 13: How is it proposed to implement steps 1a, 1b (collection search) in the green interface called "STAC API" taking into account that STAC does not define a (STAC) Collection level search but only searching for "items" (returning a GeoJSON FeatureCollection, also a STAC collection is defined in JSON, not in GeoJSON)? Do you propose to have an OGC API interface instead possibly aligned with OGC API Common and/or OGC API Records with a search supported at the /collections endpoint and not only at the /items endpoint? Christophe replied that STAC has collection level searching. [Example](https://cmr.earthdata.nasa.gov/stac/GES_DISC/collections?datetime=2009-04-22T00:00:00.000Z/2009-04-23T00:00:00.000Z)

Y. Coene asked regarding slide 12: are you proposing that data providers support the "collections" queryable (defined for the /search STAC endpoint) and another search interface at their /collections/{id}/items endpoint (thus two different endpoints for performing granule searches)? Or are you proposing that the /search endpoint is provided at the global level (at NASA EOSDIS (IDN) CMR)? Otherwise, the "collections" that can appear in a search (slide 12) are limited to those provided by the same data provider? Answer: global.

Y. Coene (Spacebel/ESA): slide 15: does the "collection-level search" take any HTTP query parameters allowing to search (e.g., equivalent to {searchTerms}) or is this a flat list of "collections" as in the current STAC/OGC API Features specification? Answer: Free text is not currently supported by the base STAC API specification, or any extensions.

## [GEOSS Platform – Data Brokering](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_11.45_GEOSS%20Platform%20Data%20Brokering.pptx)

Roberto Roncella (CNR/ESA) gave a presentation on data brokering on the GEOSS Platform. He explained the focus of the GEOSS Platform and its main components. Roberto elaborated on the views and filters, and the supported standards. He concluded saying that there are three brokered catalogues using OpenSearch:

* CEOS International Directory Network (IDN)
* CEOS WGISS Integrated Catalog (CWIC)
* Federated EO Gateway (FedEO) – CEOS

Possible enhancements are to avoid empty collections, use custom filters on granules search, and custom views focused on specific use-cases.

## [Service Metadata and Discovery OpenSearch Best Practices](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.00_Service_Discovery_Best_Practices.pptx)

Yves Coene (Spacebel/ESA) gave a presentation on service metadata a discovery best practice:

* Current “WGISS Connected Data Assets” is made possible thanks to cooperation within WGISS/SLT:
	+ Prepared “CEOS OpenSearch Best Practice”
	+ Best Practice implemented by partners and integrated with IDN
* Cooperation within WGISS/SLT proposed at WGISS-52 for future “CEOS Best Practice for Service Metadata and Discovery”
* Current status: the initial draft (November 2021 was updated March 2022 and is available resolving review comments from SLT group on initial draft. The current document is presented at this meeting.

Christophe Noel added that Data Storage and Format on Cloud additional presentations and demonstration videos can be found [here](https://github.com/christophenoel/geozarr-spec).

# Data INTEROPERABILITY and USE

## [Explore ARD with the Data Cube for Atmospheric Composition](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.00_Service_Discovery_Best_Practices.pptx)

Federico Fierli and Julia Wagemann (EUMETSAT) gave a presentation on exploring ARD with a data cube for atmospheric composition. They discussed the concept of ARD, and its practical application to data cubes.

EUMETSAT’s Atmospheric Composition DC (ACDC) includes three data services, three instruments on four different satellites, and six atmospheric variables. The ACDC includes 18 different data products in three different data formats.

Users can explore the Atmospheric Composition Data Cube (ACDC) with the [Jupyter Training platform](https://epct.ltpy.adamplatform.eu). This is a work in progress, and input from WGISS is welcome.

## [Internet of Things: Monitoring High Quality Geographical Products for Environment and Sustainability](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_12.45_Internet%20of%20Things.pptx)

Liu Chuang (NRSCC) gave a presentation on Monitoring High Quality Geographical Products for Environment and Sustainability. She discussed the 2021-2030 Joint Decadal Program on Geographical Indications Environment and Sustainability (GIES). A research gap exists whereby people focus on geographical products themselves but pay limited attention on the habitat and environment. To fill this gap, the team will focus on the original geographical environment, adopting a multi-stakeholder joint action.

Makoto asked if the service is similar to a knowledge hub. Liu Chuang said yes, with four catalogs: physical geography, product data, associated economic data, and traditional culture.

# Data Preservation and Stewardship

## [Object Storage Benchmark Report](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_12.45_Internet%20of%20Things.pptx)

Daniele Iozzino (Rhea/ESA) gave a report on object storage benchmark, a data storage architecture for handling large amounts of unstructured data. Objects contains the data and the metadata. Access to data is performed using a data transfer system over HTTP. Custom metadata can be used to tailor the technology to specific use cases. Extensive studies and benchmarks are taking place to assess whether Object Storage is suitable for Long Term Data Preservation of ESA EO Data holdings. It has been discovered that the drawbacks are the same with the already present Tape Libraries. Benchmark shows the benefit of usability and costs against small lack of performance. Object Storage is being implemented in the “Heritage Software Hub” as a proof of concept.

Doug commented that separating the data from the metadata could improve latency; storing the metadata 'with the object' is an interesting proposition but the limit on the size of metadata makes it unsuitable for the current needs. The metadata is still limited to 2K.

