Open Meteorological Data with OGC and INSPIRE

INSPIRE 2014 Aalborg
Finnish Meteorological Institute
Roope Tervo, Mikko Visa, Mikko Rauhala, Tarja Riihisaari
FMI Open Data

Finnish Meteorological Institute opened its data in 2013.

Basically everything FMI owns was opened.

Data is provided in freely in machine readable format.

https://en.ilmatieteenlaitos.fi/open-data
FMI Open Data Portal

FMI Open Data Portal follows INSPIRE requirements.

The very same data portal works as Open Data and INSPIRE portal.
## Example of Data Sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Description</th>
<th>Time Interval</th>
<th>Estimated publish date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weather Observations</strong></td>
<td>Temperature, Wind, Humidity, Ground Temperature…</td>
<td>10 min</td>
<td>Open, older data to be added</td>
</tr>
<tr>
<td><strong>Sun Radiation</strong></td>
<td>UV, Short and Long Term Radiation…</td>
<td>1 min</td>
<td>Open</td>
</tr>
<tr>
<td><strong>Marine Observations</strong></td>
<td>Waves, Sea Temperature, Sea Level…</td>
<td>1 h</td>
<td>Open</td>
</tr>
<tr>
<td><strong>Weather Radars</strong></td>
<td>Precipitation Rate, Precipitation Amount…</td>
<td>5 min</td>
<td>Open, older data to be added</td>
</tr>
<tr>
<td><strong>Lightning</strong></td>
<td>Thunder Strikes in Finland</td>
<td>5 min</td>
<td>Open</td>
</tr>
</tbody>
</table>
## Example of Data Sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Description</th>
<th>Time Interval</th>
<th>Estimated publish date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Time Observations</td>
<td>Real Time Observations from specific location(s)</td>
<td>AWS 2010 – Soundings 1959 – Flashes 1998 – Sea Level 1971 – Waves 2005 –</td>
<td>Open older data will be added</td>
</tr>
<tr>
<td>Climatological Observations</td>
<td>Dayly and monthly temperature mean and extreme values from weather stations</td>
<td>1959 -</td>
<td>Open</td>
</tr>
<tr>
<td>Climatological Observations</td>
<td>Monthly temperature and precipitation rate mean values interpolated to grid</td>
<td>1961 -</td>
<td>Open</td>
</tr>
</tbody>
</table>
# Example of Data Sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Description</th>
<th>Time Interval</th>
<th>Estimated publish date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather forecast model HIRLAM RCR</td>
<td>Point forecasts and grid data</td>
<td>Latest model run (4 times a day) 0…54 h</td>
<td>Open</td>
</tr>
<tr>
<td>Sea forecast models</td>
<td>Sea level point forecasts, Wave (WAM) and current (HBM) as grid data</td>
<td>Latest model run (4 times a day) 0…54 h</td>
<td>Open</td>
</tr>
<tr>
<td>Environmental Monitoring Facilities</td>
<td>Weather observation stations, radars…</td>
<td></td>
<td>2014</td>
</tr>
<tr>
<td>Aviation Observations</td>
<td>METAR</td>
<td>30 min</td>
<td>2014</td>
</tr>
<tr>
<td>Ground &amp; mast observations</td>
<td>Special observations from ground and masts</td>
<td></td>
<td>2014</td>
</tr>
</tbody>
</table>
# Example of Data Sets

<table>
<thead>
<tr>
<th>Data set</th>
<th>Description</th>
<th>Time Interval</th>
<th>Estimated publish date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Quality Observations</td>
<td>Air Quality Observations</td>
<td>1h</td>
<td>2015-2016</td>
</tr>
<tr>
<td>Silam Model</td>
<td>Dispersion Model for Air Quality, Forest Fire and Pollen</td>
<td>Latest model run (once a day) 0…96h</td>
<td>2014-2015</td>
</tr>
<tr>
<td>HELMI Ice Model</td>
<td>Ice forecast model</td>
<td>Latest model run (4 times a day) 0...54 h</td>
<td>2014</td>
</tr>
<tr>
<td>Soundings</td>
<td>Temperature, Humidity, Pressure, Wind from ground to 25 km height</td>
<td>2 times a day</td>
<td>2014</td>
</tr>
</tbody>
</table>
View Service (WMS)
- Based on GeoServer
- Only the most common layers published

Catalog Service (CSW)
- Based on GeoNetwork
Download Service (WFS 2.0)

- Web Feature Service (WFS) 2.0 Simple Profile
- Based on stored queries
  - Predefined data sets with possibility for additional parameters (i.e. time and area)
- In-house production
Data Models

- Observations and point forecasts as GML
  - The same data is published in two different formats.
- Gridded data is provided in appropriate binary format (grib, NetCDF, HDF…)
  - WFS members contains the metadata ‘envelope’ with a link to a actual data
Data Models

gmlcov:MultiPointCoverage

gml:rangeSet
  gml:doubleOrNilReasonTupleList

The data is listed for every point defined in domain set.

