GEOSS Interoperability Workshops

Brought to you by the GEOSS Standards and Interoperability Forum (SIF)

Steve Browdy – SIF Chairperson

October 23, 2017

Past Workshops

- Virtual workshops held online

- Held for those interested in GEOSS/GCI, but specifically for data practitioners and data providers involved in, or interested in, the implementation of the GEOSS DMPs

- Slides of the presentations, as well as the recordings, are available at the SIF Workshop site
  - https://sites.google.com/site/geosssifworkshops/

GROUP ON
EARTH OBSERVATIONS

Past Workshops

- 2016 GEOSS SIF Interoperability Workshop
  - 2016-OCT-20
  - Over 60 registrants
  - Focus on GEOSS DMPs
    - Metadata for Discovery
    - Traceability
    - Quality Control
  - Also presentations on:
    - Open Science
    - FAIR DMPs
    - Community Portal Interaction with GCI
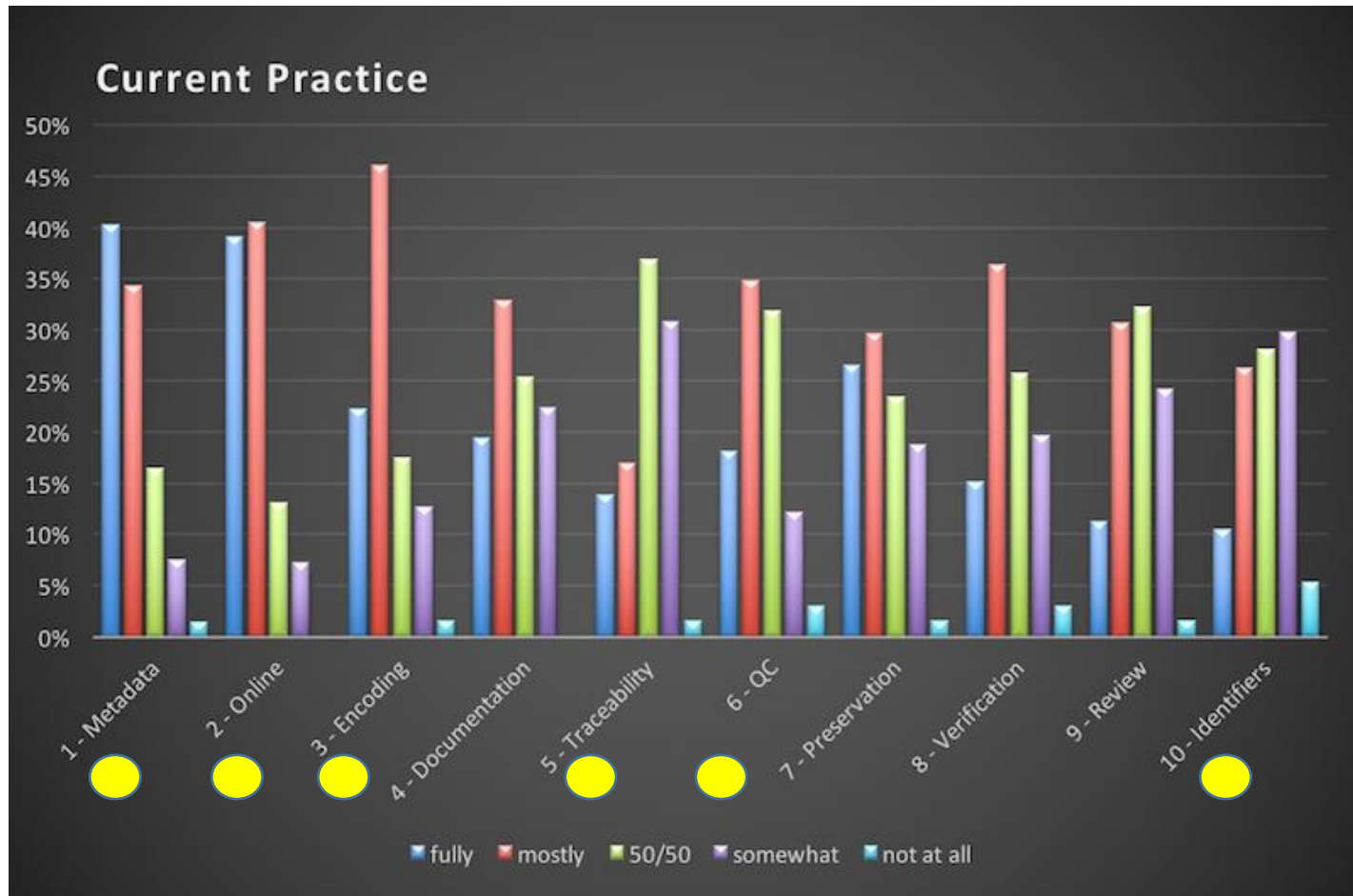   - Capacity Building