## [WGISS Data Management and Stewardship Maturity Matrix: Application Use Case](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_12.45_Internet%20of%20Things.pptx)

Iolanda Maggio (Rhea/ESA) discussed an application use case of the WGISS Data Management and Stewardship Maturity Matrix (DMSMM). She began with the main concept and continued with the application of the DMSMM step by step, using the MM to measure status of processes and also to define future objectives: For each component go through the checklist, ultimately compiling the score. Using the preservation area as an example, there are a total of 15 steps for the three levels (L0, L1, L2, L3). When each level is complete, a score can be computed. Next, she showed the final view of the DMSMM for this example.

A use case for ENVISAT MIPAS L2 Data:

* First Step: Analyse the relevant Goals and Objectives to be reached and visualise the final measurement on the table.
* Intermediate Steps: Periodically the operator can verify the status of the Data Management and Stewardship processes highlighting the missing steps to reach the specific goals.
* Last Step: Closure of the process and creation of the final score.

Makoto asked what the expected progress by WGISS-54 would be? Iolanda replied that they plan to simulate use of MM with a real example: AVHRR. For the quality aspects they will work with WGCV, interacting with them to ensure full alignment.

Ken asked how long it takes to complete the analysis? NOAA has found it took a long time to do a similar assessment. Iolanda replied that it depends on the mission: For Envisat with different sensors it can take some time and a team of experts is needed to see if objectives have been reached; also, it is an iterative process.

Iolanda added that this new version of the DMSMM has 12 components and a complete landscape, including the quality.

Mirko added that this will be reviewed with WGCV at WGISS-54, and he will ask NOAA to review the assessment.

## [Session on Data Integrity and Authenticity on Cloud](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_10.30_Data%20integrity%20and%20Authenticity%20on%20Cloud%20SESSION.pptx)

Mirko Albani (ESA) introduced the session on data integrity and authenticity on cloud highlighting the need to understand how to handle a data replica on the cloud and how to address the integrity and authenticity. At WGISS-52 a comparison of the different approaches was shown; this session will review different agency experiences.

### [ESA Approach](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_10.40_ESA%20Traceability%20Service%20for%20EO%20Data.pptx)

Mirko Albani (ESA) discussed the ESA approach to handling data integrity and authenticity on the cloud. He provided the following summary.

* Prototyping activity is implemented under ESA contract.
* Scope: to store signed traces of EO data to ensure their integrity and to guarantee their origin to users.
* TRACE: Record with signed information on products, (e.g., Name, ID, size) and data content HASH: Owner signature guarantees product authenticity, trust in the owner is assured by a certificate
* HASH: Computation that maps data to a bit string ‘hash’: Used to uniquely identify files.
* Digital signature used for verifying authenticity of digital messages, based on pair of public and private keys: Data producer will sign traces, end-users will be able to verify data integrity and origin
* Certification Authority (CA): delivers digital certificates to prove the ownership of a public key

Prototype activities have been completed and have been used in the transfer of Sentinels Level-0 data between two archives. A new archive could check authenticity and integrity of the data received from the old archive.

Operational service is under consideration/preparation and might be used between a production and archiving entity (production checking data received from archive). The production centre will then generate a new trace which will be embedded in the product and allow another entity to verify authenticity/integrity of the product itself through getting a certificate confirming that the new trace is authentic. The new entity will also get information about the product history (traceability back to initial archive/owner, provenance information maintained). Final concept for end user utilization is to be refined.

Ken asked who provides the certificate authority? Mirko replied that it is the guidance system, and the final needs to be decided.

### [NASA Approach](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_10.50_Data%20Integrity%20NASA.pdf)

Manil Maskey (NASA) discussed the NASA Earth Observation data traceability study. Its scope is to:

* Investigate problems that arise due to duplication and mirroring from a data user’s perspective
* Explore a range of technical solutions to address this data traceability problem
* Focus only on addressing a set of specified use cases

Study facets are data integrity checks (authenticity, data deletion, data modification), non-repudiation, ease of implementation, and performance.

Approaches are filename lookup, data upload and verify, Hash lookup, provable data possession, and signed hashes.

Key takeaways are that no approach addresses all the identified needs, and both hash lookup and signed hashes are viable practical approaches to address some of the issues. Hash lookup is suitable if the data user wants to process the full archive of a dataset for any kind of large-scale processing.

### [JAXA Approach](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_11.00_Data%20Integrity%20and%20Authenticity%20on%20Cloud_JAXA%20Approach.pptx)

Makoto Natsuisaka (JAXA) discussed the JAXA approach, noting that there are a few JAXA replicas on cloud so far. PIDs for standard products will be soon introduced, but those will not be included in the metadata, nor will hash code be in the metadata. PIDs in the metadata could be considered as a realistic way for JAXA, since they can give integrity, data authenticity, provenance and traced modifications.

### [USGS Approach](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_11.05_Integrity%20USGS.pdf)

Tom Sohre (USGS) discussed the approach at USGS, noting that Landsat data is provided and replicated in multiple locations. Ensuring data integrity and authenticity of Landsat data in each location is of utmost importance to both Landsat and customers. USGS enabled Landsat in the Cloud in 2020 and utilizes a hybrid cloud approach for data processing, storage and access; checksums are currently utilized by internal Landsat processes to ensure data integrity.

Landsat users that are retrieving data from the USGS cloud archive are provided a checksum value; USGS will be investigating strategies and technologies to specify and verify replica archive copies as well as individual products for data security, data integrity, digital provenance, and authenticity.

Daniele commented that ESA is investigating/testing blockchain and can share the outcomes.

