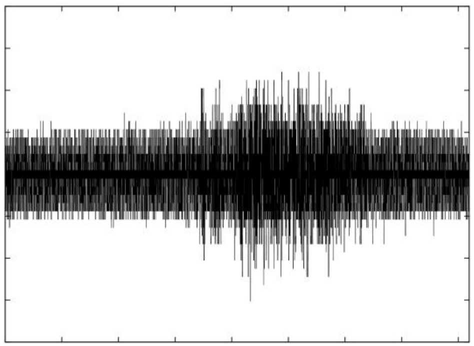
Markus Stocker
Aerosol Physics Group • University of Eastern Finland • March 21, 2013

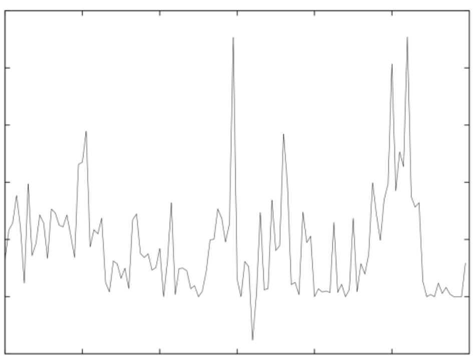
Wavellite

Continuous and automated representation of real-world knowledge acquired from sensor data

UNIVERSITY OF EASTERN FINLAND

Introduction
Sensor data makes *no* sense (almost)






```
  7.3316302167824074e+005  6.7295417637024639e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
1.1248345678061637e+003  7.6451655783892170e+002  6.6489916883398178e+002  5.0996188060151417e+002  5.3037337771849309e+002
2.5990152166875021e+002  2.6354791501170808e+002  2.6316883341227492e+002  3.3696093295963192e+002  3.2708190669952091e+002
3.9391899945110441e+002  3.2426846166900322e+002  2.0141200942601449e+002  1.4700101128227615e+002  4.0928587811142940e+001
1.2831620686413419e+001  0.0000000000000000e+000
  7.3316303024305555e+005  5.9954073981185900e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
8.1548383245909952e+002  6.4755047038178782e+002  4.7342333956288417e+002  5.0497438804107304e+002  4.8678378833469156e+002
2.1558831592426984e+002  2.4469781600550951e+002  2.4637216889127018e+002  3.9338495014282574e+002  2.8801709503956096e+002
4.1110643983031497e+002  2.8396322218576239e+002  2.1375040618051182e+002  1.4957123471247169e+002  3.7975823102376701e+001
2.0382122539321763e+001  1.6943213617983506e+001
  7.3316303880787035e+005  6.4135476612374202e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
7.9725650443997608e+002  4.0630047999881816e+002  3.5247269226184869e+002  3.7902085695621594e+002  2.8805461752290518e+002
1.8132905435068977e+002  2.6028214884093450e+002  1.6268304363454374e+002  3.3463583010398571e+002  2.0169500653114130e+002
4.6380154662366067e+002  2.6412454349776425e+002  2.0127541006127689e+002  1.3503244293786514e+002  2.3154902150570045e+001
7.6393422067327084e+000  1.7201363075124494e+001
  7.3316304687500000e+005  6.1036574617458143e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
7.2736971215268670e+002  3.9684227323502876e+002  3.4903661050199901e+002  3.4531808153726166e+002  3.0686816846877008e+002
1.5193881810750614e+002  2.4133238139702240e+002  2.4228704440975204e+002  2.8825745672488563e+002  2.0412591958392611e+002
4.0962054312850012e+002  2.8308360749792962e+002  2.2063151348349507e+002  1.2405735268505646e+002  1.0751097756079318e+001
1.1881689141194544e+001  1.3547412000218912e+001
  7.3316305520833330e+005  6.4313854345743607e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
8.0539618874151654e+002  3.8626777986640963e+002  2.8741638600588004e+002  2.5488895887853258e+002  2.6289312410222209e+002
1.6973869269222212e+002  1.7064856680396792e+002  2.6528938061547774e+002  2.7436008754172127e+002  2.3416327840132379e+002
4.5037513571727550e+002  3.4357911665879971e+002  2.1968465764114876e+002  1.0707033953281511e+002  4.2106918027684735e+001
0.0000000000000000e+000  2.5075705754016887e+001
  7.3316306354166672e+005  6.9018149795028319e+002  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000  0.0000000000000000e+000
1.0122740400440272e+003  4.7588170401009631e+002  3.6163701941444793e+002  3.8492799229831303e+002  3.1633325134373962e+002
1.8305150227202395e+002  2.2529123956614126e+002  2.0836586101380522e+002  3.1634735061210313e+002  2.5690547494848136e+002
3.6567505629777969e+002  3.0461779733996269e+002  1.8490080930780519e+002  1.1544775603278390e+002  4.1043905610232025e+000
1.3740708174213207e+001  2.9403296138810457e+001
```

