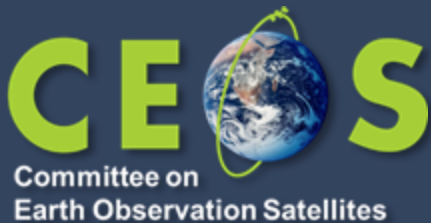


Interoperability Handbook Vocabulary (Semantics)



Peter Strobl, EC-JRC

Agenda Item 11.2

WGISS-60

13-17 October, 2025

DLR, Oberpfaffenhofen, Germany

How it started

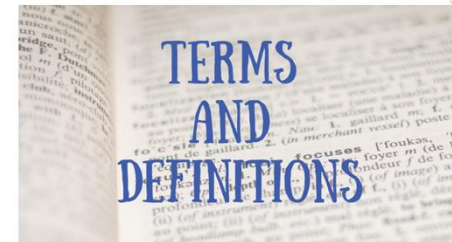


Background

- Joint initiative of CEOS WGCV, CEOS WGISS, and CEOS LSI-VC
- **Mandate: CEOS WGCV Action Item 49-06 (June 2021) “[...] to advance the idea of a CEOS common online dictionary, with a view to eventually reaching out to launch a broader community effort”**

Summary of activities

- CEOS Common Terminology group established August 2021
- Workplan drafted
- Integrated WGISS Data Stewardship Glossary (pdf) and NOAA NESDIS Data Management Lexicon (pdf) into existing WGCV ‘Terms and Definitions’ Wiki
- Discussing categorization of terms-prototyping approach using base terms and high impact terms
- Discussing governance
- Reaching out to promote initiative internally with focus on WGISS, WGCV, LSI-VC externally with ISO/OGC
- Summarize findings – publish – move on!



http://calvalportal.ceos.org/ca/t-d_wiki

Format of terminologies



TC 211 Geographic information/Geomatics

ISO/TC 211 Multi-Lingual Glossary of Terms (MLGT)

NEWS & ARTICLES REGISTERS STATISTICS ABOUT FEEDBACK SEMANTIC SEARCH

All Concepts

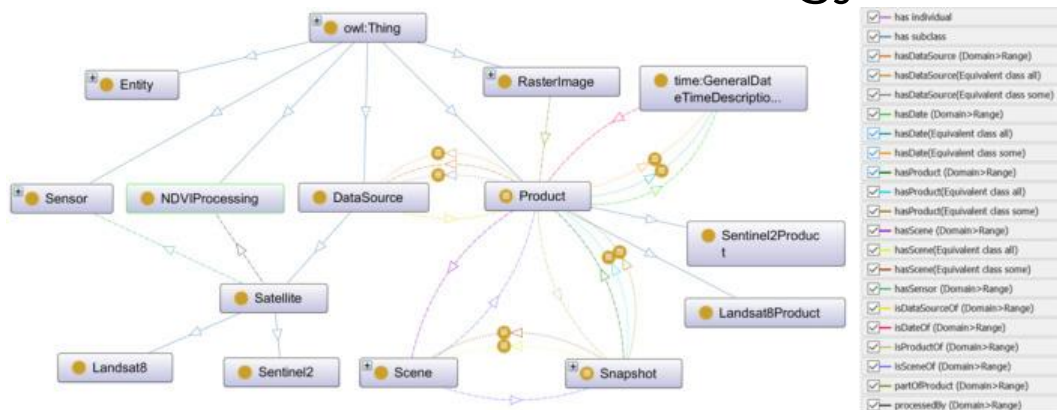
The concepts below are sorted by term name, alphabetically, in a case-insensitive way.

Term ID	Term
815	a posteriori classification
816	a priori classification
2	abbreviation
1928	absolute accuracy
2528	abstract
2028	abstract root
2471	abstract test case
3	abstract test case
4	abstract test method
5	abstract test module

❖ Mostly alphabetical lists

- Often in PDFs
- Hard to 'explore'

❖ Or formal ontology



- ❖ Simple **lists of words** are not used often – structure is important
- ❖ Lack of **versioning** (or keep older definitions) – and definitions change
- ❖ **Ambiguity** of definitions for some terms – e.g. ISO online browsing platform has 77 versions of ‘observation’
- ❖ **Inconsistent** definitions (e.g. in-situ, observation, sample, ...)
- ❖ **Superficial** definitions (e.g. interoperability) lacking full framework
- ❖ **Circular** definitions – and poor use of the foundational ‘base terms’
- ❖ Development process – **isolated** efforts – creates these problems

Lost in translation: The need for common vocabularies and an interoperable thesaurus in Earth observation sciences

P.A. Strobl¹, E.R. Woolliams² and K. Molch³

¹EC-JRC, ²NPL, ³DLR (now ESA)

Journal: Surveys in Geophysics

Online: 1 October 2024

<https://doi.org/10.1007/s10712-024-09854-8>

- ❖ First draft (July 2023) reviewed by 12 observation experts
- ❖ Journal peer review included significant challenge from ontology expert
- ❖ Challenging review procedure (8 months)
- Final paper has benefited from the insights of all these perspectives.



