

# Connecting WDCA / WDCRG to WIGOS & OSCAR: Lessons and Issues

M. Fiebig, K. Tørseth, R.O. Rud



Norsk institutt for luftforskning  
Norwegian Institute for Air Research

# Agenda

1. Filling in the record.
2. Vocabulary for observed variables – the need for modifiers
3. How to document licenses
4. Some more minor issues & news

# Filling in the Record

- WMDR can be very comprehensive. Implies often that many vocabulary items need to be defined.
- Also possible to fill a minimum about the observation, and extend as vocabulary becomes available (header, observation).
- Leave out items which are controlled centrally.
- Concept of «Deployment» may not be transferable (granularity), may lead to fragmentation.
- Online validator should be updated at all times!
- OSCAR doesn't recognise "inapplicable" for observed variable, even though defined in code list, and CF name is attached.

# Vocabulary for observed variables – the need for modifiers

**Example:** particle light scattering coefficient

- Size fraction: pm 1, pm2.5, pm10, TSP
- T / p reference conditions: standard, ambient, lab
- Sample conditioning: dry, dried, ambient
- Some use cases require modifiers, others don't.
- WMDR now includes modifiers with variable names -> 1700+ new variables for WDCA! (only the ones actually used – not feasible)
- GALION and WDCPC have similar challenges.

# Vocabulary for observed variables – approach to modifiers

1. Don't define further variable names including modifiers. This would lock us into an inflexible solution.
2. Simultaneously
  - a. Start process amending WMDR to include modifiers. Can be a simple key / value pair. Once defined, solution can be used in “pre-approval” mode.
  - b. All observed variables defined now will only comprise core variable.
3. Collect requirements for modifiers needed.

# Licenses in WMDR

- Today, WMDR knows: DataPolicy, Attribution
- Data policy (ethical statement) and license (legal statement) live in different realms, can even be combined.
- Need for documenting data policy and data license separately.
- License to be identified uniquely to facilitate brokering (later).

# Possible License Metadata

## Data license metadata [0..\*]

- License identifier (key/value)(controlled vocabulary preferred, free text for legacy data)  
DataCite: 16 Rights (free text), 16.b rightsIdentifier (SPDX license list preferred)
- Link to license (online resource)
- Attribution (free text) [0..\*]  
Having the option of having several acknowledgement texts allows for attribution stacking (necessary feature), and allows to test whether an acknowledgement already exists before adding another one.
- Licensee (key/value)[1..\*], include option for PID of licensee, “general public” as default, clarify whether it needs to be mandatory.

# Minor Issues & News

- Use of term “aerosol”
- Link to external provenance metadata
- Work on DOIs (would fill separate presentation)
- ...