Past Workshops

- 2017 GEOSS SIF Interoperability Workshop
  - 2017-AUG-28
  - Over 70 registrants
  - Focus on GEOSS DMPs
    - Online Access
    - Encoding
    - Persistent and Resolvable Identifiers
  - Also presentations on:
    - NextGEOSS H2020 Project
    - Community Portal Interaction with GCI
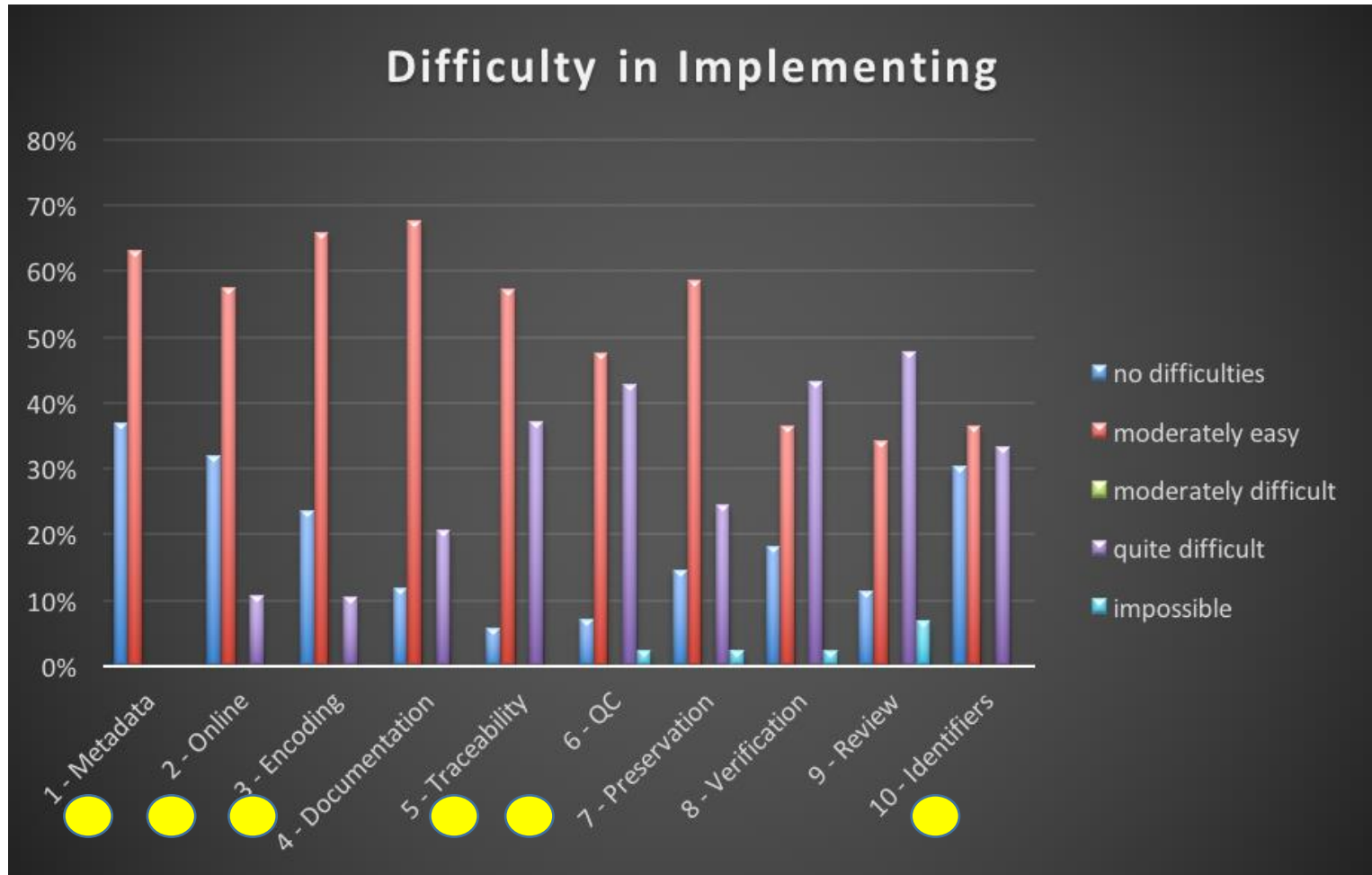   - Capacity Building

