



Enhancing data quality and data publication with the Meteolo library

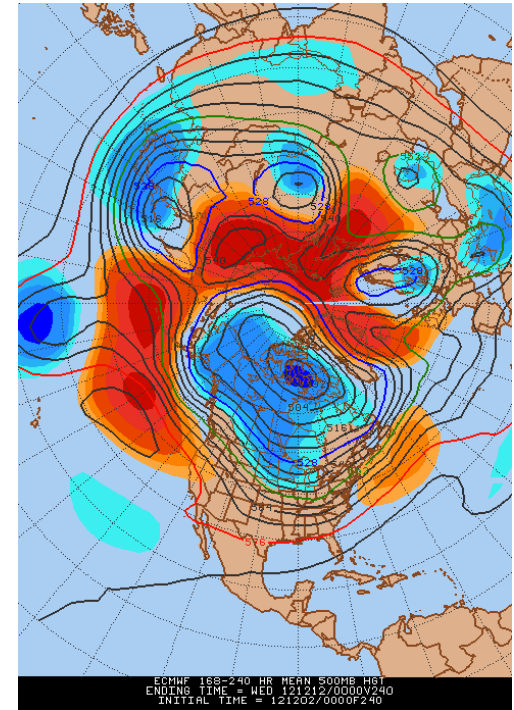
M. Bavay, J. Fiddes, Ch. Fierz, Ø. Godøy, R. Nitu

At the beginning...



Raw Data

- Often from weather stations
- But also from models (reanalysis, forecasts)
- Different formats / protocols
- Sometimes very good
- Sometimes, very bad
- Extract gold out of junk!



At the beginning...

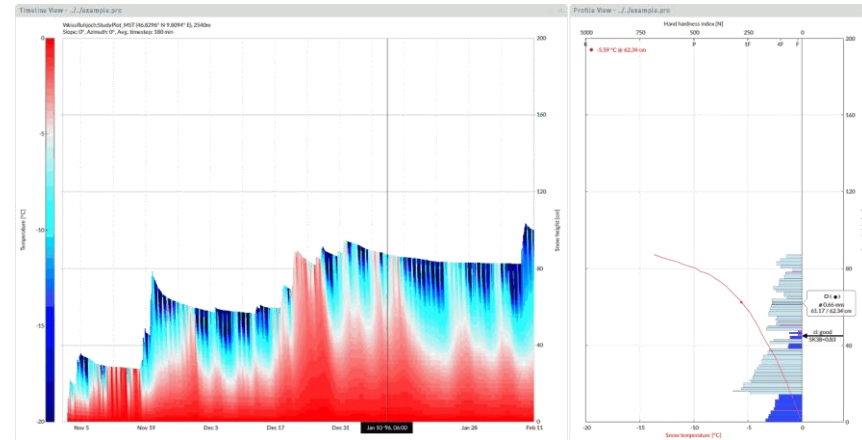
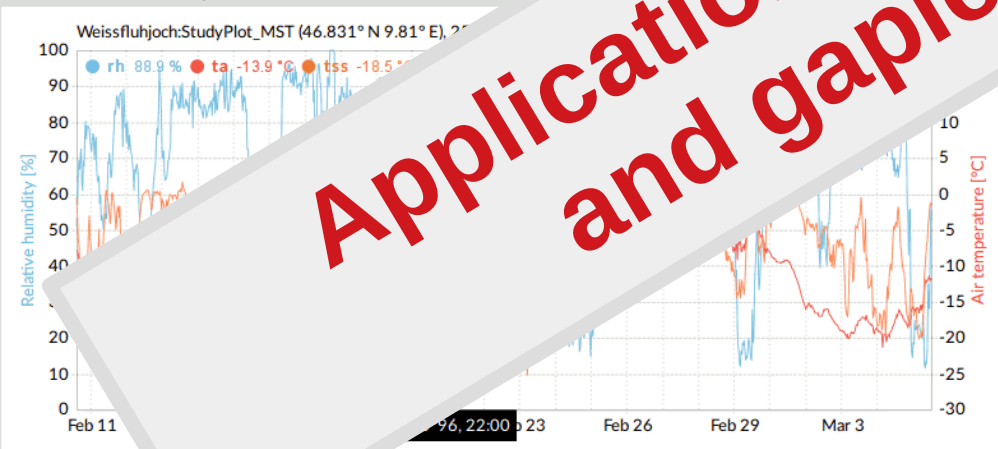
- Web dashboards
- Models (ex. Snowpack, GeoTop, Amund)

Applications require good and gapless data!

