(Towards)

A metadata model for atmospheric data resources

Anne De Rudder and Jean-Christopher Lambert
Belgian Institute for Space Aeronomy (IASB-BIRA), Brussels
The context

Projects that have contributed to the GMES/Copernicus Atmospheric Service

- EU FP7
  - Ground-based atmospheric observing system
    - GeMon
- Ground-based data for MACC validation
  - NORS
  - Validation
- Core & downstream service development
  - PROMOTE
    - Ozone, UV, AQ, climate study, aviation control
  - PROMOTE 2
- Ground-based atmospheric observing system
  - GEMS
    - GHG, reactive gases, aerosol, AQ, solar radiation
- MACC
  - Core service
    - Input and validation
  - MACC II
    - Final products
- PASODOBLE (Air Quality)
  - Final products
- EVOSS (Volcanos)
  - Final products
- Downstream service
- End users
  - Research, monitoring, environmental policy, aviation control, health, sport, leisure, …

Final products

Obligation of INSPIRE compliance

INTEROPERABILITY
EU contribution to

the

Global Earth Observation System of Systems (GEOSS)
Quality

Quality Assurance Framework for Earth Observation

is a contribution to the GEO GROUP ON EARTH OBSERVATIONS

‘providing guidance on the quality assurance of Earth Observation data’

- Provision of quality indicators (e.g. uncertainties) allowing the user to assess whether the product is « fit-for-purpose »
- Traceability of instrument calibration and data generation
- Traceability of validation (i.e. of quality indicator calculation)
- Accordance with standard procedures and, as far as possible, SI
Dimensionality

Atmospheric (and more generally Earth Observation) data are usually 4-dimensional.

Atmospheric data may depend on a variety of independent variables (not only longitude and/or latitude) which, as such, have all the same status and should be documented symmetrically.
Integration of existing concepts and models

Generic Earth Observation Metadata Standard (GEOMS) metadata scheme

- HDF-based
- promoted by ESA
- used by space agencies, NDACC, Cal/Val databases EVDC (ESA), AVDC (NASA)

Climate Forecast (CF) Metadata Convention

- NetCDF-based
- widely used within the climate community

Forthcoming metadata guidelines from the GEO Air Quality (AQ) Community of Practice (CoP)
Preliminary questions, options, decisions

Granularity
What is the type of object we want to describe with one given metadata set? In other terms, if a dataset is the entity we want to describe, what do we define as a dataset?

Purpose
What is the metadata intended for?

Volume
How detailed (= voluminous) do we want the metadata scheme to be?

Compliance with standards
Which standards? What are the implications?
Metadata quality

What do we need to care for in order to get good (pertinent, useful, complete) metadata?

• A good metadata scheme does not guarantee good metadata! – It does not even guarantee compliance with some adopted standard
• Provide clear definitions of metadata elements (+ possibly examples)
• Where possible, provide lists of preselected metadata values
• Make room for atypical cases by offering options such as
  o other
  o not applicable
  o unknown (yet)
  o [void] ↔ either appears as not provided in the output or does not appear at all
  Applies even to mandatory metadata fields, except if the field value determines which branch of the scheme must be followed (e.g. if the resource is a dataset, the applicable metadata fields will be different from those for a service)
• Avoid redundancies
An atmospheric metadata model

Derived from the PASODOBLE detailed metadata model

A Product description
   A1 Identification and nature
   A2 Domain of definition
   A3 Domain of values
   A4 Quality
   A5 History
   A6 Presentation, coding, format
B Product generation
   B1 Developer
   B2 Derivation chain and intermediate validation
C Product availability, distribution and use
   C1 Provider
   C2 Applications
   C3 Maintenance and update
   C4 Delivery
   C5 Terms and conditions
   C6 Support
D References
   D1 Publications informing on the resource
   D2 Standards of reference
E Metadata
INSPIRE discovery metadata

• Described in the INSPIRE Metadata Regulation (*).
• Includes the set of top-level metadata supposed to apply to any spatial resource – Completed by data specifications – one for each INSPIRE data theme – there is one for Atmospheric Conditions and Meteorological Geographical Features.

Mapping with INSPIRE

A  Product description
   A1  Identification and nature
   A2  Domain of definition
   A3  Domain of values
   A4  Quality
   A5  History
   A6  Presentation, coding, format

B  Product generation
   B1  Developer
   B2  Derivation chain and intermediate validation

C  Product availability, distribution and use
   C1  Provider
   C2  Applications
   C3  Maintenance and update
   C4  Delivery
   C5  Terms and conditions
   C6  Support

D  References
   D1  Publications
   D2  Standards

E  Metadata

INSPIRE discovery metadata
1. Identification
2. Classification
3. Keyword
4. Geographic location
5. Temporal reference
6. Quality and validity
7. Conformity
8. Constraint related to access and use
9. Organisation responsible for the establishment, management, maintenance and distribution of spatial data and services
10. Metadata on metadata
Compliance issues to be solved

A Product
   A1 Identification and nature
   A2 Domain of definition
   A3 Domain of values
   A4 Quality
   A5 History
   A6 Presentation, coding, format

B Product generation
   B1 Developer
   B2 Derivation chain and intermediate validation

C Product availability, distribution and usage
   C1 Provider
   C2 Applications
   C3 Maintenance and update
   C4 Delivery
   C5 Terms and conditions
   C6 Support

D References
   D1 Publications
   D2 Standards

E Metadata

Existence of quality assessment
Validation method
Validation datasets
Validation results
Content analysis
Quality indicators
a major user requirement

The nature of the physical quantities and their units need to be documented

The horizontal resolution may not be identified to data quality
Mapping with INSPIRE (cont.)

A Product description
A1 Identification and nature
A2 Domain of definition
A3 Domain of values
A4 Quality
A5 History
A6 Presentation, coding, format

B Product generation
B1 Developer
B2 Derivation chain and intermediate validation

C Product availability, distribution and use
C1 Provider
C2 Applications
C3 Maintenance and update
C4 Delivery
C5 Terms and conditions
C6 Support

D References
D1 Publications
D2 Standards

E Metadata

INSPIRE discovery metadata

1. Identification
2. Classification
3. Keyword
4. Geographic location
5. Temporal reference (*)
6. Quality and validity
7. Conformity
8. Constraint related to access and use
9. Organisation responsible for the establishment, management, maintenance and distribution of spatial data and services
10. Metadata on metadata

(*): Any of the following as long as it is a date in the format yyyy-mm-dd

- First date of data collection
- Dataset creation date
- Resource publication date
- Metadata publication date

URI
URL
Conclusions

- Various efforts are currently going on to harmonise Earth Observation metadata and comply with the INSPIRE Metadata Regulation
- Compliance with the principles of QA4EO
- Integration of metadata models in use (GEOMS, CF)
- Ability to encompass data depending on any independent variable(s)
- Proper handling of quality concepts / quality indicators
- The information content should prevail on formal aspects
Thank you