DMPs Considered

- DMP-1 *Metadata for Discovery*
- DMP-2 *Online Access*
- DMP-3 *Data Encoding*
- DMP-5 *Data Traceability*
- DMP-6 *Quality Control*
- DMP-10 *Persistent & Resolvable Identifiers*

- Some of the survey pictures provided by:
  - Siri Jodha Khalsa
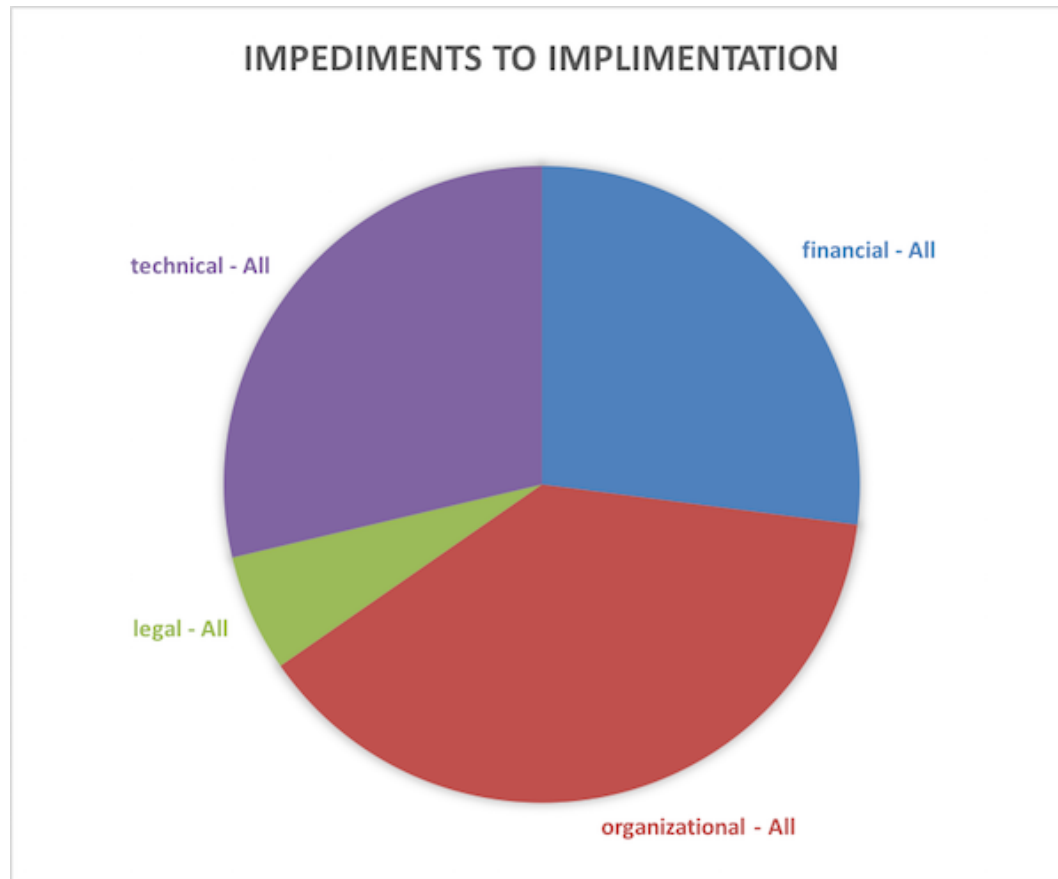 - Robert Downs
  - Marie-Francoise Voidrot