Applications



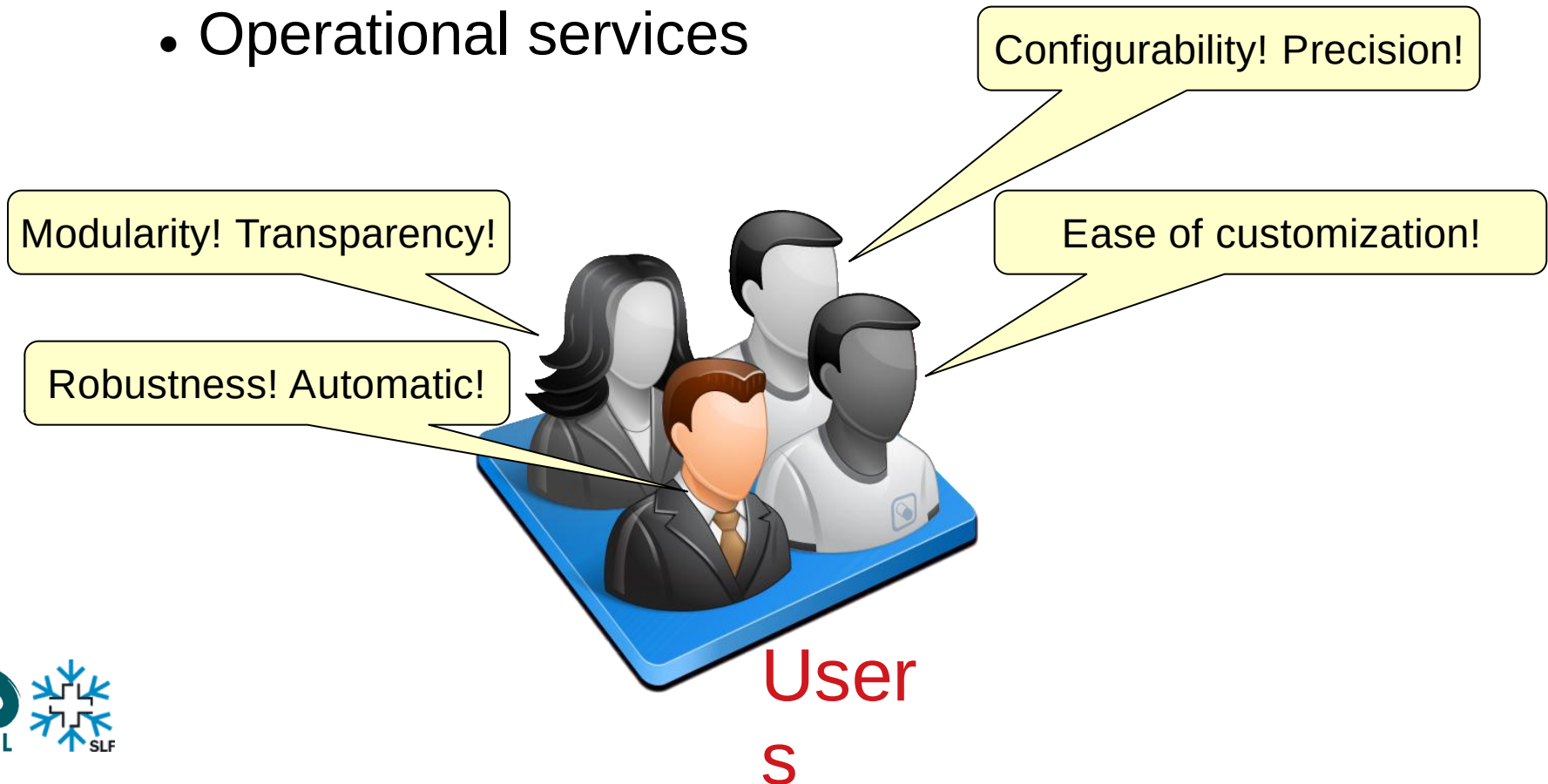
Time series - ../example.smet



At the beginning...

Various types of users:

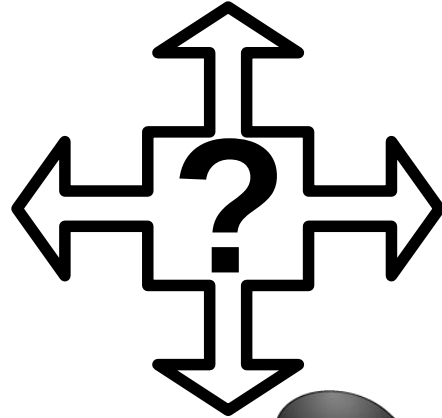
- Researcher
- Model developer
- Casual developer
- Operational services



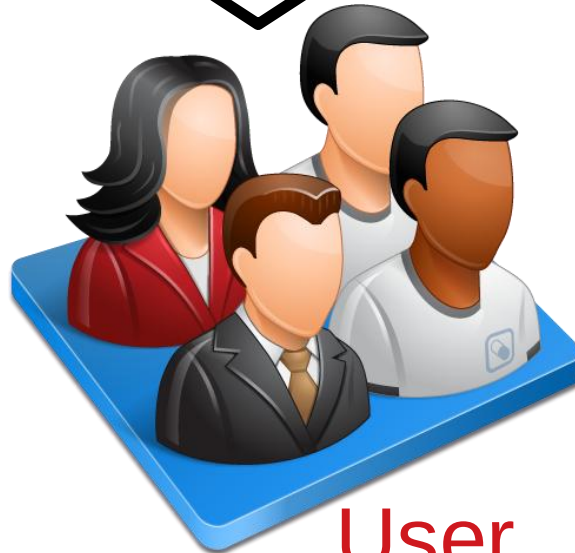
At the beginning...



