Machine readable semantic information about observed environments

Markus Stocker
PANGAEA Data Publisher in Earth and Environmental Science
MARUM Center for Marine Environmental Sciences
University of Bremen

@envinf
This years’ motto

“Geospatial Sensor Webs - from data to information in the Sensor Web”
From data to information: A rather common notion ...
FixO3

Provide access to data products as well as derived information
JERICO-NEXT

Delivery of high quality environmental data and information products
ICOS
INTEGRATED CARBON OBSERVATION SYSTEM

Knowledge through observations
This year’s motto … clarified

… from observational data to information about the observed environment …
In other words

Semantic information about the observed environment derived from primary observational data by research communities
Semantic information

- Consists of data
- Data must be well-formed
- The well-formed data must be meaningful
- Meaningful well-formed data must be truthful
- False information is not a kind of information

Use case

Aerosol science for the study of atmospheric new particle formation
734544 0 0 0 0
734545 0 0 0 1
734546 0 1 0 0
734547 0 0 0 1
734548 0 0 0 1
734549 0 0 0 1
734550 0 0 0 1
734551 0 0 0 1
734552 0 0 1 0
MATLAB datenum

Class lb event
Record data and meaning gained in interpretation
from smear.datafetcher import fetchdata
from smear.dataplotter import plotdata

# Fetch and plot concentration data for the given time and location
# from SmartSMEAR, https://avaa.tdata.fi/web/smart
plotdata(fetchdata('2013-04-04', 'Hyytiala'))
from smear.datafetcher import fetchdata
from factory import assess

# Automated assessment for whether or not an event occurred
assess(fetchdata('2013-04-04', 'Hyytiälä'))

['Event']

from factory import record, event

# Record information about the new particle formation event
record(event('2013-04-04', 'Hyytiälä', '11:00', '19:00', 'Class Ia'))
from processing.visualization import imap
from factory import events

# Visualize event locations on an interactive map
imap(events())
from processing.statistics import duration
from factory import events

# Compute the average duration of events, possibly at specific locations
duration(events(), fun='avg')

7:00:00

from processing.description import describe
from factory import events

# Describe an event in plain English text
describe(events(place='Hyytiälä'), format='text')

A Class Ia event occurred at Hyytiälä (FI) [http://www.geonames.org/656888/hyytiaelae.html] on 2013-04-04 starting at 11:00 and ending at 19:00.
from factory import events
from processing.description import describe

# Describe an event with information in machine readable format
describe(events('2013-04-04', 'Hyytiälä'), format='rdf')
Environment Ontology

- Contributed new concept to recent release
- "formation of particles in an atmosphere"
  - a particle formation process
  - that occurs in an atmosphere
  - with aerosol as output
- Planning to adopt this concept in the use case
Research Data Alliance

- IG “From Observational Data to Information”
- rd-alliance.org/groups/observational-data-information
- Session at Montréal plenary, Breakout 5
- Monthly conference calls
Take aways

- Data interpretation results in semantic information
- Truthful and meaningful well-formed data
- Acquire semantic information in infrastructures
- Integrate with tools researchers use