Surveys in Geophysics
<https://doi.org/10.1007/s10712-024-09854-8>

Lost in Translation: The Need for Common Vocabularies and an Interoperable Thesaurus in Earth Observation Sciences

P. A. Strobl¹  · E. R. Woolliams²  · K. Molch³

Received: 15 December 2023 / Accepted: 17 July 2024
© The Author(s) 2024

Abstract

The Earth Observation sciences are highly multidisciplinary with long value chains from the development, characterisation and deployment of sensors, through data processing and modelling, to the information services provided to decision makers in, for example, governments, companies and non-governmental organisations. A prerequisite to any multidisciplinary collaboration is effective communication and many communities involved in the value chains have developed vocabularies or terminologies to define terms from a particu-

A good thesaurus is:



Consistent

On a foundation of base terms. No ambiguity. No circular or overlapping definitions. Clear preferred terms. Alternative definitions explained clearly.

Interrelated

Clearly shows relationships between terms – in a definitional, contextual and ontological way. Mutually exclusive parent / sibling / child relationships

Understandable

Clear definitions, with differentiation/disambiguation of controversial terms. Checked for understandability by multidisciplinary teams

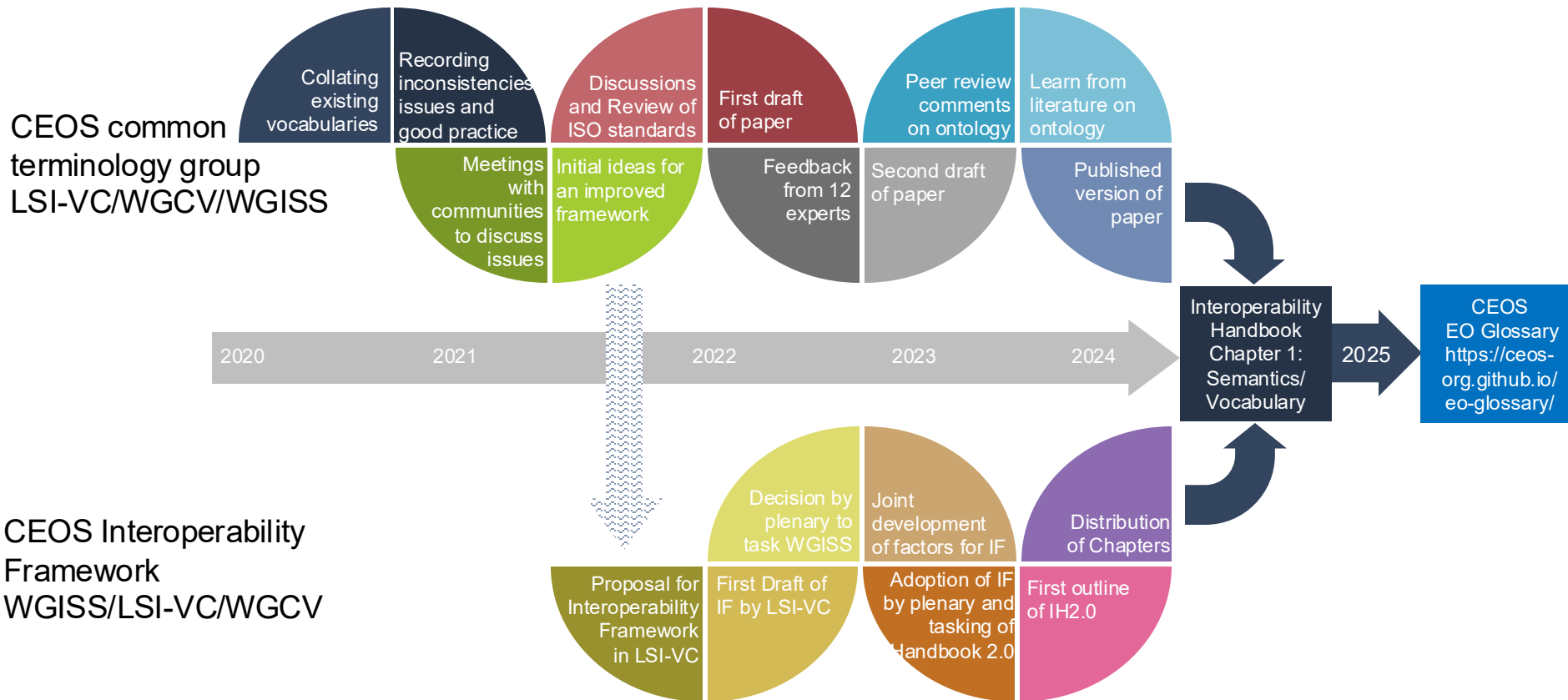
Educational

Not just for computers! Promotes a conceptual framework, shows linkages, satisfies curiosity, helps communication of concepts. Links to detailed frameworks for high-impact terms.