Raw
Data

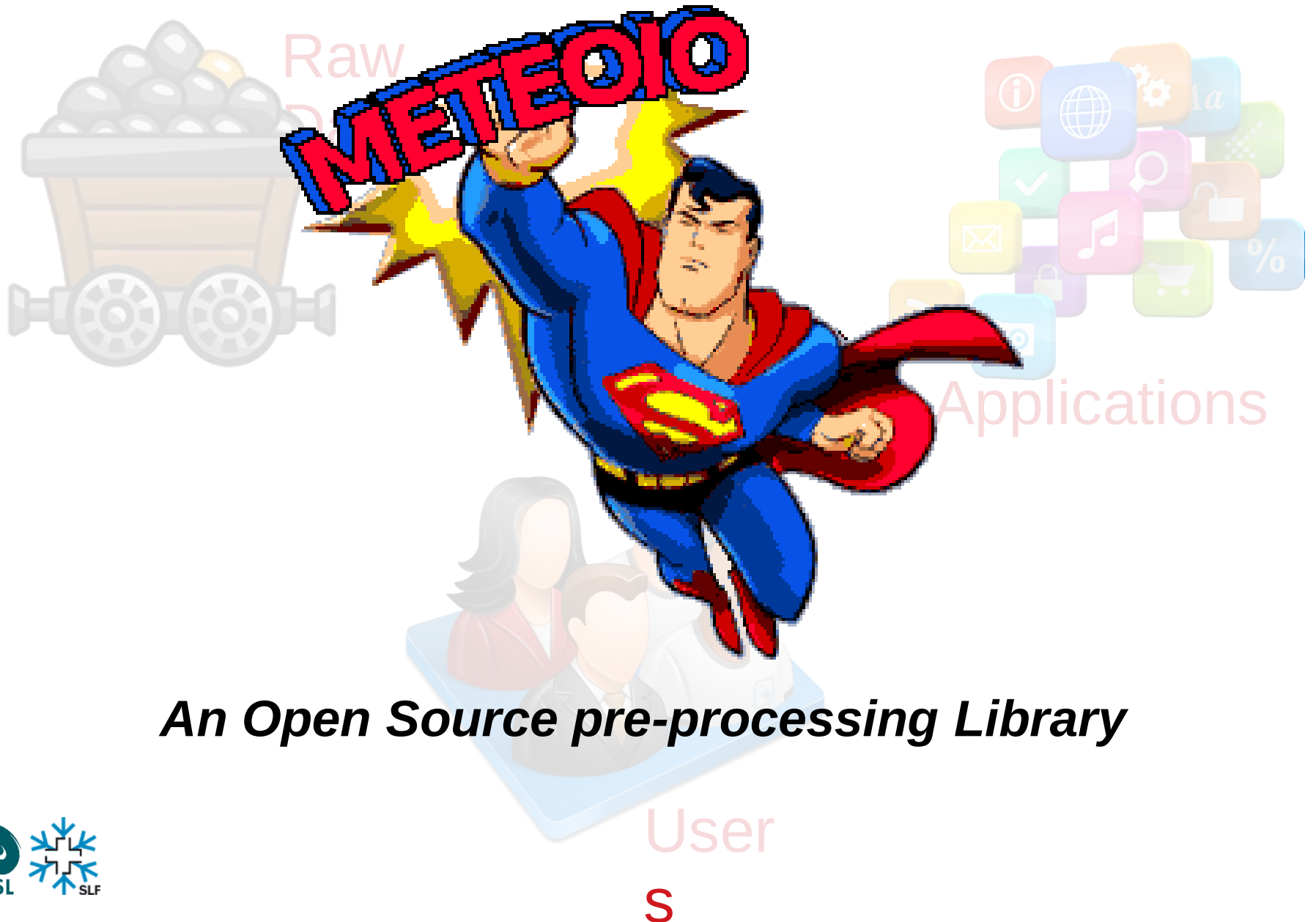


Applications



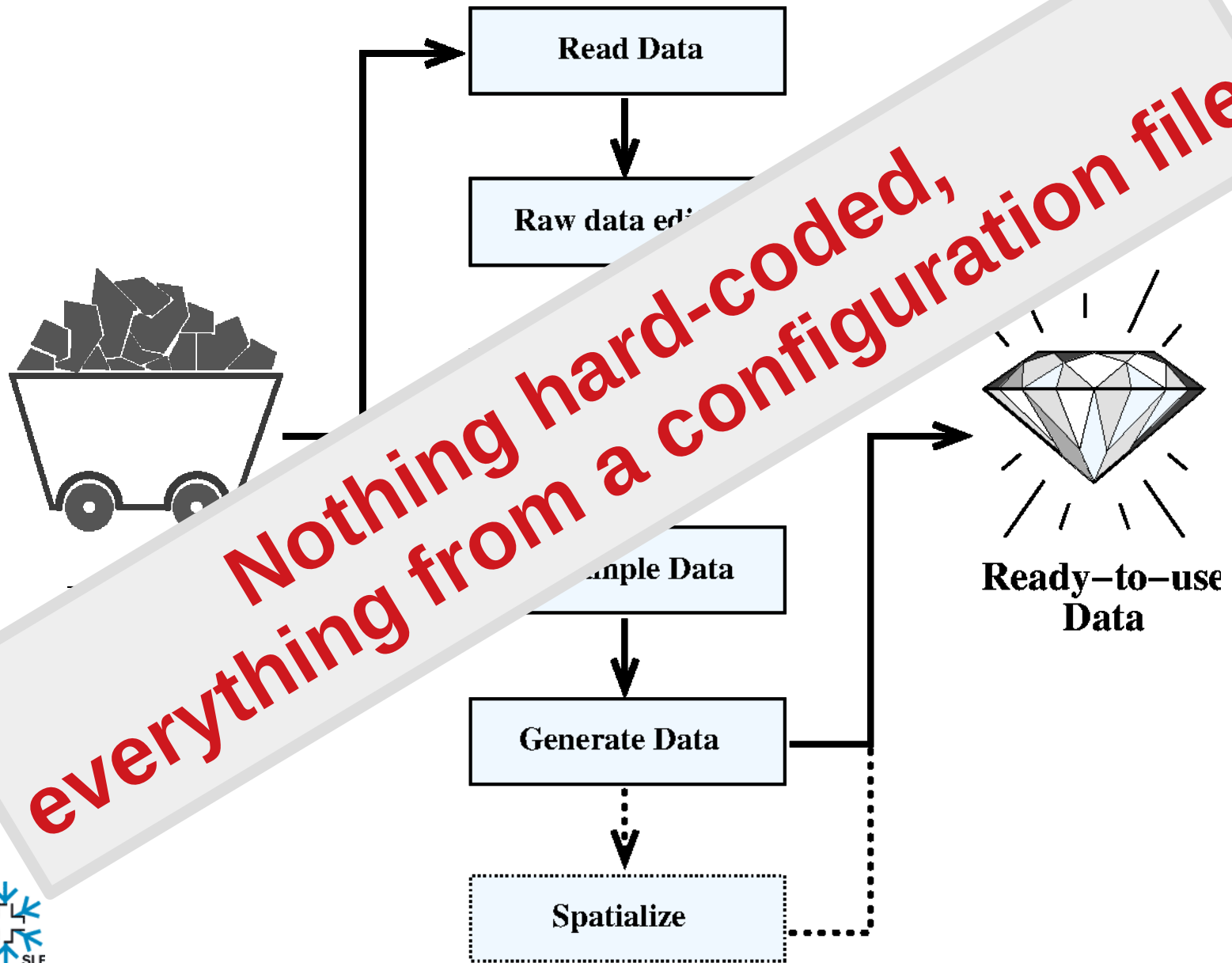
User
s

At the beginning...

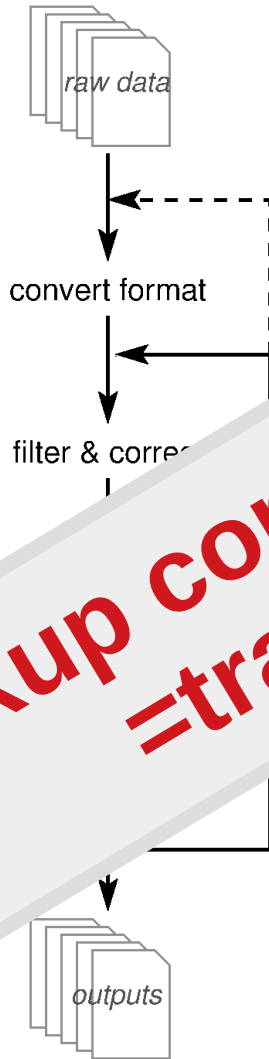


An Open Source pre-processing Library

MeteoIO's Workflow



User's workflow



**Backup configuration file
=traceability**



We now have a great toolbox sitting
between the raw data & the application...
Let's play!



