International Conference on Knowledge Engineering and Semantic Web
September 30, 2015, Moscow, Russia

Knowledge-based environmental research infrastructure with Semantic Web technologies

Markus Stocker
markus.stocker@uef.fi
@markusstocker and @envinf

UNIVERSITY OF
EASTERN FINLAND

About me

- PhD candidate at UEF
- MSc in Informatics at UZH
- SPARQL optimization at UZH and HPLB
- Owlgres and PelletSpatial at Clark & Parsia
- PhD in Environmental Informatics at UEF
- MSc in Environmental Science at UEF

Acknowledgments

- Collaborators
  - Mikko Kolehmainen (UEF)
  - Mauno Rönkkö (UEF)
  - Elham Baranizadeh (UEF)
  - Kari Lehtinen (UEF)
  - Jussi Nikander (LUKE)
  - Hanna Huitu (LUKE)
- Organizations
  - University of Eastern Finland (UEF)
  - Natural Resources Institute Finland (LUKE)
- Funders
  - Academy of Finland
  - Tekes

Overview

- Knowledge engineering in environmental science
- Environmental research infrastructure (ERI)
- Knowledge-based ERI
- Using Semantic Web technologies
- Applications
- Q&A

Knowledge engineering in environmental science

Review

Advancing ecological research with ontologies

Joshua S. Madin[1,2], Shawn Bowers[3], Mark P. Schildhauer[1] and Matthew B. Jones[1]

[1] National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, CA 93101, USA
[2] Department of Biological Sciences, Macquarie University, New South Wales 2109, Australia
[3] UC Davis Genome Center, University of California, Davis, CA 95616, USA

Ontologies have assisted other disciplines (e.g. molecular biology); ecology can benefit from similar approaches

Ambiguous terminology prevents incorporating data into broader-scale studies

ECOLOGICAL COMPLEXITY 5 (2008) 272–279

ELSEVIER

available at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/ecocom

Short Note

An ontology for landscapes

Christopher A. Lepczyk[a,*], *Christopher J. Lortie*[b], *Laurel J. Anderson*[c]

[a] Department of Forest Ecology and Management, University of Wisconsin-Madison, Madison, WI 53706, USA
[b] Biology Department, York University, 4700 Keele St., Toronto, ON, 3J 1P3 Canada
[c] Department of Botany/Microbiology, Ohio Wesleyan University, Delaware, OH 43015, USA

Few concepts in ecology convey such a wide range of meanings as the term landscape

Different usage by scientists creates linguistic uncertainty and hinders automated synthesis of datasets

IOS Press

Building a volcano-domain ontology

Volcán de Colima, México, case study

JRG Pulido [a], MA Aréchiga [b]
[a] *Faculty of Telematics, University of Colima, México, jrgp@ucol.mx*
[b] *Faculty of Telematics, University of Colima, México, mandrad@ucol.mx*

Semantify large volume seismic data such that software agents can carry out inference and forecasting

Use data mining (SOM) to extract taxonomy for volcanic events (eruptions, tremors, ...) represented in OWL

ECOLOGICAL INFORMATICS 2 (2007) 279–296

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/ecolinf

An ontology for describing and synthesizing ecological observation data

Joshua Madin[a,e,*], *Shawn Bowers*[b], *Mark Schildhauer*[a], *Sergeui Krivov*[c], *Deana Pennington*[d], *Ferdinando Villa*[c]

[a]*National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, California 93101, USA*
[b]*Genome Center, University of California, Davis, California 95616, USA*
[c]*Gund Institute for Ecological Economics, University of Vermont, Burlington, Vermont 05405, USA*
[d]*University of New Mexico, Albuquerque, New Mexico 87131, USA*
[e]*Department of Biological Sciences, Macquarie University, New South Wales 2109, Australia*

Details of observational data are not recorded; contextual information is implicit

Extensible Observation Ontology (OBOE) for capturing semantic information of observational datasets

An Ontological Representation of Time Series Observations on the Semantic Sensor Web