Sam Pepler and Ken commented that much of this is a technology evolution, a process improvement. MD5 is considered weak and hackable. A discussion on these technologies could be suitable topics for WGISS-54.

Makoto Natsuisaka asked if they used PIDs, Hash, watermark and blockchain for CSC Level-0? Mirko replied that they use DOIs (PIDs) for all data and generate a landing page providing all information on the data itself. This is most likely not enough so two other options were explored to be added on top. The prototyping consisted of Blockchain and Traces and Digital Signatures for the Sentinel Level-0 archive and it proved to work well. Implementing this as an operational service is being considered.

Adrian Guzman (AEM) noted that they did a small prototype for Data Provenance which could be used for NFT´s with differentiated databases (both Academic and Commercial), in order to secure Source Authenticity (like Harvard Dataverse) and Ownership in Commercial. He offered to present this to WGISS next year.

Ken noted that NOAA is interested in learning results from blockchain testing; this would also be a good topic for WGISS-54. There is broad interest to continue the topic of using technologies for data integrity and authenticity.

### [STFC/CEDA Approach](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_11.15_Data_Integrity_Authenticity_%20STFC_CEDA.pptx)

Sam Pepler (STFC/CEDA) gave a summary of current processes and actions promoting integrity and authenticity at STFC/CEDA. He commented that they are principally concerned with integrity, and not doing so well on the authenticity; researchers are unlikely to take up any complex protocols to ensure integrity or authenticity.

## [CEOS Common Online Dictionary](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_11.50_CEOS-Common-Online-Dictionary.pptx)

Katrin Molch (DLR) gave a presentation on the CEOS Common Online Dictionary Initiative. She gave the background and status of activities. The team are proposing two lists of terms and are starting to assemble candidate lists. Proposals are highly welcome - with or without a definition, and can be sent to: katrin.molch@dlr.de

* BASE terms illustrate the fundamental concepts, they should use understandable, unambiguous language and be built only on other base terms.
* HIGH IMPACT terms refer to expressions used within a domain to define more complex concepts or make important distinctions.

The team have reached out to ISO and OGC to align approach and learn about most recent standard developments.

Katrin listed ongoing discussion points:

* Hierarchy of terms
* Classification or scales
* Words that need to keep multiple meanings
* Processing levels re-consideration

Iolanda reminded about the issue with the term ‘level’ in terms of maturity level vs. processing level.

Makoto asked how many are involved in this initiative; Katrin said that currently it is seven people. They coordinate using chat, and have a specific working day of the week for this; they also have regular meetings.

The following inputs were given as sources:

* Sam Pepler: NERC vocabularies:  <http://vocab.nerc.ac.uk/>

 Governance: a tool to coordinate updates: https://cfconventions.org/standard-names.html

 http://cfeditor.ceda.ac.uk/proposals/1?status=active&namefilter=&proposerfilter=&descfilter=&filter+and+display=filter

* Yves Coene: <https://thesauri.eo.esa.int/thesaurus/en/> and <https://thesauri.spacebel.be/en/> The ESA thesauri are using Fuseki (SPARQL) and the SKOSMOS tool as Web client.
* Jonathan Hodge: Ocean Data Interoperability Platform (www.odip.eu) investigated a lot of these topics in the ocean domain
* Michael Morahan: the GCMD Keyword: https://wiki.earthdata.nasa.gov/display/CMR/GCMD+Keyword+Access

## [NRSCC’s Methodology for Data Publishing and Repository](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_12.05_Data%20Publishing%20and%20Reporsitory.pptx)

Liu Chuang (NRSCC) discussed NRSCC’s Data Publishing and Repository, supporting about 5000 academic journals, research data from 20 national data centers, and the three journals coordinating program. This program results in global networking, and reducing the digital divide.

Mirko noted that this is an interesting and important initiative. WGISS is proposing a session to identify datasets not accessible; it would help to have any input on this.

## [Topics/Sessions for Future WGISS Meetings](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/2.Wednesday/2022.03.23_12.15_Topics-Sessions%20for%20Future%20WGISS%20Meetings.pptx)

Iolanda Maggio (Rhea/ESA) led a discussion on possible topics for future WGISS meetings.

1. Data authenticity and integrity: discuss the specific guideline in the LTDP Preservation Guideline Best Practices: Ensure that the content of the archived data and associated information remains unchanged and, if changes are made, that these are documented, preserved and made available as well (provenance information). Also continue discussion on ESA KSI Blockchain pilot, and on small prototype on Source Authenticity (like Harvard Dataverse) and Ownership in Commercial.
2. Archive Holdings and Technology Session, to discuss current technology and infrastructure; archive volume detailed per mission; archiving operations concept, flows and processes; data format/packaging for long term archive; management and archiving of the relevant associated information; archive technology and media evolution and future trends; challenges and needs.
3. Historical Data Recovery: Climate applications are requiring more and more to extend critical long-term science observation back in time through recovery and use of historical data sets. Objective is to recover and possibly harmonise the identified data sets not accessible to users (e.g., because on old media, not kept online, etc.). The proposed topic: Identification of heritage datasets to be recovered for future joint projects with the following outline: location and status, format, volume, etc., and prioritization based on uniqueness, applications, needs, impact, etc.
4. Approaches, Processes and Tools for Information Content Preservation. Long-term accessibility and exploitability of Earth Science data requires that not only sensed data, but also technical content and associated information needs to be properly preserved and made accessible. The following topic with relevant outline is proposed: Long Term Preservation Platforms (Open source and/or licensed); Preservation processes; Preservation Metadata and Formats; Provenance and linked network; Software and Tools Long term preservation

Mirko noted that the order given above is the level of priority. Suggestions are welcome.