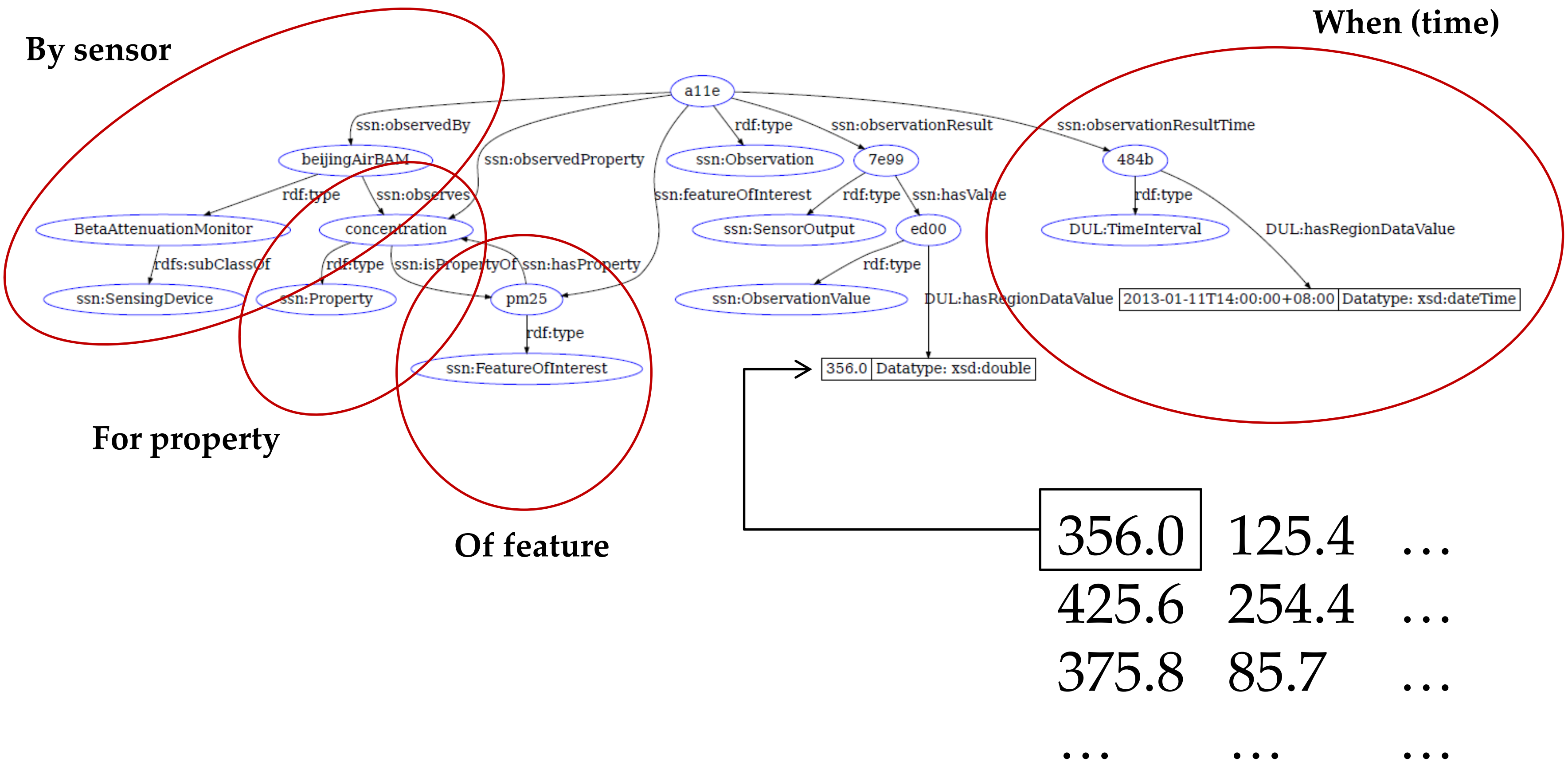
Introduction
… and

- There is lots of it, including in environmental science
- Typically automatically sampled, sometimes at "high frequency"
- More sensors; more data
- More information/knowledge?
- Probably, if we continue developing and improving
  - Management: processing, storage, and retrieval of sensor data
  - Acquisition: information and knowledge from sensor data
  - Integration, reuse, and sharing of data, information, and knowledge