Cory A. Henson[1], Holger Neuhaus[2], Amit P. Sheth[1], Krishnaprasad Thirunarayan[1], and Rajkumar Buyya[3]

[1] Kno.e.sis Center, Department of Computer Science and Engineering
Wright State University, Dayton, OH 45435, USA
{cory, amit}@knoesis.org

[2] CSIRO Tasmanian ICT Centre
GPO Box 1538, Hobart, TAS, 7001, Australia
holger.neuhaus@csiro.au

[3] GRIDS Lab, Department of Computer Science and Engineering
University of Melbourne, Australia
raj@csse.unimelb.edu.au

OGC O&M facilitates syntax-level integration
but lacks the ability of semantic-level integration

Early work toward W3C SSN ontology

Environmental Modelling & Software 24 (2009) 577–587

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

ELSEVIER

Review

Modelling with knowledge: A review of emerging semantic approaches to environmental modelling

Ferdinando Villa[a,*], Ioannis N. Athanasiadis[b], Andrea Emilio Rizzoli[b]

[a] *Ecoinformatics Collaboratory, Gund Institute for Ecological Economics and Department of Plant Biology, University of Vermont, 617 Main Street, Burlington, VT, USA*
[b] *Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, Lugano, Switzerland*

The understanding of an environmental system is usually implicit to models; it resides outside model specification and implementation

This severely limits the options in reusing environmental models

Combining OWL with RCC for Spatioterminological Reasoning on Environmental Data

Rolf Grütter and Bettina Bauer-Messmer

Swiss Federal Institute WSL, An Institute of the ETH Board,
Zürcherstrasse 111, 8903 Birmensdorf, Switzerland
{rolf.gruetter, bettina.bauer}@wsl.ch

Spatial relations between regions can be computed geometrically (GIS)

How to make the resulting qualitative spatial relations accessible to a logic formalism

Process queries that combine thematic and spatial aspects

Transactions in GIS, 2003, 7(3): 371–391

Research Article

The Role of Knowledge Representation in Geographic Knowledge Discovery: A Case Study

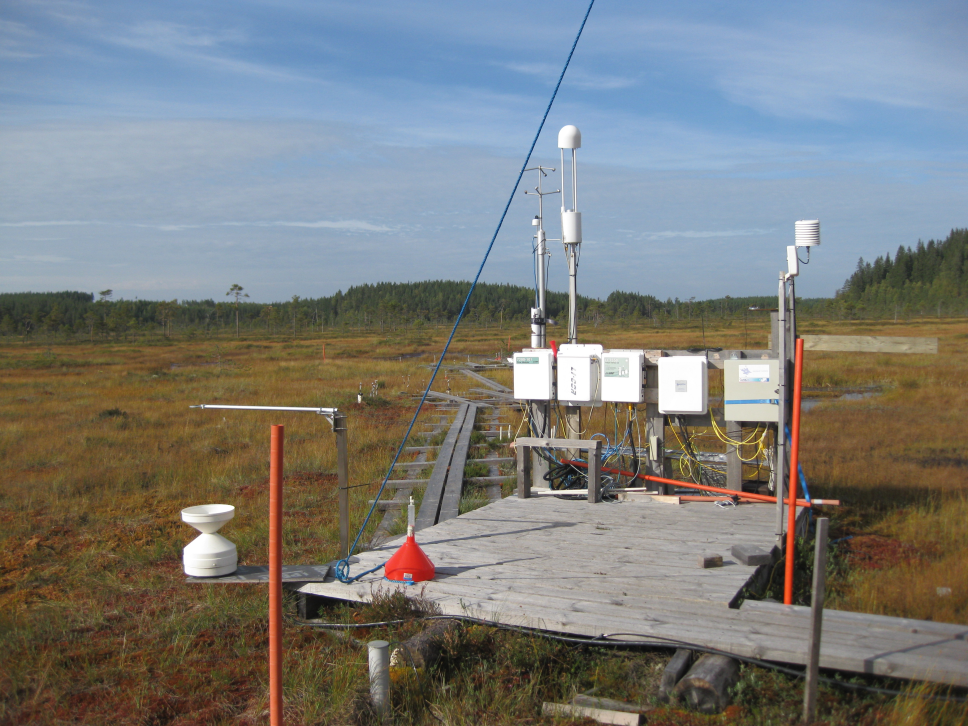
Jeremy Mennis
Department of Geography
University of Colorado, Boulder