# Technology Exploration

Yusuke Ikehata (JAXA) introduced the session.

## [CEOS Jupyter Notebooks Best Practice and Exemplars](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.15_JupyterExemplar_BPs.pptx)

Esther Conway (UKSA) began her presentation on CEOS Jupyter Notebooks best practice document, explaining who and why would create a BP; she listed the proposed outline. The BP will encourage the development of good workflow and structure within notebooks along with quality documentation.  She listed technical dependencies and virtual environments.

Esther noted that currently there is no guidance on how a Jupyter Notebook should cite input data, explain access requirements, and location and structure of data. Other guidance areas are the incorporation of Jupyter Notebooks with data cubes, version control, preservation and archival, publishing software, and getting a DOI and license.

Esther concluded saying that a Binder repository can be built by a BinderHub, which will generate a link that you can share with others, allowing them to interact with the content in your repository.

## [Jupyter Notebooks Approach at EUMETSAT](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.25_Jupyter_notebooks_at_EUMETSAT.pdf)

Julia Wagemann (EUMETSAT) discussed the approach for Jupyter Notebooks at EUMETSAT. The motivation is that Earth Observation training @EUMETSAT has to cater for different levels of data, thematic and programming literacy.

EUMETSAT has developed a series of Jupyter notebook trainings on different application areas, e.g., atmospheric composition. Learning Tool for Python (LTPy) on atmospheric composition consists of over 70 notebooks related to data access, analysis, data discovery, case studies, exercises and thematic modules. TrainHub is EUMETSAT’s Jupyter Notebook portal and Jupyterbook covers ‘Dust Aerosol detection, monitoring and forecasting’.

Using computational notebooks for training/education brings in a set of additional requirements that require the integration of didactical concepts, instructional design patterns and best practices for coding.

Julia identified the need for defining, sharing and implementing quality standards and best practices on how to make notebooks effective, reproducible and educational:

* Systematic use of recommended Python libraries
* Modularisation of content, e.g., outsource functions
* Use of instructional design patterns
* Follow a naming nomenclature to order notebooks in a specific sequence

## [Jupyter Notebooks Day and Survey](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.35_JupyterNotebookDay.pptx)

Esther Conway (UKSA) announced the “CEOS Jupyter Notebooks for Earth Observation: Best Practice and Capacity Development (OpenSource Science, toolboxes and Jupyter technologies in EO)” at the ESA Living Planet Symposium in May (<https://lps22.esa.int/frontend/index.php>). She continued with results of a questionnaire on training, a Hackathon, and the development of a Best Practice.

Esther concluded with the proposed agenda for the planned Jupyter Notebooks Day.

## [WGCapD and Support for Webinar Plus](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.45_WGCapD%20and%20Support%20for%20Webinar%20Plus.pptx)

Kenton Ross and Lauren Childs (NASA) gave a presentation on WGCapD support for Jupyter Notebook Webinar Plus. Kenton began with a list of the relevant open CEOS Work Plan deliverables:

* [CB-22-13](http://deliverables.ceos.org/task_manager/deliverables/715/): Jupyter Notebook Foundations Webinar
* [CB-22-10](http://deliverables.ceos.org/task_manager/deliverables/712/): Sentinel selected applications: practical training with Jupyter Notebooks on the ESA EO Platform
* [CB-22-07](http://deliverables.ceos.org/task_manager/deliverables/709/): Open-Source Science Outreach Plan and Training

Kenton suggested the following potential WGISS-WGCapD collaboration points:

* Relevancy Ranking of Data Search Results
* Data Cubes for Large Scale Data Analytics
* The Burgeoning Role of Python for EO Data Analysis
* Explore Capacity Development with the EAIL
* Joint Support for CEOS ARD
* Jupyter Notebook Best Practice Documentation/Training

Potential Joint WGISS-WGCapD webinars are:

* Open Geospatial Consortium (OGC) Coverage Standards Suite: Introduction & Overview
* Agile Development and Scaled Agile Framework
* Future Data Access and Analysis Architecture Initiative Webinar (Joint WGISS/WGCapD Activity)

## [Topics/Sessions for Future WGISS Meetings](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.55_Topics_Sessions%20for%20Future%20WGISS%20Meetings.pptx)

Yousuke Ikehata (JAXA) presented the following topics for Technology Exploration sessions at future WGISS meetings:

* Jupyter Notebooks, with a goal to develop a CEOS Jupyter Notebooks Best Practice in collaboration with WGCapD, and Jupyter Notebooks Day.
* AI/ML/DL, with a goal to develop an AI/ML/DL White Paper, which will consolidate past presentations and updates.
* Federation: a new topic, which needs to be defined and discussed.