DMPs in Practice

DMP Implementation

Difficulty in Implementing

Chart legend: no difficulties, moderately easy, moderately difficult, quite difficult, impossible

Categories: 1 - Metadata, 2 - Online, 3 - Encoding, 4 - Documentation, 5 - Traceability, 6 - QC, 7 - Preservation, 8 - Verification, 9 - Review, 10 - Identifiers

Impediments to Implementation

IMPEDIMENTS TO IMPLIMENTATION

technical - All
financial - All
legal - All
organizational - All

GROUP ON
EARTH OBSERVATIONS

DMP-1 Metadata for Discovery

Data and all associated metadata will be discoverable through catalogues and search engines, and data access and use conditions, including licenses, will be clearly indicated.

GROUP ON
EARTH OBSERVATIONS

DMP-1 Survey results

- Compliance with DMP-1

Fully, Mostly	50/50	Somewhat / Not At All
75%	16%	9%

- Difficulty in Implementing DMP-1

| No difficulties	Moderately Easy	Moderately Difficult	Quite Difficult	Impossible
37%	63%	0%	0%	0%

- Challenges

Financial	26%
Organizational	43%
Legal	5%
Technical	25%

GEO GROUP ON EARTH OBSERVATIONS

@GEOSEC2025
www.earthobservations.org

DMP-1 Interoperability Issues

- Access and Use Conditions (licenses & waivers)
  - Semantics involved
  - Impacts legal interoperability
  - Currently only GEOSS Data-CORE compatible licenses and waivers are considered
    - Work going on to expand this
- Catalogs and metadata must be maintained appropriately and validated to achieve reliable discovery over time
  - Costly and time consuming

DMP-2 Online Access
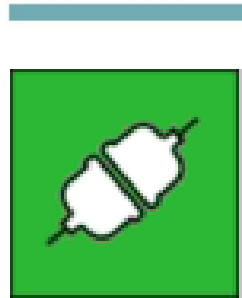
Data will be accessible via online services, including, at a minimum, direct download but preferably user-customizable services for access, visualization and analysis.

DMP-2 Survey results

- Compliance with DMP-2

Fully, Mostly	50/50	Somewhat / Not At All
77%	13%	7%

- Difficulty in Implementing DMP-2

| No difficulties	Moderately Easy	Moderately Difficult	Quite Difficult	Impossible
21%	38%	25%	7%	0%