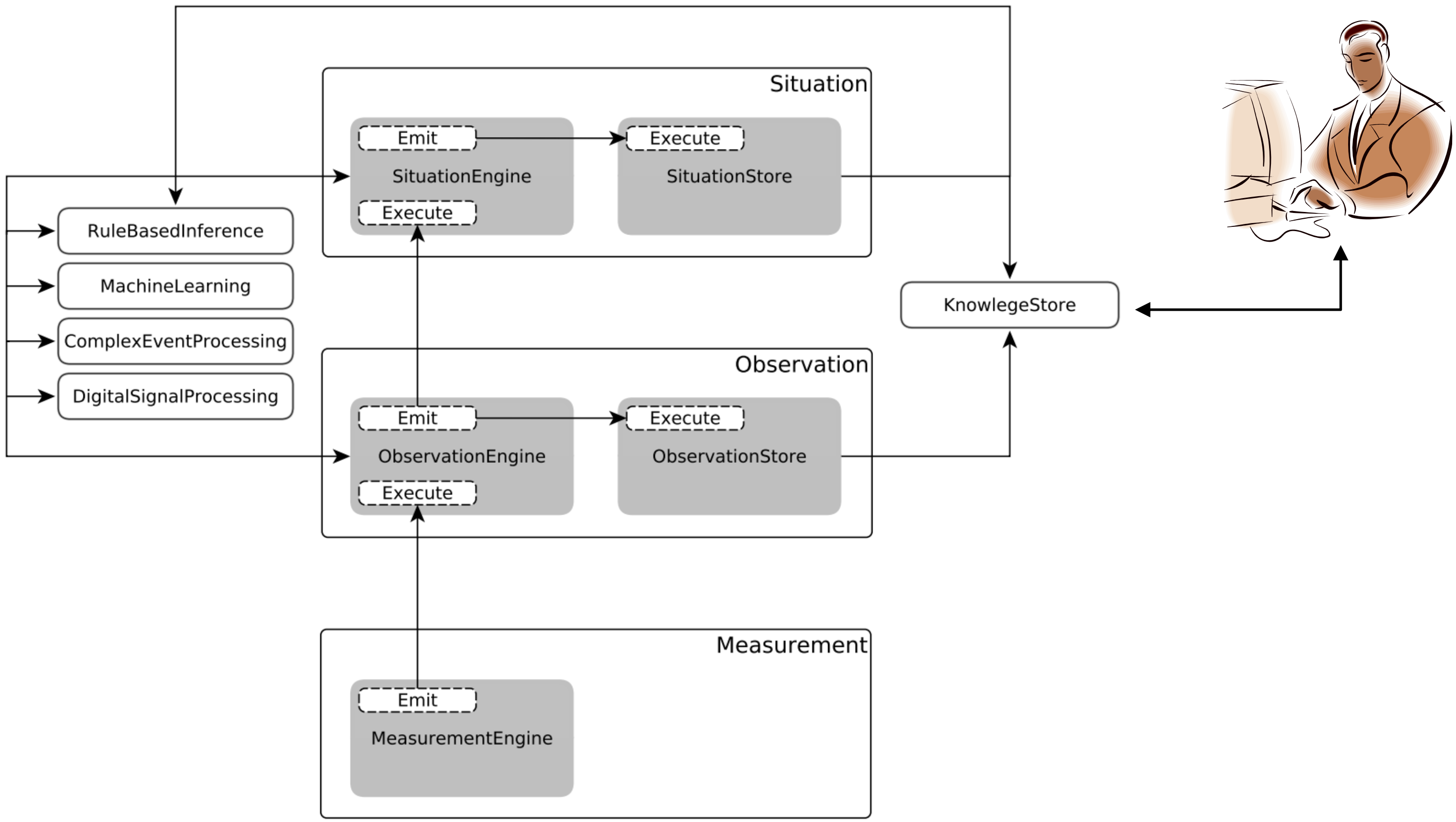
Introduction
Wavellite

- Manage sensor data *and* its metadata
  - Sensor data *and* the when-by-for-of (self-describing sensor data)
  - In other words, sensor data *and* information on
    - when is the sampling time
    - the sensor by which sampling occurs
    - the property for which measurement is performed
    - and the feature of which the property is measured
  - For instance, data sampled now by this beta attenuation monitor for the concentration of outdoor $PM_{2.5}$
- Represent knowledge for real-world phenomena observed by sensors
  - Events, episodes, scenes, changes: more generally situations
  - For instance, episodes of unhealthy exposure to $PM_{2.5}$
- Interact with represented observations and situations

Introduction
Self-describing sensor data



Materials and methods
Architecture



Materials and methods
Implementation (back end, server side)

- Continuous and automated
  - Sensor data acquisition and processing
  - Real-time knowledge acquisition and representation
- Computational intelligence for knowledge acquisition
  - ML, CEP, DSP
- Ontology for knowledge representation
  - Generic vocabulary
    - Used for all observations and situations
    - Shared by all software system components

Materials and methods
Implementation (back end, server side)

- Framework may be deployed on a single computer
  - Or on a cluster, for distributed processing (scalability)