# Agency Reports

## [NASA](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/1.Tuesday/2022.03.22_12.55_Topics_Sessions%20for%20Future%20WGISS%20Meetings.pptx)

Andy Mitchell (NASA) gave a report on recent developments at NASA. He identified the following key points:

* Delivered nearly billion products to over 4 million users in FY20 ((Oct. 1, 2020 to Sept. 30, 2021) to every country in the world representing an increase of 100 million products from FY20. 1,502 new and reprocessed datasets were added to the EOSDIS collection in FY21.
* NASA Earth Observing System Data and Information System (EOSDIS) data have been distributed to 195 independent countries
* Provided data stewardship to almost 13,000 unique data sets for which we have minted over 9700 digital object identifiers that enable users to track data sets through publications and documentation.
* By the end of FY21, had archived over 59 Petabytes of Earth Science data at a rate of 53 terabytes per day. This was a 34% increase from FY20. Over 15 petabytes are available in Amazon Web Services (AWS).
* Scored 81 on the American Customer Satisfaction Index (ACSI) survey an increase of two points from last year’s score of 79. This continues the trend of high scores for EOSDIS performance, despite the pandemic, a shift to working from home, and changes in our NASA missions and the research community schedules and plans.
* Undergoing an Open-Sourced Science for Earth System Observatory (ESO) Mission Science Data Processing Study with the goal of: Identify and assess potential architectures that meet the ESO mission science data processing objectives, promote open science principles, enable data system efficiencies, and support earth system science and applications.

## [USGS](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_10.55_USGS%20Agency%20Report.pdf)

Tom Sohre (USGS) gave a report on recent developments at USGS. The following key points were identified:

* Landsat 9 was launched in 2021 and is operational and meeting mission requirements. The U.S. Geological Survey made Landsat 9 data available from the Landsat archive beginning February 10, 2022. As of February 15, 2022, Landsat 9 Collection 2 Level-1, Level-2, and U.S. Analysis Ready Data (ARD) data are available via the commercial cloud. The Landsat 9 Data User Handbook is now available for download from the Landsat Missions website.
* NASA and USGS have set up formal projects to pursue Landsat Next; RFIs and Instrument studies have been initiated. Currently, over 10 million Landsat scenes are available.
* The USGS National Land Imaging (NLI) Program delivers a national and global capability to ensure broad public and scientific availability of observations of the Earth’s land surface.
* A new science product easily separates snow from vegetation in processed Landsat Collection 2 scenes from 1982 to present.
* As of December 1, 2021, the locations and filenames of RSS feeds that alert you to new Calibration Parameter Files (CPF) for Landsat 8 and Landsat 7 data will be changing.
* Since 1972, Landsat satellites have continually acquired data about the Earth’s land surface. On November 23, 2021, the Landsat Archive that stores this vital record added its ten millionth scene.
* Landsat data acquired after December 31, 2021 will not be available in Collection 1.

## [JAXA](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_11.05_JAXA_AgencyReport.pdf)

Yosuke Ikehata (JAXA) gave a report on recent developments at JAXA. The following key points were discussed:

* Missions’ status: Six (complete), six (in operation), and four (to be launched).
* JAXA Ground System for Earth Observation Missions is operational.
* JAXA Portals and Data Provision to Partner Portals
* JAXA Data Dissemination System: “G-Portal”
* Japanese commercial satellite data platform “Tellus”
* Science and Applications through Cooperation with International Partners, and Earth Observation Contributing to Humanities and Social Sciences.
* Cooperation for development of Global Biomass Map, cooperation with Google Earth Engine, and JAXA-NASA-ESA cooperation in response to COVID-19.
* JAXA is gradually addressing to open-source science and has started with modification or improvement of data dissemination system. JAXA has also started cooperation with various partners including service providers to promote open science.
* Conditions of intellectual properties, particularly software and tools, vary each satellite mission and many software and tools are not intended to open to the public. Cost of processing of huge volume satellite data, i.e., ALOS-2, is also a challenge to have them available with open and free condition.
* Partnership is indispensable for enhancement of further scientific researches and downstream applications for the next generation. JAXA’s IT resources and capabilities are limited to address to integration of bigdata and AI in order to have satellite observation incorporated into society; promoting applications for economics and social sciences is indispensable.
* JAXA is interested in NASA’s open-source science initiative as an opportunity to promote scientific researches and applications for next generation.

## [ISRO](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_11.15_Agency%20Report_ISRO.pdf)

Nitant Dube (ISRO) gave a report on recent developments at NOAA. He identified the following key points:

* EOS-6, INSAT-3DS and NISAR are forthcoming Earth Observation missions
* Bhoonidhi: <https://bhoonidhi.nrsc.gov.in>: Regional Data Hub available for users to order and download data. Bhoonidhi Vista is data visualization and analysis tool available with Bhoonidhi for quick visualization of full resolution data
* Bhuvan: <https://bhuvan.nrsc.gov.in>: Following New Applications Released
* Bhuvan Lite: Lite weight data visualization tool, Bhuvan Wiki: Knowledge Sharing Platform, Bhuvan Yuktdhara: Regional Level Planning Tool, Bhuvan: National Hydrology Project
* VEDAS: <https://vedas.sac.gov.in>: Enhancing Geo-Spatial Analysis
* AI based Time series forecast of NDVI, Field level crop monitoring, Monitoring of Reservoirs/Inland water bodies, Web enabled Long-term data analysis
* MOSDAC: <https://mosdac.gov.in>: Ocean-Eye, SCORPIO, Netra and Safe Beach are new applications released on MOSDAC – Alert and Forewarning services, now available for Indian States and South East Asian Countries. – Working on Ocean and Weather Data Cubes – Weather and Ocean Data Analytics
* National Information system for Climate and Environment Studies (NICES) <https://nrsc.gov.in>: Data available for download from Bhuvan – Satellite-retrieved geophysical product inventory (70-products, including derived products) – Products from National Satellites (24), International satellites (15) and Model outputs (9) – Atmospheric Lightning ECV – using ground-based Detection Sensor Network

