SCIENTIFIC KNOWLEDGE INFRASTRUCTURES

INTEGRATING RESEARCHERS, MACHINE LEARNING AND SEMANTIC TECHNOLOGIES

Scientists use research infrastructures to generate scientific knowledge. Fundamental is the interpretation of observational data for their meaning in the context of research investigations. Uninterpreted data are thereby evolved into information, meaningful data. By integrating machine learning and semantic technologies, research infrastructures can support the extraction of information and the curation of meaning.

// KNOWLEDGE INFRASTRUCTURES

Robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds.

Paul Edwards (2010). A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming, p. 17

// RESEARCH INFRASTRUCTURES

Facilities, resources and services that are used by the research communities to conduct research and foster innovation in their fields.

REGULATION (EU) No 1291/2013, Article 2 - Definitions (6), p. L 347/109

// USE CASE

Environmental science, specifically aerosol science and the investigation of atmospheric new particle formation (NPF) events using particle size distribution observational data. Central to NPF events is the formation of aerosol particles at specific spatio-temporal locations and the growth of particle diameter size over the course of a few hours. NPF events are studied for their relevance in climate science and human health.

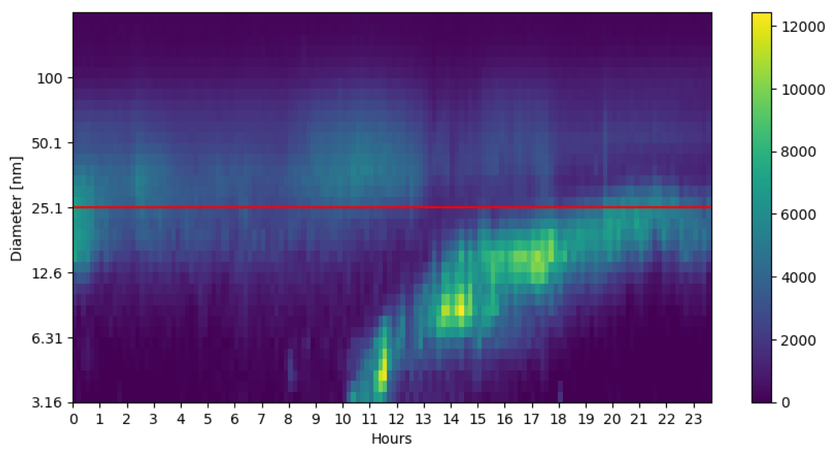
In [1]:

```
from pynpf.smear.datafetcher import fetchdata
from pynpf.smear.dataplotter import plotdata
from pynpf.factory import record, event, assess

day = '2013-04-04'
place = 'Hyytiälä'
```

In [2]:

```
data = fetchdata(day, place)
plotdata(data)
```



In [3]:

```
assess(data)
```

```
['Event']
```

In [4]:

```
beginning = '11:00'
end = '18:00'
classification = 'Class Ia'
```

In [ ]:

```
record(event(day, place, beginning, end, classification))
```

Fig. 1. Fetching and visualizing particle size distribution data for April 4, 2013 at Hyytiälä, Finland. Automated detection of the occurrence of an event using a MLP ANN classifier. Recording information about the event.

```
[] a lode:Event ;
  smear:hasClassification [
    rdfs:label "Class Ia"^^xsd:string ;
    rdfs:comment "Very clear and strong event"^^xsd:string
  ] ;
  lode:atPlace [
    a gn:Feature, DUL:Place ;
    gn:countryCode "FI"^^xsd:string ;
    gn:locationMap <http://www.geonames.org/656888/hyytiaelae.html> ;
    gn:name "Hyytiälä"^^xsd:string
  ] ;
  lode:inSpace [
    a sf:Point, wgs84:SpatialThing ;
    geosparql:asWKT "POINT (24.29077 61.84562)"^^geosparql:wktLiteral ;
  ] ;
  lode:atTime [
   a time:Interval ;
    time:hasBeginning [
      a time:Instant ;
      time:inXSDDateTime "2013-04-04T11:00:00+03:00"^^xsd:dateTime
    ] ;
    time:hasEnd [
      a time:Instant ;
      time:inXSDDateTime "2013-04-04T18:00:00+03:00"^^xsd:dateTime
    ]
  ] .
```

Fig. 2. Semantic machine-readable description resulting from recording information about a NPF event, here for the event that occurred on April 4, 2013 at Hyytiälä, Finland.

Dr. Markus Stocker
Head of Knowledge Infrastructures Research Group
Lange Laube 30
30159 Hannover
markus.stocker@tib.eu
@envinf

ENVRI plus

TIB LEIBNIZ-INFORMATIONSZENTRUM TECHNIK UND NATURWISSENSCHAFTEN UNIVERSITÄTSBIBLIOTHEK