Donna J Peuquet
Department of Geography
The Pennsylvania State University

We suggest that geographic data models that support knowledge discovery must represent both observational data and derived knowledge

Hierarchy of storm types (expert knowledge) is formally represented (within a database context) to extract instances of storms from observational data

Environmental research infrastructure


7200-101
LI-COR
LI-7550
LI-COR

SMEAR

ICOS
INTEGRATED CARBON OBSERVATION SYSTEM

european multidisciplinary seafloor & water column observatory
emso

neon
National Ecological Observatory Network

GLEON
global lake ecological observatory network

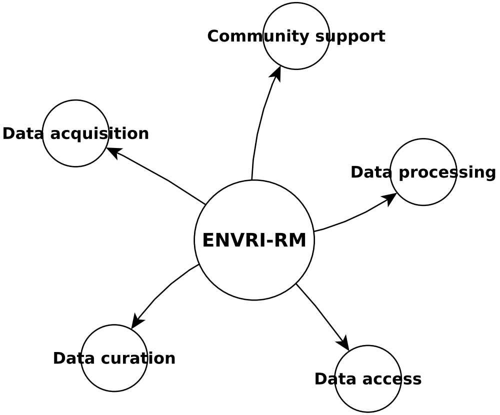
EuroGOOS
European Global Ocean Observing System

OCEAN NETWORKS CANADA

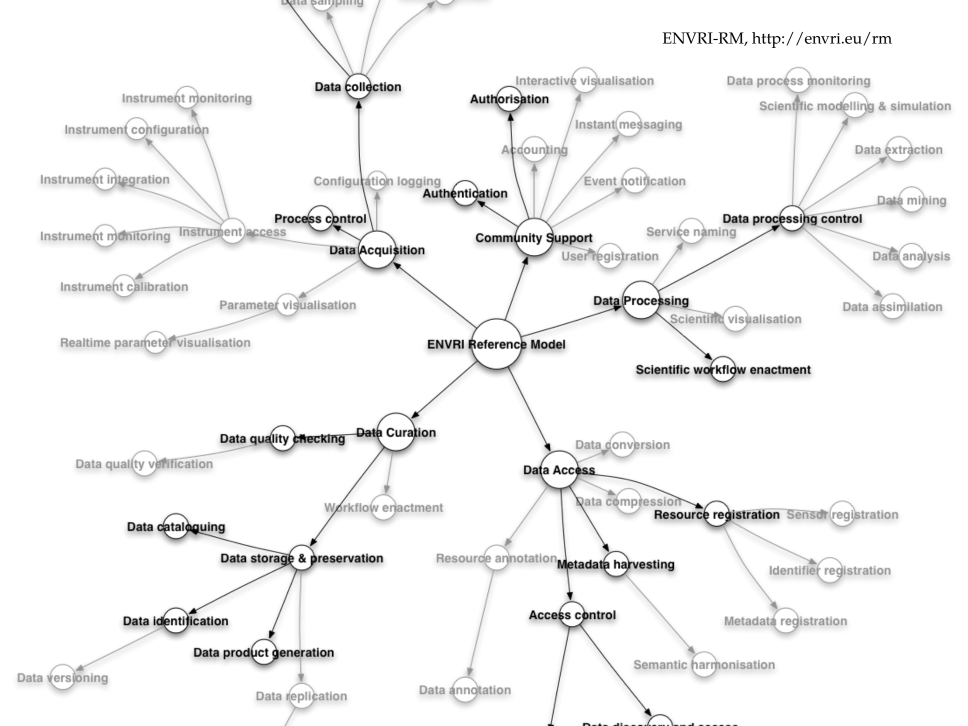
TERN
Terrestrial Ecosystem Research Network

```
  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
```