## [NOAA](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_11.25_NOAA%20Agency%20Report.pptx)

Ken Casey (NOAA) gave a report on recent developments at NOAA. He identified the following key points:

* As part of changes across the US Federal Government, NOAA is substantially revising its data governance
* Update on NESDIS Common Cloud Framework (NCCF)

NCCF continues progress

NESDIS Cloud Archive Program integrated into NCCF and on track for late 2022 initial operating capability

New knowledge graph supports the emerging Cloud Archive

* Updates from NCEI

OneStop NOAA Data Catalog update to V3 in May-June time frame

Collection Metadata Editing Tool (CoMET) continues progress, integrating Data Stewardship Maturity tools

# WGISS Cooperation with other Working Groups

## [Updates from GEO](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_10.00_Updates%20from%20GEO.pptx)

Paola De Salvo (GEO) gave a presentation updating WGISS on pertinent GEO activities.

* Data Working Group/Data Sharing:

- Data Management Principles:

Review of Data Sharing and Data Management Principles Documents.

Finalize the DMPs Crosswalk (GEO -FAIR-TRUST-CARE).

DMPs Success stories – Living document.

Towards a GEO DMPs self-assessment tool.

Establishment of GEO Educational Series – launch at the GEO Symposium.

Data Analysis Survey and Engagement calls.

- In-situ data: Bringing together selected experts to focus on topics of direct relevance to the in-situ data community within the GEO context.

- Law and policy: Working to Ensure the Legal Interoperability of Shared Data.

- Data Ethics: Definition of the sub-group ToR and next deliverables; study on possible Data Ethics issues with the use of cloud computing platforms/infrastructures.

* GEO Knowledge Hub latest additions:

- Using EO to calculate ration of land consumption rate to population growth rate

- Masada: Massive Spatial Automatic Data Analytics

- Open DataCube cloud statistics

* GEOSS Platform:

- GEOSS Platform Plus (GPP, an H2020 co-funded project, started on January 1st 2022):​ GPP contributes to the implementation of the Global Earth Observation System of Systems (GEOSS), aiming at evolving the European GEOSS Platform components to enable access to tailor-made information and actionable knowledge and will do so in close collaboration with the GEO partners (via the GEOSS Infrastructure Development Task Team) with a user-centric approach.​

- Ongoing developments:

Rethinking of several components and functionalities to address GDPR compliancy

YP: rethinking of YP form and implementation of a widget to be hosted on YP owner premises   ​

Mirror Sites: rethinking of mirror sites (MS as a widget) and a new wizard to support semi-automatic request and creation of MS on Communities premises ​

GEOSS Portal and GEO DAB: Implementation of flags to identify and manage empty collections

GPP first cycle of requirements collection: functionalities, architecture, services   ​

* Rethinking GEOSS

- EAG: timeline and approach: Review, assessment, recommendation

Doug Newman asked if they are using Graph DB technologies to implement this? If so, what specific technologies? Paola replied that they are using InvenioRDM.

## [Progress on ISO 19124-1](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_10.15_ISO19124-1%20and%20ARD%20Standards.pptx)

Liping Di (ISO/NASA) reported on progress on ISO 19124-1:

* ISO 19124-1: Fundamentals

The first part of ISO 19124, a multi-part ISO standard on calibration and validation of remote sensing data and derived products

Defining the common framework, methods, and UML structure

Allowing plug in other parts of ISO 19124 to be developed

CEOS, particularly the Cal/Val WG, is heavily involving in and contributing to the project

The Working Draft has been submitted to TC 211 for DTS vote in 3/2022; CEOS Cal/Val and WGISS shall provide comments on the standard draft

ISO 19124-1 standard is expected to be published before 6/2023

* ISO & OGC standard on Analysis Ready Data (ARD)

Plan to develop a multi-part ISO standard on ARD, which will also be an OGC standard

Part 1: ARD Framework

Part 2: ARD for Land

Many other parts (e.g., ARD for Ocean, Model output, ARD services)

Based on CEOS ARD initiative with CEOS support and to meet EO community needs

OGC Disaster Pilot 21 is developing the roadmap on ARD standardization and a new working item proposal (NWIP) to ISO TC 211 for part 1 of the ARD standard by 6/2022

Form an ISO project team and the OGC ARD Standard Working Group (SWG) to undertake the standard development, expected to start in July 2022 if funding is available

Need funding resource to support the work

* ISO TC 211 54th Plenary will be held in Vienna, Austria, May 12-13, 2022

Need CEOS WGISS input to provide liaison report to ISO TC 211; this should be a presentation that explains the work of WGISS, and summarizes activities from the last year.

## [WGDisasters](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_10.30_WG%20Disasters.pptx)

Laura Frulla (CONAE, WGDisasters vice-chair) gave a presentation on the activities of the working group. She began with the mission statement, objectives, and Chair priorities. She detailed each of the following activities, each of which has a group of participants comprising a subgroup:

* Landslide Demonstrator
* Volcano Demonstrator
* Seismic Hazards Demonstrator
* GeoHazards Lab
* GEO-DARMA
* Flood Pilot
* Geohazards Supersites and Natural Laboratories (GSNL)
* Recovery Observatory Demonstrator
* Wildfire Pilot

Makoto noted that the Flood Pilot is listed in the EAIL activities of WGISS.