In the following examples, no “external” scripts are used, only MeteorIO with some applications

Data quality monitoring

Enable the “DATA_QA” option, then request some data...

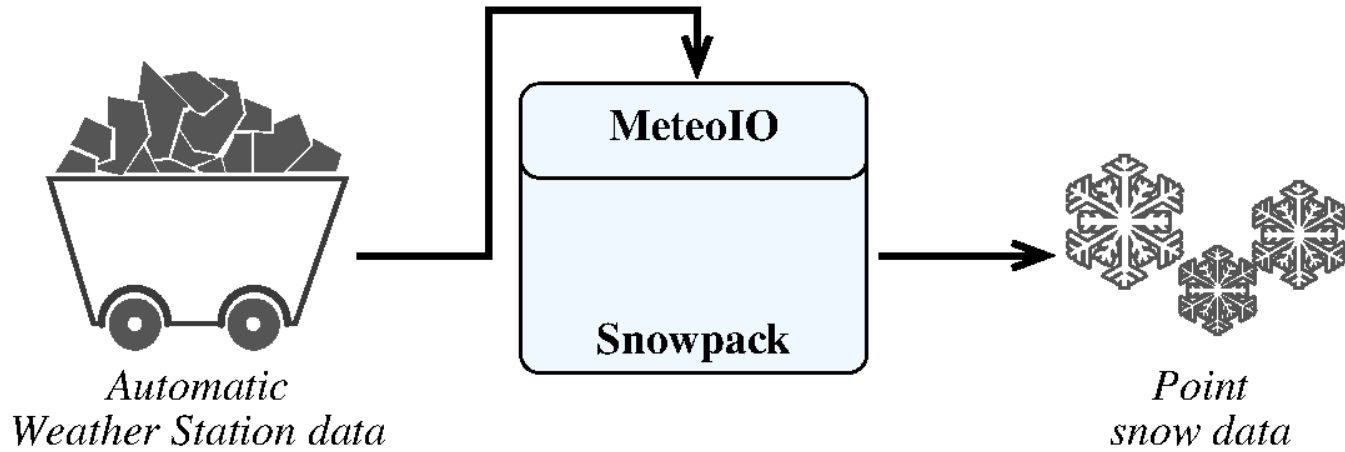
Filtering	AR02::HS::MAD	2018-09-30T22:00:00
Filtering	SIM2::HS::MIN	2018-09-25T14:00:00
Resampling	ALI2::RH::LINEAR	2018-09-27T06:00:00
Resampling	PAR2::RSWR::LINEAR	2018-09-27T06:00:00
Missing	DAV5::TSG	2018-09-27T06:00:00
Missing	ELS2::RH	2018-09-27T06:00:00