- Challenges

Financial	24%
Organizational	25%
Legal	21%
Technical	30%

GROUP ON EARTH OBSERVATIONS

@GEOSEC2025
www.earthobservations.org

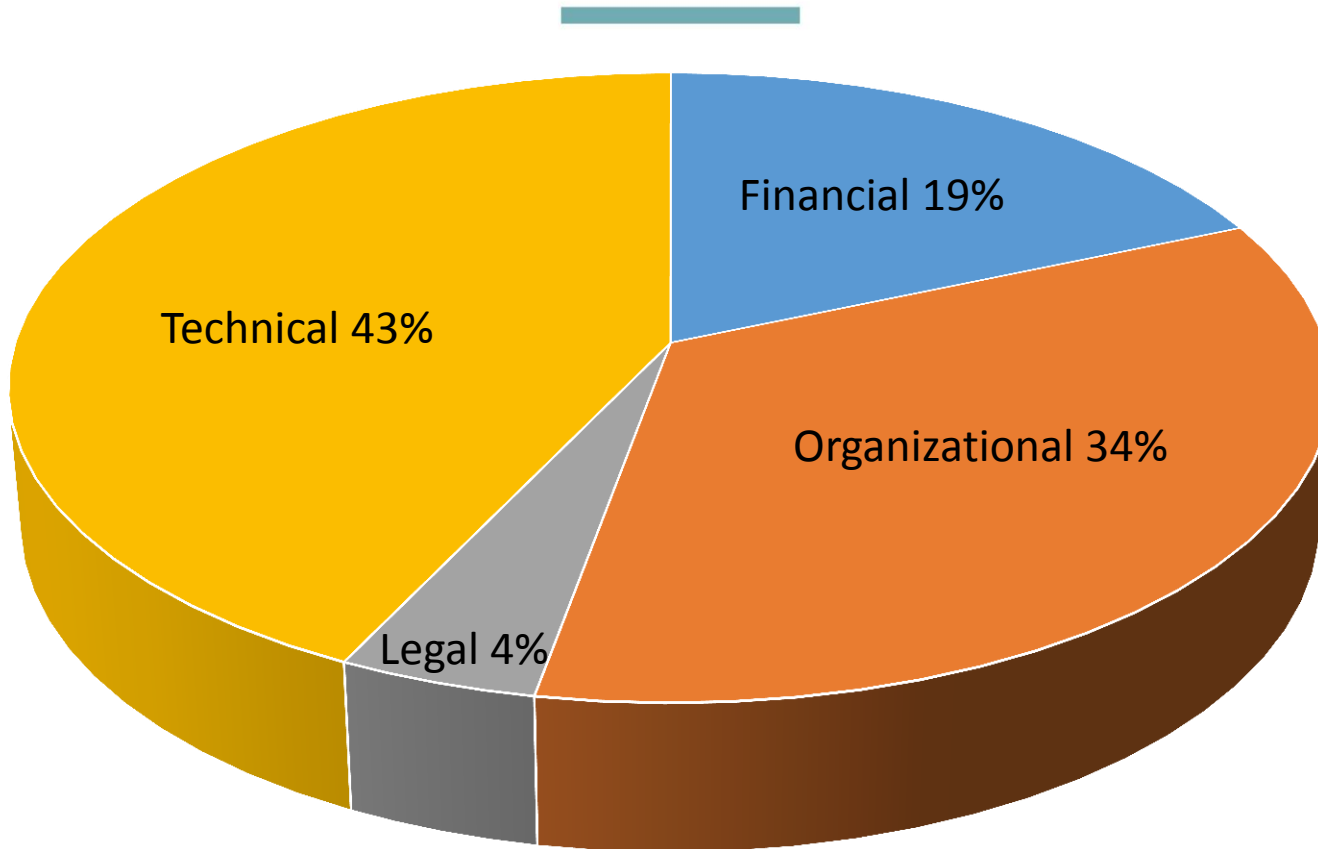# DMP-2 Interoperability Issues

- Single Sign-On
  - Authentication
  - Authorization
  - GEOSS-wide federation
    - Trust issues
    - Variety of mechanisms

- In-place processing
  - Becoming more present (minimize data movement)
  - Takes burden off of network and local machines/servers

- Semantic concerns
  - Need better mediation/mapping between vocabularies
  - Maybe too many vocabularies

DMP-3 Data Encoding
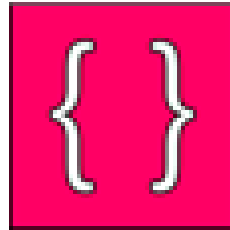
Data should be structured using encodings that are widely accepted in the target user community and aligned with organizational needs and observing methods, with preference given to non-proprietary international standards.