# WGISS Plenary Session, Part 2

## [Future Meetings](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_10.30_WG%20Disasters.pptx)

Tom Sohre (USGS) outlined upcoming WGISS meetings as follows:

WGISS-54 - October 4-6, 2022 - Hosted by the Japan Aerospace Exploration Agency (JAXA) in Tokyo, Japan. This meeting will include a joint symposium with CEOS Working Group on Calibration and Validation (WGCV). The venue is The International House of Japan, 5-11-16 Roppongi, Minato-ku, Tokyo <https://www.i-house.or.jp/eng/facilities/>

WGISS-55 - April 2023 - Hosted by Comisión Nacional de Actividades Espaciales (CONAE) in Buenos Aires, Argentina. The meeting will be held at the Centro Cultural de la Ciencia – (C3) in Palermo, a lively neighborhood of Buenos Aires city. <http://c3.mincyt.gob.ar/> Godoy Cruz 2270, CABA [https://goo.gl/maps/QgEp FX5rxDFVNFPa7](https://goo.gl/maps/QgEp%20FX5rxDFVNFPa7). Tourism: <https://turismo.buenosaires.gob.ar/en>

WGISS-54 proposed schedule:

 Premeeting (Confirmation of Final Agenda, Logistics) Oct. 3 (Mon.), 2022

 Meeting (Including joint symposium with WGCV) Oct. 4 (Tue.)-6 (Thu.), 2022

 Facility Tour in Tsukuba Space Center (Option) Oct. 7(Fri.), 2022.

## [WGISS Summary and Discussion](https://ceos.org/document_management/Working_Groups/WGISS/Meetings/WGISS-53/3.Thursday/2022.03.24_12.10_WGISS-53%20Summary.pptx)

Makoto Natsuisaka (JAXA) gave the following meeting summary:

Plenary 1

* CEOS Chair CNES initiative “Paths to Sustainability: from strategy to practical measures” was introduced. The initiative defined three priorities in response to the Paris Climate Agreement, the Sendai Framework for Disaster Risk Reduction, and the 2030 Agenda for Sustainable Development.
* CEOS Work Plan 2022-2024 to be approved at SIT-37 and associated WGISS part (Chapter 3.9) were reported. Near-term plan, joint collaborations and information exchanges with other groups were summarized.

Data Discovery and Access

* CEOS/IDN collaborations, IDN upgrades, and IDN metrics were reported. FedEO migration to CEOS domain and DIF-10 export status were also reported. Both systems exchange registered metadata on daily basis.
* ESA GSTP project aiming to provide a cloud-native access to hyperspectral data was introduced. The formats trade-off resulted in the selection of Zarr format and GeoZarr format was proposed. Demonstrations on Jupyter Notebooks, OpenLayers, FedEO + STAC catalogue, and advanced applications were also made.
* A parallel STAC implementation in CWIC was proposed. STAC is a great candidate for a federated search solution and has a lot of traction in web- and cloud-based user communities.
* The current status on GEOSS Platform was reported. The platform consists of four main components, GEOSS Portal, GEO DAB, GEOSS Status Checker, and GEOSS Yellow Pages. The platform is brokering CEOS catalogues from IDN, CWIC and FedEO.
* The current status on service metadata and discovery best practice was reported. Based on the draft version, the contents were introduced. The comments and ideas were solicited.

Data Interoperability and Use

* Explore ARD with the Data Cube for Atmospheric Composition developed by EUMETSAT was introduced. The system was one of the perfect implementations enabling interoperability and use in the cloud.
* Internet of Things: Monitoring High Quality Geographical Products for Environment and Sustainability was introduced.

Data Preservation and Stewardship

* Object Storage Benchmark in ESA was reported. The benchmarks showed benefit of usability (Scalability, Manageability, etc.) and costs against small lack of performance.
* WGISS Data Management and Stewardship Maturity Matrix Use Case was reported. Main concept, use case for ENVISAT, etc. were introduced.
* Considering increase of data replication on cloud, a session on “Data Integrity and Authenticity on Cloud” was organized. To keep data integrity, data authenticity, provenance, traced modification, PID in the metadata, Hash Code in the metadata, watermark, KSI block chain were discussed in the previous meetings. Member agencies reported their approaches which include as the possibility to use Filename Lookup, Hash Lookup, Provable Data, Possession Signed Hashes, Checksum, and Non-Fungible Token (NFT). The necessity of further exchange of information on ongoing research and prototyping activities was pointed out.
* Common Online Dictionary Initiative, which is a joint activity with WGISS, WGCV, LSI-VC, and ISO/OGC, was introduced by the CEOS Common Terminology Group and called for more WGISS participation. Participant’s vocabularies and glossaries were provided by WGISS as input to the group.
* NRSCC’s Methodology for Data Publishing and Repository was reported.
* Topics for future DSIG sessions were agreed: Archiving technologies and approaches, preservation of space data associated content (information, software, tools), recovery of heritage data not accessible to users, data quality assessment and indicators (in cooperation with WGCV).

Technology Exploration

* Concept of CEOS Jupyter Notebooks Best Practice and exemplars was introduced.
* Jupyter Notebooks approach at EUMETSAT was introduced.
* “Jupyter Notebooks Day” at the end of June, a kind of hands-on event including CapD training, was proposed.
* Expectation for further webinars was expressed by the WGCapD representative.
* A co-lead of Technology Exploration Interest Group solicited future topics.