- Domain-specific extensions
  - The framework implements generic functionality
  - Extend it to accommodate domain-specific requirements
    - Measurement engine implementation to acquire data from sensors
    - Situation engine implementation for knowledge acquisition tasks
- Use of
  - Semantic web technologies (RDF, OWL, SPARQL, knowledge store)
  - Storm as distributed real-time computation system
- Written in Java

Materials and methods
Implementation (front end, client side)

- Web enabled graphical user interface
  - Interaction between users and represented knowledge
  - Retrieve observations
    - Made by a sensor for a property of a feature within a time interval
    - Show results in tabular form, time-series plot, heat map, …
    - Show summary statistics (e.g. mean, sd, max, min, …)
    - Export data to various formats (e.g. Matlab, WEKA, …)
  - Browse situations along spatio-temporal dimensions
    - For instance on a map or timeline
- REST
  - Interaction between software and represented knowledge
  - HTTP requests and accept of various MIME media types

Example BeijingAir

- Sensor data acquired from @bejingair Twitter stream
- U.S. Embassy to Beijing runs a BAM for $PM_{2.5}$ monitoring
- One status update every hour, e.g.

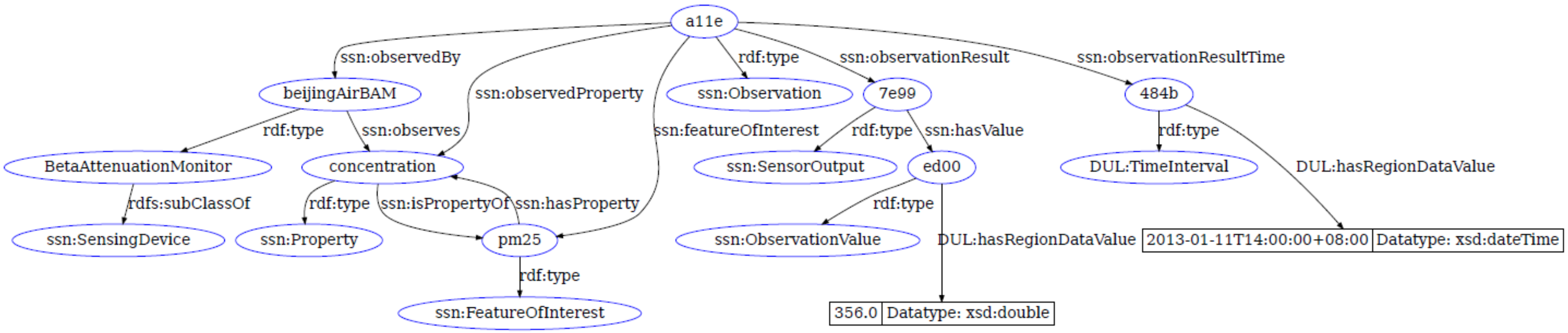
BeijingAir @BeijingAir 1h
04-02-2013 19:00; PM2.5; 108.0; 176; Unhealthy (at 24-hour exposure at this level)
Expand