Data Quality Tool

Station	PSUM	RSWR	TA	TSG	SUM
BER3	128	128	128	128	512
SCH2	1			336	337
OBW3	47				47
DIA2					0
CAM3	0
ATT2					0

Ex.: IMIS Network

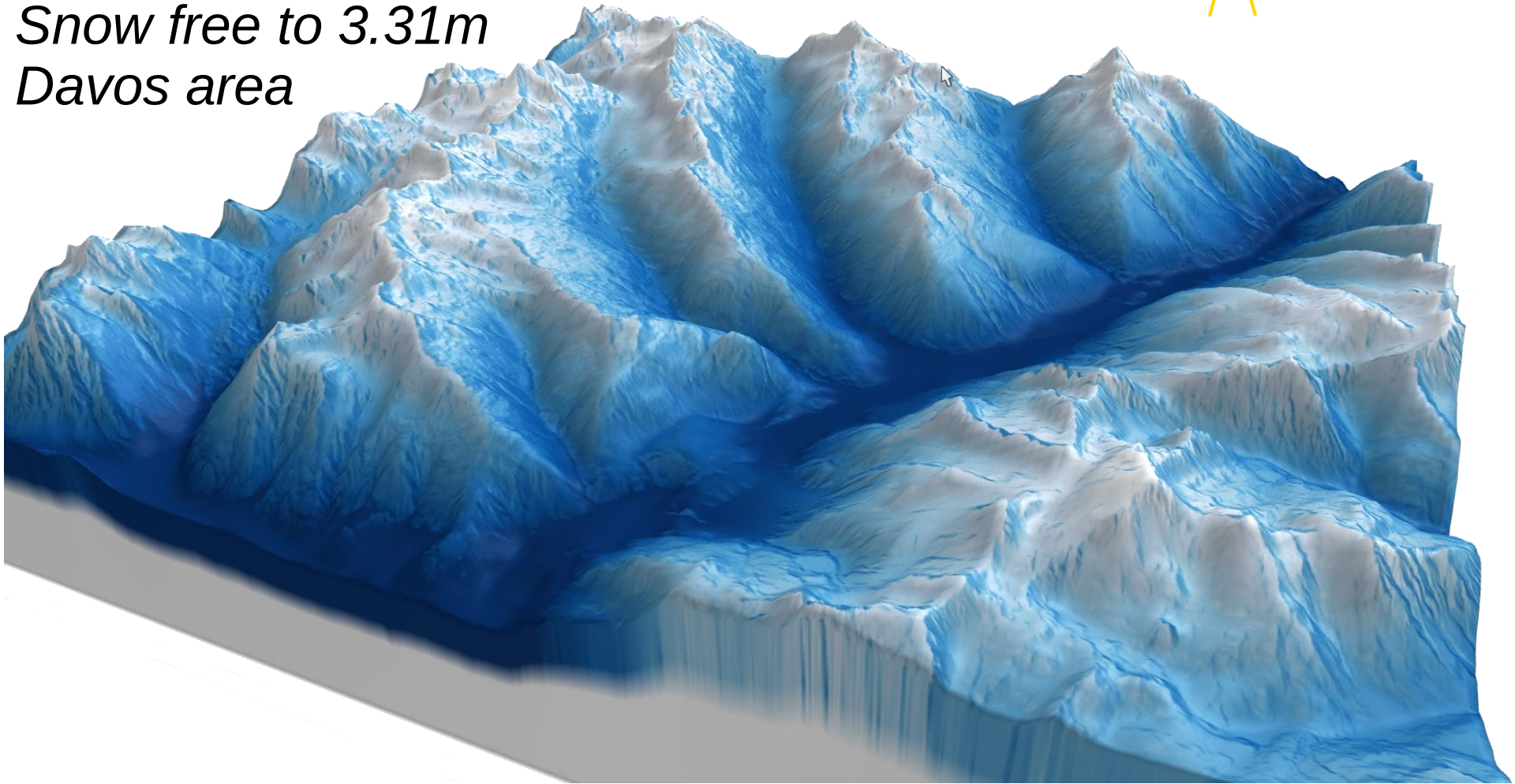
Model chains



Model chains

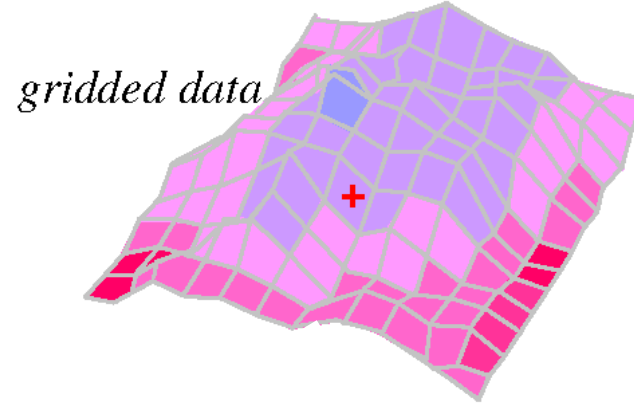
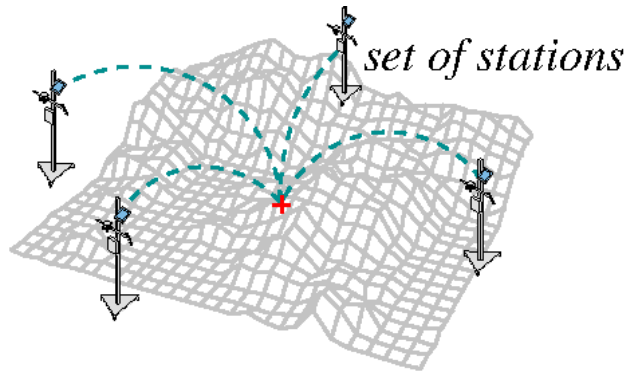


Snow Depth on 01.05
Snow free to 3.31m
Davos area



Ex.: operational snow maps??

Downscaling



Downscaling

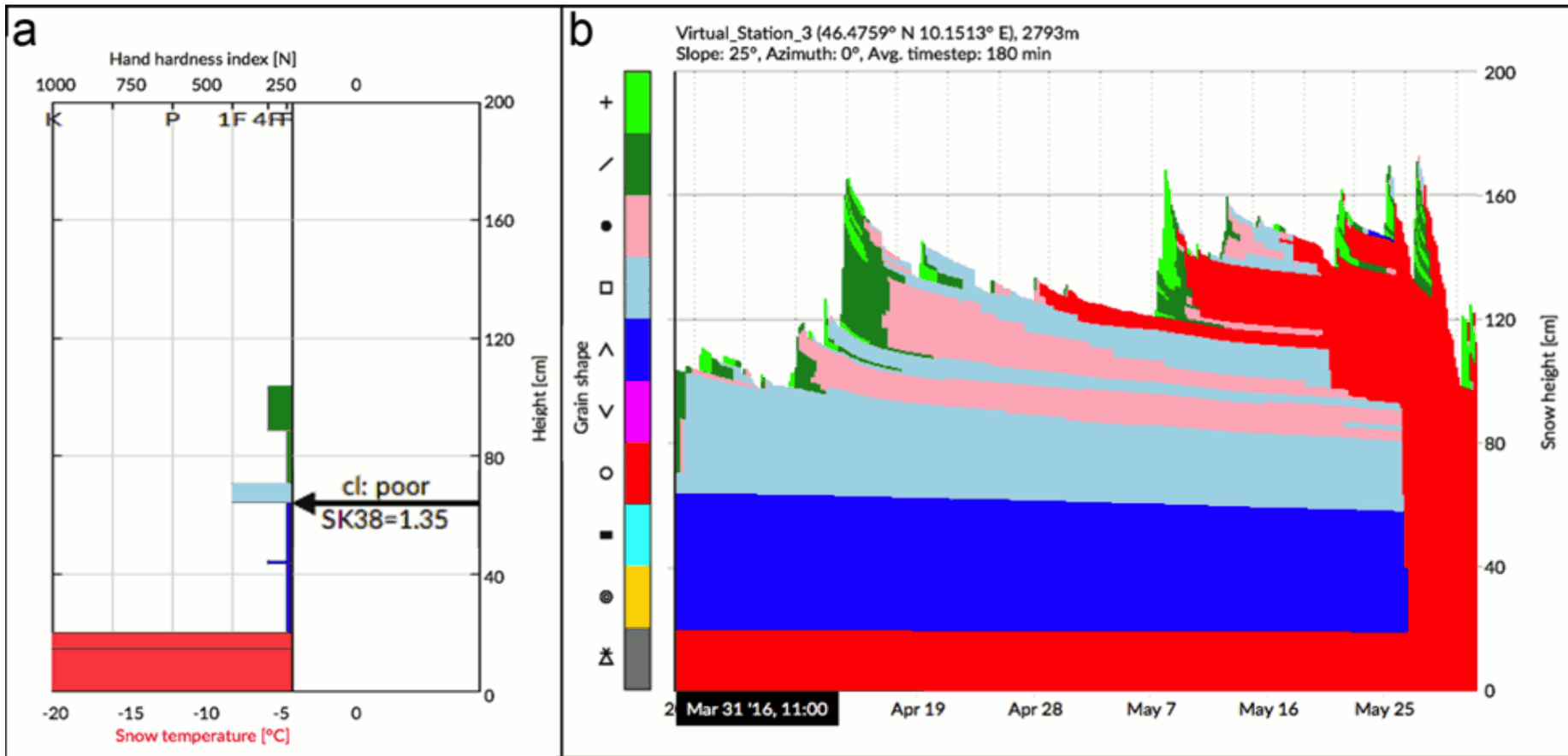


Fig. 2: a) Manual profile collected on Monte Vago on 18 March 2016. b) Manual profile evolution simulated by the snow cover model SNOWPACK forced with the weather data derived by MeteolO for the specific location. Symbols and colors are accordingly to Fierz et al. (2009).

Ex.: ISSW2016

In Depth:



Global Cryosphere Watch

Global Cryosphere Watch

- WMO effort to support all key cryospheric in-situ and remote sensing observations;
- started in 2011, should be fully operational by 2020;
- focus on standardization and sharing of data;
- The goal is for everybody to be able to contribute its data;
- Online data portal;

<http://get-cool-data.org> !?