WGISS Cooperation with other Working Groups

* Latest information on GEO Data Working Group was updated. DWG consist of four subgroups; Data Management Principles, in-situ Data, Law and Policy, and Data Ethics. The status on developments of GEO Knowledge Hub and GEOSS Plus were shared.
* Progress on ISO 19124-1 and the New Joint ISO-OGC Standard Initiative on Analysis Ready Data was reported. The latter initiative is aiming at standardizing CEOS ARD. WGISS activities will be input to the ISO TC 211 54th Plenary will be held in Vienna, Austria on May 12-13, 2022.
* The activities of WGDisasters were introduced. WGDisasters consist of 9 subgroups, Landslide Demonstrator, Volcano Demonstrator, Seismic Hazards Demonstrator, GeoHazards Lab, GEO-DARMA, Flood Pilot, GSNL, Recovery Observatory Demonstrator, and Wildfire Pilot.

Plenary 2

* Member agencies reported the current activities.
* WGISS-54 will be held in-person in Tokyo in the week of Oct. 3rd, 2022. The logistics will be delivered later.
* Special Sessions in WGISS-54:

Joint symposium with WGCV

Data Integrity and Authenticity on Cloud (TBD)

AI/ML (TBD)

EAIL (TBD)

Conclusions

* WGISS-53 was held virtually on 22, 23, and 24 March 2022.
* Forty-four people from WGISS, CEO, CEO, WGCapD, WGDisasters, EUMETSAT, etc. participated in the meeting.
* Key issues, which are necessary for “data interoperability and co-use”, were identified. “Cloud” is the inevitable element and associating technologies such as data archiving in the cloud (including ARD and data cube), data analysis in place in the cloud (including AI/ML), data integrity and authenticity, knowledge share with Jupyter Notebooks should be continuously pursued.
* Further discussions in Living on Planet Symposium would be productive in addition to the WGISS regular meeting.

## Review of WGISS Actions

Michelle Piepgrass (JAXA) presented the open actions from past meetings. Comments were added and status updated:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Number** | **Category** | **Description** | **Actionees** | **Comments (March 24, 2022)** | **Status** |
| **Action WGISS-52-01** | Data Interoperability and Use | Continue working on the sustainability of the EAIL and possible expansion for nodes to include other similar systems hosted by other CEOS teams (true interop tests!) |   |   | Ongoing |
| **Action WGISS-52-02** | Data Interoperability and Use | Explore the Jupyter Notebooks theme with a view to EAIL providing a hosting environment for development and interop tests | Robert Woodcock, Esther Conway |   | Ongoing |
| **Action WGISS-52-03** | Data Interoperability and Use | Explore addition of JAXA data into EAIL | Robert Woodcock, Makoto Natsuisaka |  Makoto to discuss this with Jonathan Hodge | Ongoing |
| **Action WGISS-52-04** | Data Interoperability and Use | Prepare a symposium/workshop on EAIL Use-cases and vision for WGISS-53  | Robert Woodcock | Jonathan Hodge will present their work with a number of groups at WGISS-54. | Ongoing |
| **Action WGISS-52-05** | Data Interoperability and Use | Consider development of a Best Practice for Cloud data discovery and use (and demonstrate in EAIL). | Robert Woodcock and SLT | WGISS-54 | Ongoing |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Action WGISS-52-06** | Data Preservation and Stewardship | OSS & FDA inventory: verify with SLT the current inventories metadata and define how to include all inventory items in the IDN. In the meantime, update and sustain both inventories (maintain them in CEOS webpage or in WGISS and publicise them). | DSIG |   | Ongoing |
| **Action WGISS-52-07** | Data Preservation and Stewardship | WGISS DMSMM: prepare joint presentation with WGCV for WGISS-53 | DSIG | To be presented at WGISS-54 | Ongoing |
| **Action WGISS-52-07** | Data Preservation and Stewardship | WGISS DMSMM: prepare a webinar for WGCapD and technology session describing the maturity matrix use case after WGISS-53. | DSIG | WGISS-54 | Ongoing |
| **Action WGISS-52-09** | Data Preservation and Stewardship | CEOS common online dictionary: Progress with joint activity with WGCV on common online. CEOS Terms and Definitions wiki. | DSIG | To be presented at WGISS-54 | Ongoing |
| **Action WGISS-52-10** | Data Discovery and Access | Consider the possibility of having an alignment in implementing STAC (in all forms API, static search or metadata modelling) by the different stakeholders. | SLT | Consider the possibility of defining a guideline/best practice for the STAC at CEOS level, (as already done by WGISS for OpenSearch and are currently starting for service and tools discovery). This might be related to the STAC BP. | Ongoing |

Doug Newman commented that STAC discovery is regarded as an essential part of NASA's cloud evolution, and Michael Morahan noted that the IDN presentation showed how to refine for collections (data Products) on AWS.

## Concluding Discussion and Remarks

Makoto Natsuisaka (JAXA) closed the meeting, thanking everyone for their participation.

# Glossary of Acronyms

API Application Programming Interface

CEO CEOS Executive Officer

CEOS Committee on Earth Observation Satellites

CWIC CEOS WGISS Integrated Catalogue

DAAC Distributed Active Archive Center

DC Data Cube

DIF Directory Interchange Format

ECV Essential Climate Variable

EO Earth Observation

GEO Group on Earth Observations

GEOSS Global Earth Observation System of Systems

GIS Geospatial Information System

IDN International Directory Network

ISO International Standards Organisation

LSI Land Surface Imaging

OGC Open Geospatial Consortium

PI Persistent Identifier

PoC Point of Contact

SEO Systems Engineering Office

SIT Strategic Implementation Team

SLT System Level Team

ToR Terms of Reference

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