BeijingAir @BeijingAir 2h
04-02-2013 18:00; PM2.5; 123.0; 184; Unhealthy (at 24-hour exposure at this level)
Expand

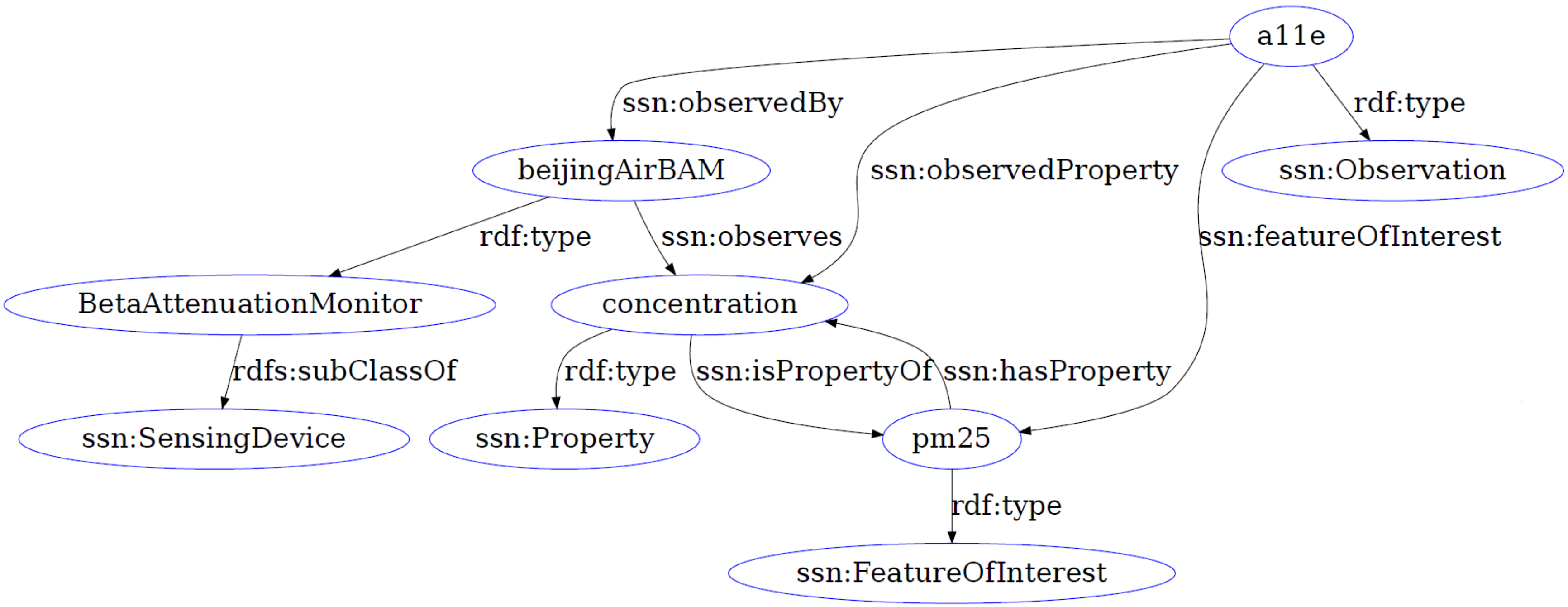
BeijingAir @BeijingAir 3h
04-02-2013 17:00; PM2.5; 130.0; 188; Unhealthy (at 24-hour exposure at this level)
Expand

- Statuses -> Measurement -> Observations -> Situations
- Situations of unhealthy exposure (> 65.5 µg/m^3; 24-hour)