Global Cryosphere Watch



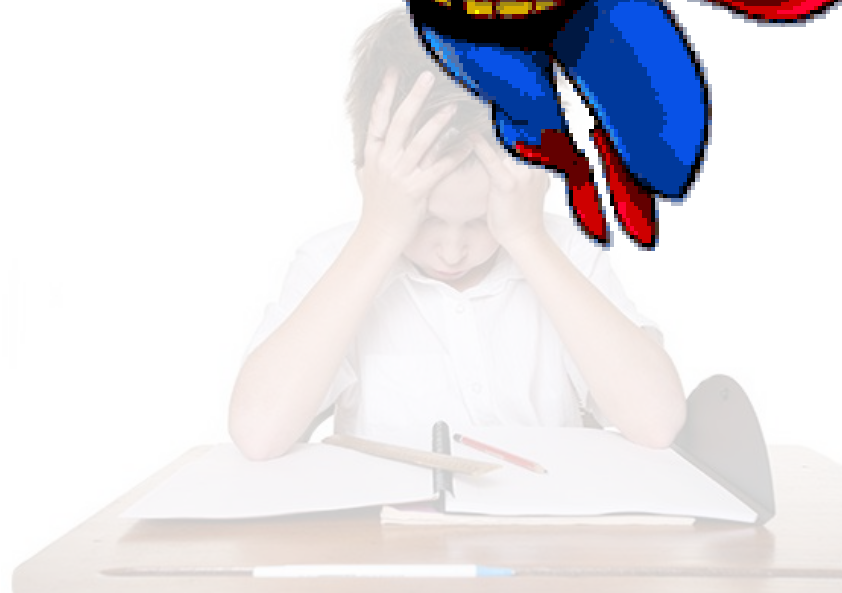
But...

- lots of stations managed by small institutions (ie. little manpower to invest);
- WMO requirements are complex to satisfy;
- preparing a quality dataset is *very* time consuming;
- how can small institutions contribute?

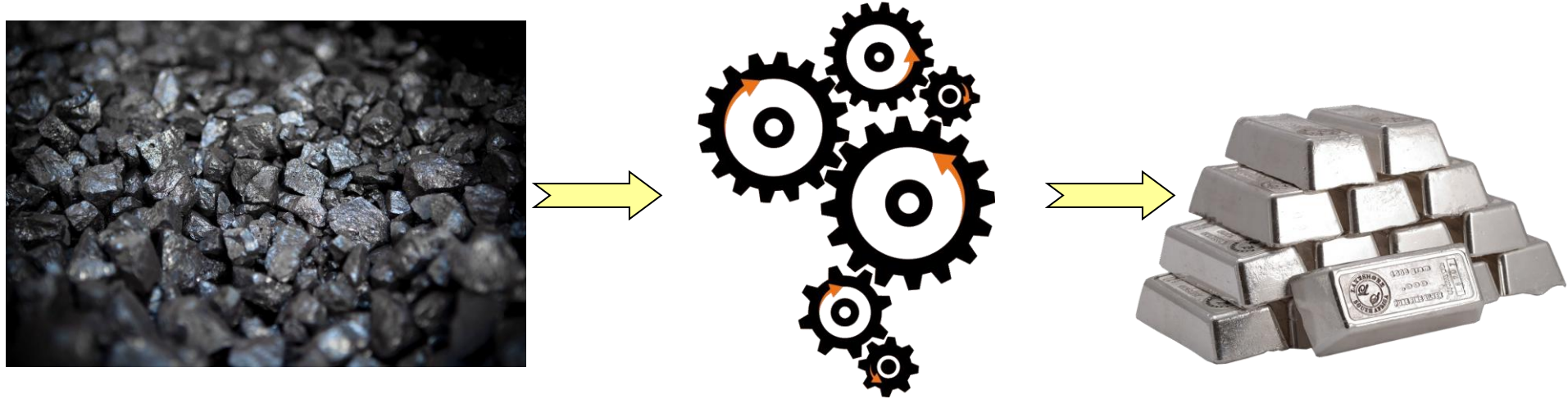


But...

- lots of stations, many small institutions (ie. little man)
- WMO requirements are complex to satisfy;
- preparing a quality report is very time consuming;
- how can small institutions cope?



MeteoIO for Data standardization



Raw data,
vast variety
of formats
and
protocols
(csv, db,
NetCDF...)

- Read with different plugins
- Process to standard levels
- Add search metadata
- Write with a common plugin

Standard
product

Standard product: NetCDF

- international standard of the Open Geospatial Consortium;
- for array-oriented scientific data

Dimensions, variables, attributes
NOT standardized

```
netcdf mynetcdf{
dimensions:
  X=4;
  Y=5;
  Time=UNLIMITED;
variables:
  float Y;
  float X;
}
```

Temperature =
100, 110, 120, 130, 140;
150, 160, 170, 180, 190;

Time = 1 to 3
Y = 1 to 4
X = 1 to 4

Standard product: CF-1

- Climate & Forecast metadata (based on CF-1)
- Standardize variable names, dimensions, and attributes (data origin, units, data type)
- Some missing elements for standardization, need to get them standardized

Standard product: ACDD

- Attribute Catalogue for Dataset Discovery (based on CF-1)
- Describe the dataset (where, when, sampling frequency, history...)

many claim to do NetCDF CF-1, nobody is even remotely doing it!