DMP-3 Survey Results

Financial 19%

Organizational 34%

Legal 4%

Technical 43%
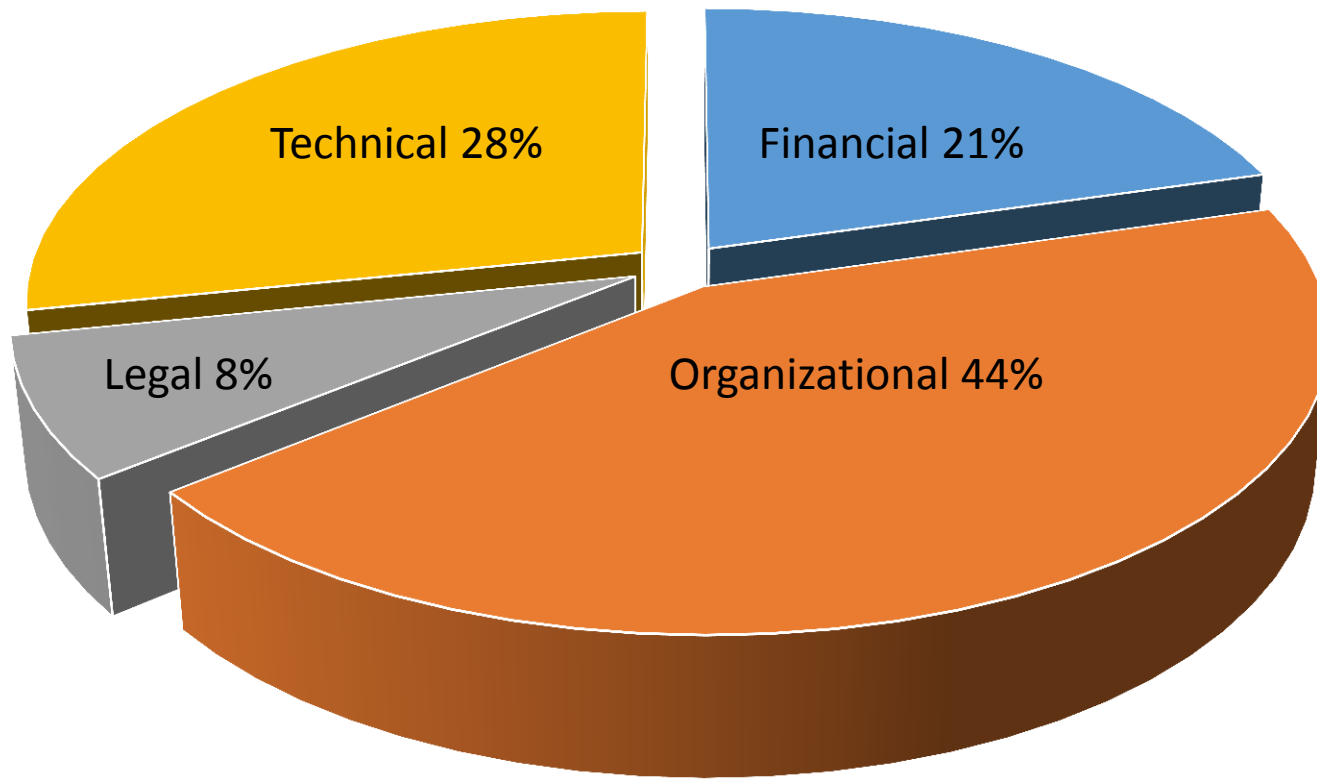
DMP-3 Interoperability Issues

- Semantic interoperability
  - Need better mediation between:
    - communities and disciplines
    - Ontologies, vocabularies, taxonomies, etc.
  - Affects analysis, visualization, and integration

- Legal interoperability
  - Need expanded set of licenses
  - Need encodings of the legal parameters
  - Is impacted by semantics

DMP-5 Data Traceability

Data will include provenance metadata indicating the origin and processing history of raw observations and derived products, to ensure full traceability of the product chain.

GROUP ON
EARTH OBSERVATIONS

DMP-5 Survey Results

Technical 28%
Financial 21%
Legal 8%
Organizational 44%

GEO GROUP ON EARTH OBSERVATIONS

@GEOSEC2025
www.earthobservations.org

DMP-5 Interoperability Issues

- Completeness

- Correctness

- Semantic compatibility
  - Translations between incompatible provenance elements

- Interpretability
  - Standardize values for provenance elements

DMP-6 Quality Control
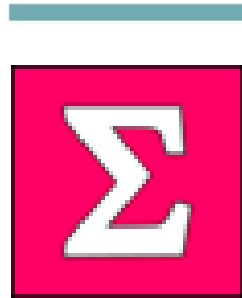
Data will be quality-controlled and the results of quality control shall be indicated in metadata; data made available in advance of quality control will be flagged in metadata as unchecked.

DMP-6 Survey results

- Compliance with DMP-6

Fully, Mostly	50/50	Somewhat / Not At All
49%	30%	14%

- Difficulty in Implementing DMP-6

No difficulties	Moderately Easy	Moderately Difficult	Quite Difficult	Impossible
4%	28%	30%	25%	1%