---

gml:domainSet
gmlcov:simpleMultiPoint

The coverage is defined as a list of points in 4 dimensional grid (lat, lon, height, time).

---

gmlcov:rangeType

The parameters listed in range set are defined in separate element.
Data Models

Pros
+ Compact
+ Efficient
+ Small file size
+ Works for many data types

Cons
- Not intuitive
- No natural structure of XML
  → XSLT and Xpath don’t work

gmlcov:MultiPointCoverage
Data Models

One member contains time series from one parameter and one location

wml2:MeasurementTimeseries
Data Models

Pros
+ Intuitive
+ Easy to use
+ XSLT & XPath works

Cons
- Lots of repetition
- Large file size
- Heavy for DOM-based parsers
- Don’t work i.e. for thunder strikes
Data Models  File size Comparison

- SimpleFeature: 81.7 MB
- MeasurementTimeseries: 52.9 MB
- MultiPointCoverage: 1.8 MB

- Compressed Document Size
  - SimpleFeature: 1.3 MB
  - MeasurementTimeseries: 1.2 MB
  - MultiPointCoverage: 0.2 MB

- 138 observation stations
- 11 parameters
- 72 time steps (12 hours)
Data Models Popularity Comparison

- Measurement Time Series: 80%
- MultiPoint Coverage: 19.8%
- Binary Grid: 0.2%

Based on one month data (04/2014)

All data sets combined
# Data Models Popularity Comparison

<table>
<thead>
<tr>
<th></th>
<th>Downloads [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast</td>
<td>54</td>
</tr>
<tr>
<td>Weather Observations</td>
<td>37</td>
</tr>
<tr>
<td>Radar Images</td>
<td>9</td>
</tr>
</tbody>
</table>

- Based on one month data (04/2014)
- All data sets combined
- Note that some data is updated more often than others
MetoLib

- Open source JavaScript library produced by Finnish Meteorological Institute
- Helps users to load and use the data
- Supports multi point coverage data format
- Python version is on the list

Easy requests

Cache

Parse the data to as JSON

Please contribute!
https://github.com/fmidev/
Registration

- Registration is required to use View and Download Services
  - Working email address is the only mandatory information
- After registration the user gets an API key which have to be added into all requests
  - POST field fmi-apikey=…
  - GET parameter fmi-apikey=…&
  - Header fmi-apikey; …
  - Part of url http://wms.fmi.fi/fmi-apikey/…/wms?
- One can create several API keys with one email
Usage Limits

With one API key it’s allowed to

- do at most **20 000** requests **per day** to Download Service
- do at most **10 000** requests **per day** to View Service
- do at most **600** requests **per 5 minutes** to both services
- If all observations from one time step is calculated to as one, little over **17 000** new data sets are published daily
- So, with one API key it’s allowed load everything once
- View service can be used for testing but can not be used as a back end for popular clients
Some Experiences

At the moment a little over 5000 registered users

And a little over 100 000 data downloads per day (1.2 req/s)
Some Experiences

There's a gap between provided data model and clients' capabilities

Practically no client supports complex features
GeoServer is modified to support stored queries in WFS 2.0 (will be integrated into the standard package)

FMI is going to open the same data as simple features to support clients
Some Experiences

Amateur and freelancer coders would prefer simple JSON API

Industry is happy to use standardized services
Some Experiences

Quite many expected a user interface to load data to i.e. to Excel instead of machine readable interface.

So far maybe even more professional interest than private.
For now, very few have been interested in forecast models as a grid data type.

Point forecasts, observations and radar images are the most interesting data types.
Finally we have all the data behind one access point behind standard interfaces and in harmonized formats.

The Open Data project has been a great boost for ourselves both technically and for working culture.