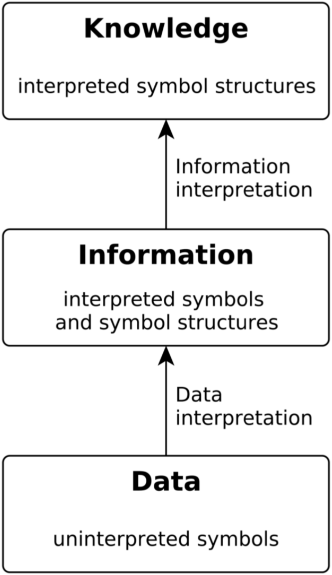
ENVRI Common Operations of Environmental Research Infrastructures



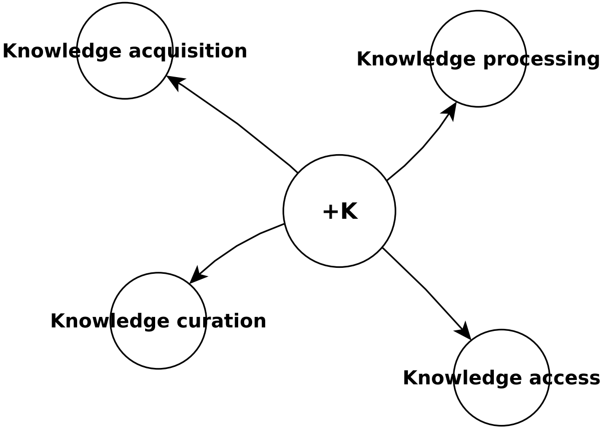
Community support
Data acquisition
Data processing
ENVRI-RM
Data curation
Data access


ENVRI-RM, http://envri.eu/rm
ENVRI Reference Model
Data Acquisition
Data collection
Process control
Configuration logging
Instrument access
Instrument monitoring
Instrument configuration
Instrument integration
Instrument monitoring
Instrument calibration
Parameter visualisation
Realtime parameter visualisation
Community Support
Authorisation
Authentication
Interactive visualisation
Instant messaging
Accounting
Event notification
User registration
Data Processing
Data processing control
Data process monitoring
Scientific modelling & simulation
Data extraction
Data mining
Data analysis
Data assimilation
Service naming
Scientific visualisation
Scientific workflow enactment
Data Curation
Data quality checking
Data quality verification
Workflow enactment
Data storage & preservation
Data cataloguing
Data identification
Data versioning
Data product generation
Data replication
Data Access
Data conversion
Data compression
Resource registration
Sensor registration
Identifier registration
Metadata registration
Resource annotation
Data annotation
Metadata harvesting
Semantic harmonisation
Access control