- Challenges

Financial	31%
Organizational	36%
Legal	1%
Technical	32%

GROUP ON
EARTH OBSERVATIONS

DMP-6 Interoperability Issues

- Appropriate definition of QC parameters
- Semantic mediation between communities and disciplines
- Impacted by:
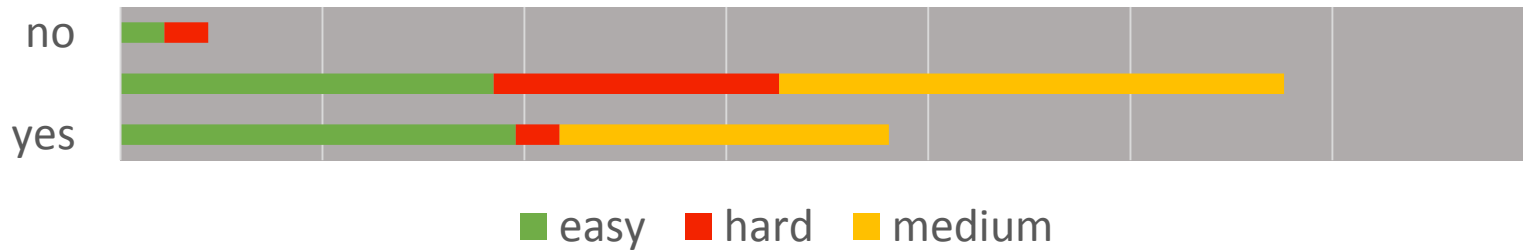 - Encoding
  - Metadata discovery
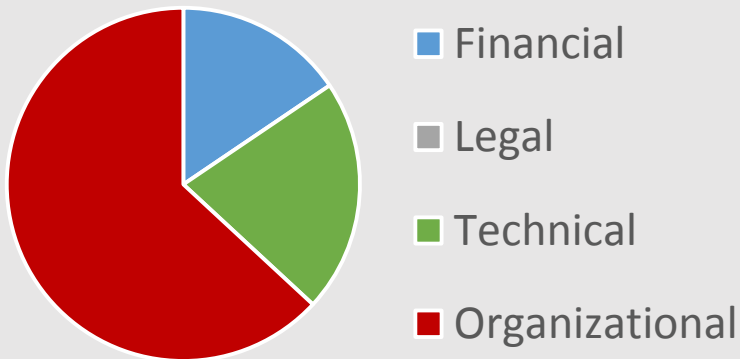
DMP-10 Persistent and Resolvable Identifiers

Data will be assigned appropriate, persistent, unique, and resolvable identifiers to enable documents to cite the data on which they are based and to enable data providers to receive acknowledgement for use of their data.
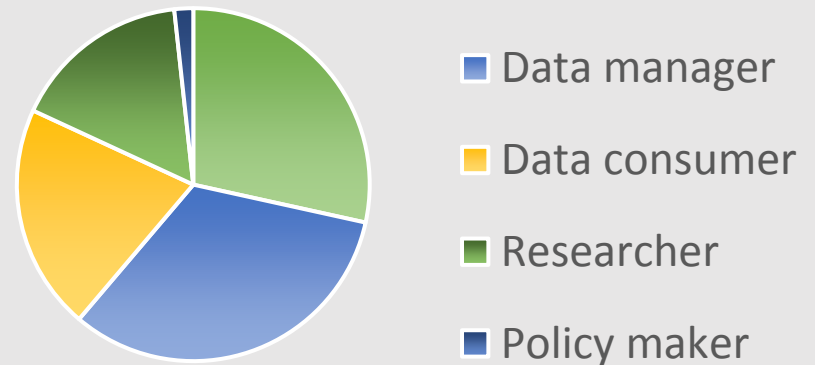
DMP-10 Survey Results

Practice and Difficulty

no

yes

■ easy   ■ hard   ■ medium

Challenges

■ Financial
■ Legal
■ Technical
■ Organizational

Positions that Practice

■ Data producer
■ Data manager
■ Data consumer
■ Researcher
■ Policy maker

GEO GROUP ON EARTH OBSERVATIONS

@GEOSEC2025
www.earthobservations.org

DMP-10 Interoperability Issues

- Automation of DOI implementation
  - Manual process is too slow and not scalable
- Dynamic linkages across multiple resources is not established
- Institutionalizing the concept of citing data takes time

Future Considerations

- ***Research towards moving to a peer-to-peer federated data sharing infrastructure***
  - ***Utilizes old and new technologies in evolutionary ways***
  - ***Mitigates issues with centralization***
  - ***Addresses security and privacy issues, if applicable***
  - ***Simplifies governance model***
  - ***Allows easier participation***
- ***Standards remain key to interoperability success***
- ***5-10 years for mainstream adoption***

Thank You

Communicate and Collaborate with GEO:

GEO GROUP ON EARTH OBSERVATIONS