MeteoIO's smallest configuration

[Input]

```
COORDSYS      = UTM
COORDPARAM    = 32T
TIME_ZONE     = 1
```

Select the input plugin

Plugin options

```
METEO = CSV
METEOPATH = ./input/meteo
CSV_DELIMITER = ;
CSV_NR_HEADERS = 15
CSV_COLUMNS_HEADERS = 2
CSV_DATETIME_SPEC = "YYYY-MM-DD HH24:MI:SS"

STATION1 = HW516.csv
POSITION1 = latlon (46.8,9.80,2200)
CSV1_FIELDS = DATETIME ILWR PSUM
```

[Output]

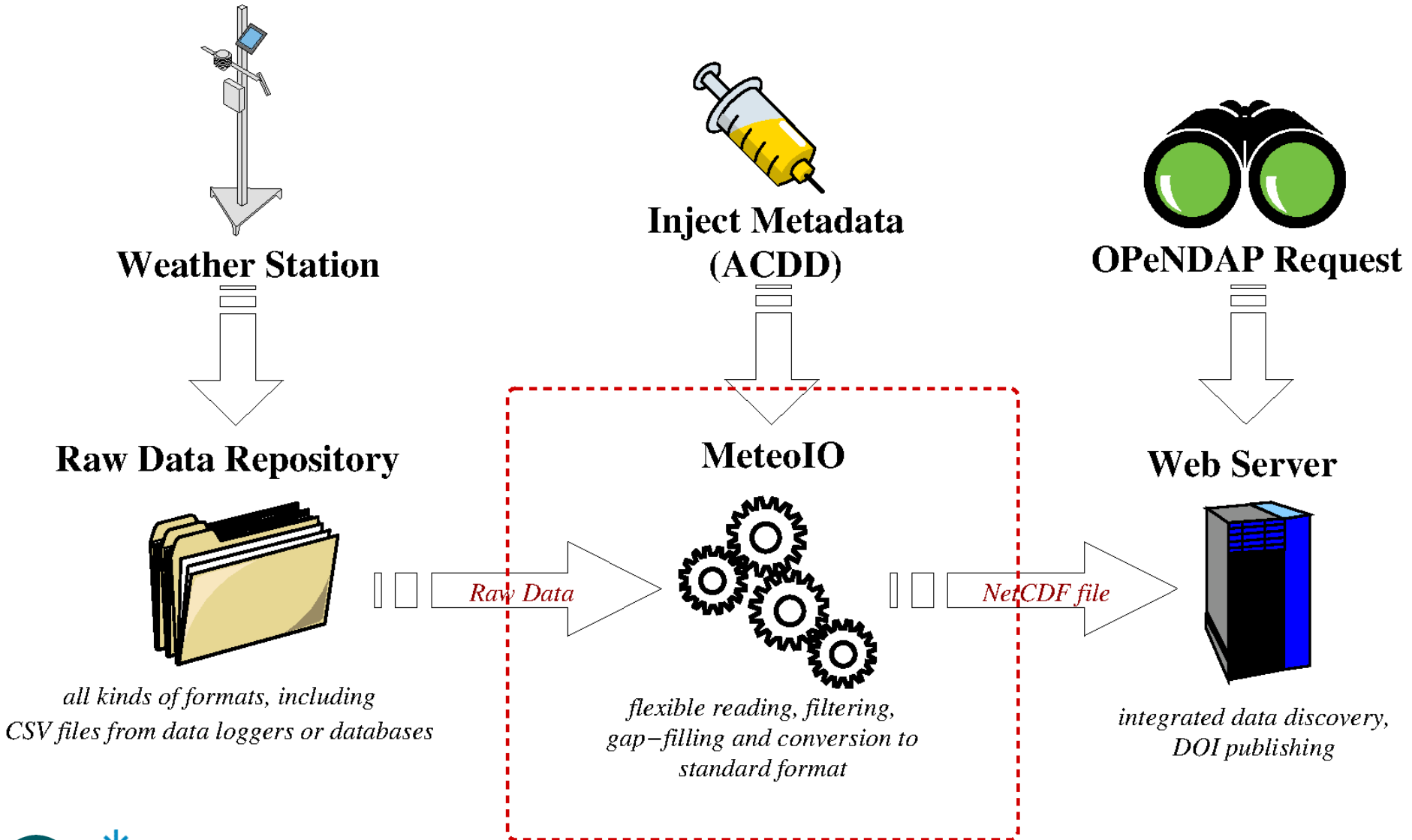
```
COORDSYS      = UTM
COORDPARAM    = 32T
TIME_ZONE     = 1
```

Selecting NetCDF CF-1.6

Injecting extra ACDD
metadata

```
METEO = NETCDF
METEOPATH = ./output
NETCDF_SCHEMA = CF-1.6
NC_CREATOR = Mathias Bavay
NC_LICENSE = CC-BY-NC
```

Conclusion: GCW's toolchain



Conclusion

MeteoIO is a flexible, powerful & robust pre-processor

It makes the creation of standardized datasets easy

Original paper in GMD:

Bavay, M., and T. Egger. *"MeteoIO 2.4. 2: a preprocessing library for meteorological data."* Geoscientific Model Development (2014).

Get MeteoIO at <http://model.slf.ch>!