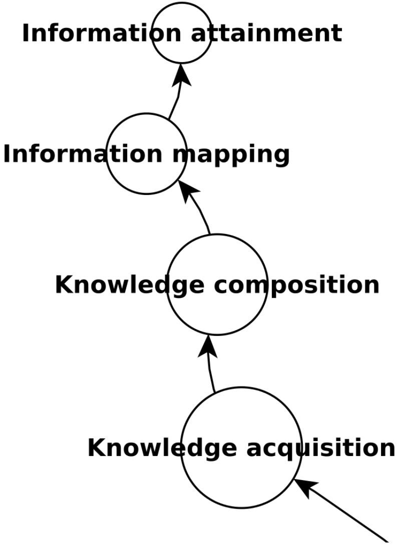
Knowledge-based ERI



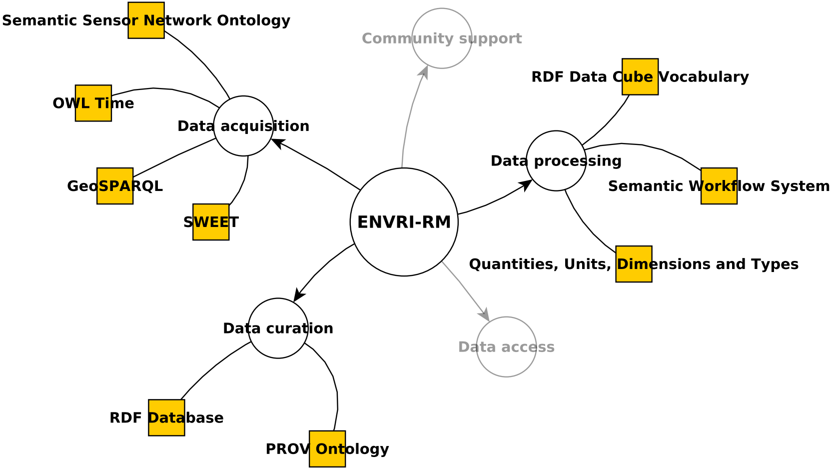
Aamodt and Nygård (1995) *Data & Knowledge Engineering* 16(3):191-222 (Adapted)

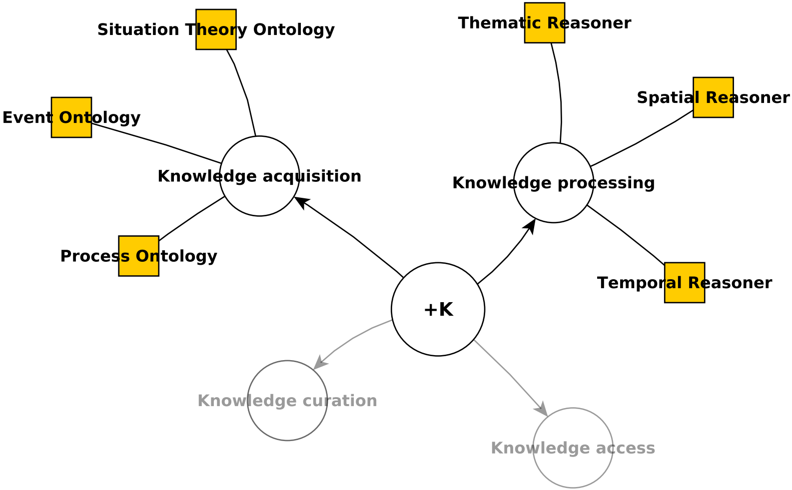

Knowledge acquisition
Knowledge processing
+K
Knowledge curation
Knowledge access


Rule-based reasoning
Conceptual reasoning
Information attainment
Spatio-temporal reasoning
Information mapping
Knowledge analysis
Reasoning
Knowledge composition
Knowledge acquisition
Knowledge processing
Knowledge visualization
+K
Knowledge curation
Knowledge quality checking
Knowledge access
Knowledge storage and preservation
Knowledge discovery and retrieval
Knowledge representation
Knowledge publication
Knowledge identification


Information attainment
Information mapping
Knowledge composition
Knowledge acquisition

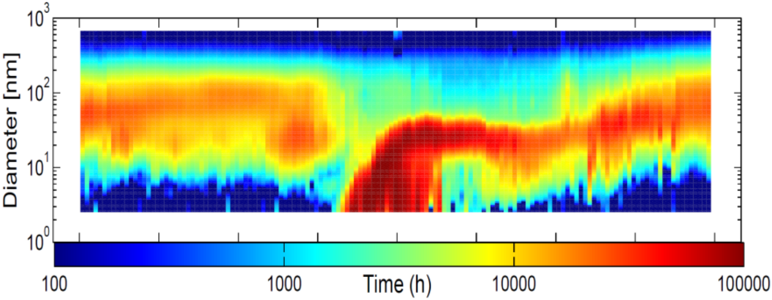
Using
Semantic Web technologies


Semantic Sensor Network Ontology
Community support
RDF Data Cube Vocabulary
OWL Time
Data acquisition
Data processing
GeoSPARQL
Semantic Workflow System
SWEET
ENVRI-RM
Quantities, Units, Dimensions and Types
Data curation
Data access
RDF Database
PROV Ontology