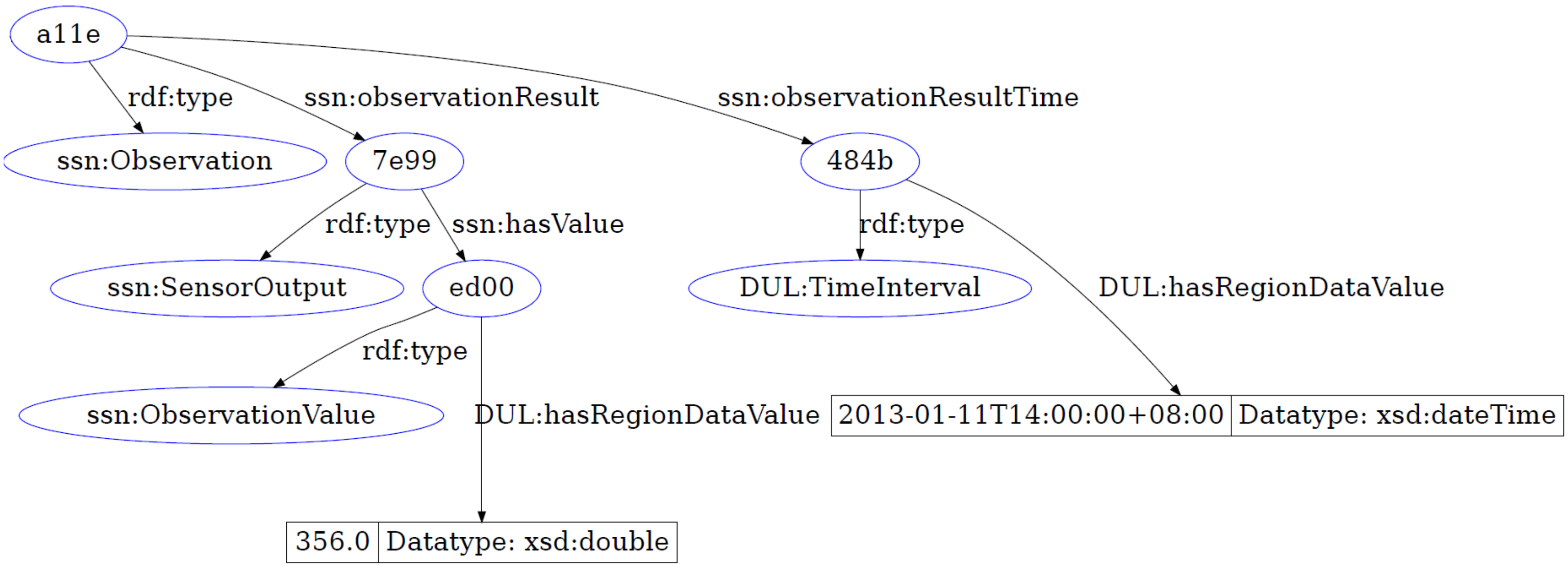
Example
BeijingAir (Observation)