Updateable

Unified thesaurus with version control at the level of individual terms, methods for providing disambiguation links, and adding new terms.

Semantics meet Interoperability



A better approach to Semantics



Factor 1 of the Interoperability Framework as phrased in the Handbook 2.0

SEM#1	... an open Earth observation thesaurus, ... through GitHub	✓
SEM#2	... public comment and discussion on existing and new terms and definitions	✓
SEM#3	... version control and change management ...and link to ... alternative definitions	✓
SEM#4	use of project or document specific vocabularies ... discouraged...	
SEM#5	... promote the common thesaurus ... to strive for domain wide adoption	
SEM#6	... online repositories for abbreviations and acronyms ... controlled vocabularies	

Full version at: <https://github.com/ceos-org/interoperability-handbook/blob/main/Vocabulary.md>

A better Thesaurus



THES#1	... the thesaurus should be consistent and divided into classes ...
THES#2	... may not contain the term itself nor other circular ...
THES#3	...clear and mappable relationship...
THES#4	... unambiguous and short ... can replace the term
THES#5	... 'Notes' and ... 'Examples' section...
THES#6	... accompanying 'Sources' section ...
THES#7	... version controlled at the individual term level ...
THES#8	... contradictory definitions ... with links ... and explanations

Clear rules for
writing clear
definitions,
in a common
Thesaurus,
**and everywhere
else!**

Not only a Thesaurus!



SEM#6 ... online repositories for **abbreviations and acronyms** ... **controlled vocabularies**

A million ways to call a mission/platform/instrument:

Sentinel-2 A, Sentinel 2A, S2-A, S-2A, S2 A, S2A, ...

Urgently needed is a scheme for harmonised identifiers across EO assets such as:

“sensor, instrument, platform, constellation, mission, campaign”

Who if not CEOS can deliver this?

Not only a Thesaurus!



Each field has 4 “incarnations” for different use cases:

fullName (≤ 256), shortName (≤ 32), acronym (≤ 16), mnemonic (≤ 8 , file/URI safe)

Design notes:

- ❖ No colons (:) in IDs to remain filename-safe (use - and _)
- ❖ Distinguish sensor vs instrument (sensor = concrete device; instrument = class/type)
- ❖ Support arrays for fused products (multiple instruments/sensors)
- ❖ Provide STAC/EO mappings (platform/mission/constellation + satellite extension)
- ❖ Extend to bands/modes.

See George Dykes’ [conversation starter](https://github.com/gamedaygeorge/ceos-db-toolkit/blob/canonical-id-demo/colab-notebooks/canonical_id_demo/canonical_id_demo.ipynb) on Github:

https://github.com/gamedaygeorge/ceos-db-toolkit/blob/canonical-id-demo/colab-notebooks/canonical_id_demo/canonical_id_demo.ipynb

Way forward



- ❖ **CEOS EO Glossary is online!**
(details this afternoon)
- ❖ Opportunity to participate and contribute **open now!**
- ❖ Links to **CEOS-ARD** and **STAC** development have been established
- ❖ **More to come!**

This glossary is currently at an early stage and still work in progress. It is not 100% reviewed yet and some things might change over time. Stay tuned in our [blog](#) or on [GitHub](#)!

CEOS EO Glossary

Home Glossary Contribute Blog Contact

EO Glossary

We are bridging Earth Observation communities. Get started with the glossary [here](#) or contribute on [GitHub](#)

Search terms...



Topology

Instead of creating a taxonomy we work with cross-links. Such a child-parent topology helps you quickly grasp the relevant context of a definition. See the diagram.

WGCV-WGISS -2024-10	WGVISS & WGCV to create a joint deliverable for the 2025-2027 CEOS Work Plan regarding the development of the CEOS/KCEO Glossary. Both working groups should ensure enough people are engaged, including from other CEOS groups, as well as from the external community. WGCV & WGISS Chairs to include this in their respective Plenary presentations, to solicit support to brand the glossary as the 'CEOS/KCEO Glossary'.	Q4 2024
WGCV-WGISS -2024-11	SEO to discuss with Paolo Castracane & Peter Strobl (and perhaps NOAA Knowledge Graph team) about transferring the current Cal/Val Portal Terms & Definition Wiki to GitHub KCEO repository, considering the appropriate format.	Q1 2025

Action Lead: Cody Anderson, Tom Sohre

Action Lead: Brian Terry, Paolo Castracane, Peter Strobl

Thank you!