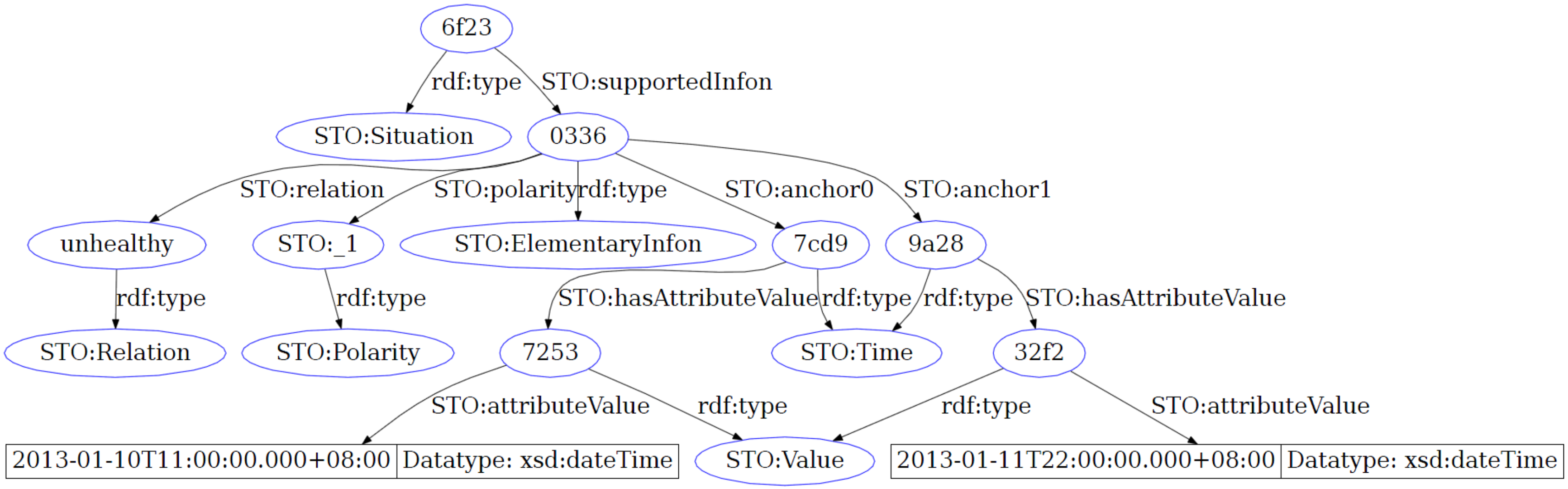
Example
BeijingAir (Observation metadata)



Example
BeijingAir (Observation data)



Example
BeijingAir (Situation)



Demo
BeijingAir in Wavellite

- Use Wavellite Web
  - visualize observations made by the BAM operated by the U.S. Embassy to Beijing
- Some basic functionality such as plotting of query results
- No situations of new particle formation, yet
- Available at
  - http://kuo-234195.uef.fi:8080/wavellite-web/
  - Under development and early stage
  - Expect bugs and limited features
  - Give feedback, including feature requests!

Next steps
New particle formation

- Get more powerful computer hardware!
- Import “all” Puijo DMPS, weather, and gas data
  - Possibly also other locations
- Develop knowledge acquisition task for NPF
  - Deploy it at situation layer
- Execute Wavellite over observation data
  - And let the framework represent knowledge for NPF
  - Situations, specifically events/episodes, of NPF with information for NPF class, as well as temporal and spatial locations
- Visualize results, e.g. by means of a timeline for NPF events

Next steps
And more …

- Continue developing Wavellite features, e.g.
  - Export functionality, heat maps, summary statistics
- Write papers and publish results
- Wavellite is in process towards open source publication at UEF
  - It should, thus, become available for free

References

[1] Compton, M., Henson, C., Neuhaus, H., Lefort, L., Sheth, A., 2009. A Survey of the Semantic Specication of Sensors, in: 2nd International Workshop on Semantic Sensor Networks, at 8th International Semantic Web Conference.

[2] Barnaghi, P., Ganz, F., Henson, C., Sheth, A., 2012. Computing Perception from Sensor Data. Technical Report. knoesis.org.

[3] De Maio, C., Fenza, G., Furno, D., Loia, V., 2012. Swarm-based semantic fuzzy reasoning for situation awareness computing, in: Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on, pp. 1--7.

[4] Doulaverakis, C., Konstantinou, N., Knape, T., Kompatsiaris, I., Soldatos, J., 2011. An Approach to Intelligent Information Fusion in Sensor Saturated Urban Environments, in: Intelligence and Security Informatics Conference (EISIC), 2011 European, pp. 108--115.

[5] Fenza, G., Furno, D., Loia, V., Veniero, M., 2010. Agent-based Cognitive approach to Airport Security Situation Awareness, in: Proceedings of the 2010 International Conference on Complex, Intelligent and Software Intensive Systems, IEEE Computer Society. pp. 1057--1062.

[6] Conroy, K., May, G., Roantree, M., Warrington, G., Cullen, S.J., Mc-Goldrick, A., 2011a. Knowledge acquisition from sensor data in an equine environment, in: Proceedings of the 13th international conference on data warehousing and knowledge discovery, Springer-Verlag, Berlin, Heidelberg. pp. 432--444.

[7] Gaglio, S., Gatani, L., Lo Re, G., Ortolani, M., 2007. Understanding the Environment Through Wireless Sensor Networks, in: Proceedings of the 10th Congress of the Italian Association for Articial Intelligence on AI*IA 2007: Articial Intelligence and Human-Oriented Computing, Springer-Verlag, Berlin, Heidelberg. pp. 72--83.

[8] Liu, J., Zhao, F., 2005. Towards semantic services for sensor-rich information systems, in: Broadband Networks, 2005. BroadNets 2005. 2nd International Conference on, pp. 967--974.

[9] Whitehouse, K., Zhao, F., Liu, J., 2006. Semantic Streams: A Framework for Composable Semantic Interpretation of Sensor Data, in: Römer, K., Karl, H., Mattern, F. (Eds.), Wireless Sensor Networks. Springer Berlin / Heidelberg. volume 3868 of LNCS, pp. 5--20.

[10] Wanner, L., Vrochidis, S., Tonelli, S., Mossgraber, J., Bosch, H., Karppinen, A., Myllynen, M., Rospocher, M., Bouayad-Agha, N., Bügel, U., Casamayor, G., Ertl, T., Kompatsiaris, I., Koskentalo, T., Mille, S., Moumtzidou, A., Pianta, E., Saggion, H., Serani, L., Tarvainen, V., 2011. Building an Environmental Information System for Personalized Content Delivery, in: Hrebícek, J., Schimak, G., Denzer, R. (Eds.), Environmental Software Systems. Frameworks of eEnvironment. Springer Boston. volume 359 of IFIP Advances in Information and Communication Technology, pp. 169--176.

[11] Wei, W., Barnaghi, P., 2009. Semantic Annotation and Reasoning for Sensor Data. Smart Sensing and Context , 66--76.

http://purl.oclc.org/NET/ssnx/ssn

http://vistology.com/ont/2008/STO/STO.owl

http://www.visitingdc.com/paris/eiffel-tower-paris-france.asp

http://storm-project.net/

Materials and methods
Situation Theory

- A situation, *s*, supports (one or more) infon(s) σ
  - Formally, $s \models \sigma$
- An infon is a tuple $\ll R, a_i, \ldots a_m, 0/1 \gg$ where
  - *R* is a relation
  - $a_i, \ldots a_m$ are objects appropriate to the relation *R*
  - And *0/1* is the polarity, whereby *1* means that $a_i, \ldots a_m$ "stand in" *R*
- Example
  - A situation of "class 1 new particle formation located at Puijo today"
  - $s \models \ll npf, class1, puijo, 2013\text{-}03\text{-}21, 1 \gg$
- Developed by Barwise and Perry, extended by Devlin
- Relates to Situational Awareness by Endsley

Materials and methods
Ontology

- An ontology formally represents knowledge as a set of concepts within a domain, and the relationships between pairs of concepts [Wikipedia: Ontology (information science)]
- A document that describes concepts and relationships of a domain; such documents are “written” by means of a (ontology) language
- An ontology is external to, and shared among, software systems
- Software systems commit to one (or more) ontology by adopting the defined terminology and semantics
- Key technology in knowledge representation and reasoning
- Developed within artificial intelligence research
- Used in many domains, e.g. bioinformatics (see Gene Ontology)

Materials and methods
Semantic Sensor Network (SSN) ontology

- Vocabulary for the representation of knowledge for sensors, properties, features, observations, …
- Adopted at the Wavellite observation layer
- Domain-specific sensors and measured properties of features are accommodated by extending from SSN ontology
- Examples
  - DifferentialMobilityParticleSizer *subClassOf* ssn:SensingDevice
  - PolydisperseAerosol *subClassOf* ssn:FeatureOfInterest
  - ParticleConcentration *subClassOf* ssn:Property
  - puijoDMPS *isA* DifferentialMobilityParticleSize
  - puijoAerosol *isA* PolydisperseAerosol
  - particleConcentrationWithDiameter800nm *isA* ParticleConcentration

Materials and methods
Situation Theory Ontology

- Vocabulary for the representation of knowledge for situations
- Borrows from Situation Theory
- Adopted at the Wavellite situation layer
- Domain-specific situations and relations are accommodated by extending from STO
- Example
  - NewParticleFormationEvent *subClassOf* sto:Situation
  - npf *isA* sto:Relation