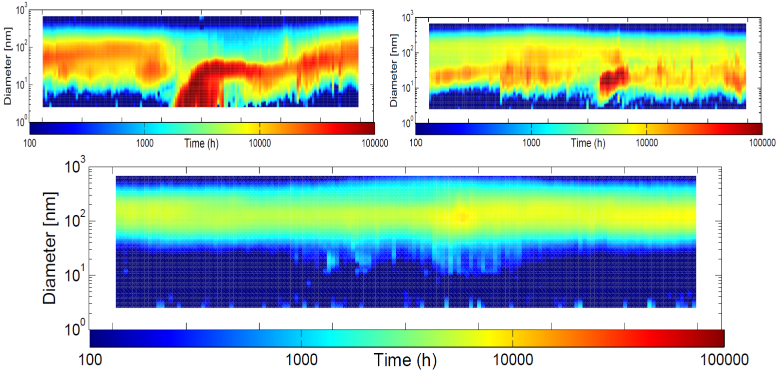

Situation Theory Ontology
Thematic Reasoner
Spatial Reasoner
Event Ontology
Knowledge acquisition
Knowledge processing
Process Ontology
Temporal Reasoner
+K
Knowledge curation
Knowledge access

Applications


DROPLET
MEASUREMENT
TECHNOLOGIES
www.dropletmeasurement.com



Hamed et al. (2007) *Atmos. Chem. Phys.*, 7, 355-376



Hamed et al. (2007) *Atmos. Chem. Phys.*, 7, 355-376

Result

Day	Event class
15.06.2014	1
16.06.2014	0
17.06.2014	3
...	...

A knowledge-based ERI would ...

- Support the researcher in data analysis
  - Automated assessment of event class
  - Allow for curation (review) of its assessment
- Relate contextual information
  - Date, event class, location, plot, other event attributes
  - Create a knowledge object
- Represent knowledge object
  - According to a formal vocabulary (ontology)
  - Using suitable data formats (e.g. RDF)
- Manage and process knowledge objects
  - Persist knowledge objects (knowledge base)
  - Support access to knowledge objects
  - Reasoning to infer new knowledge (rules)

Relation

Outbreak
Acute outbreak
Pest protection

Start date

____-__-__ __:__

End date

____-__-__ __:__

filter

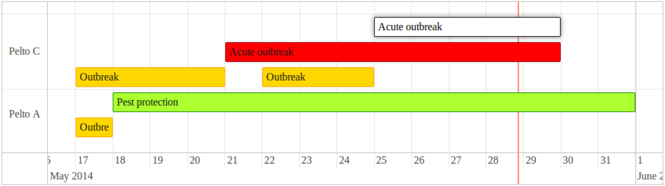
Relation
Acute outbreak

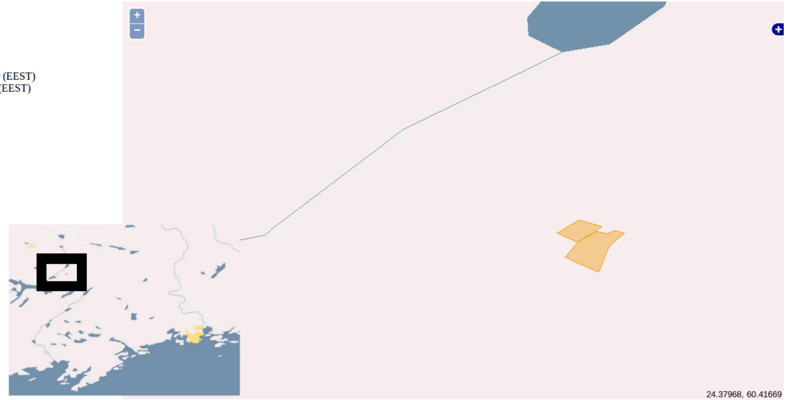
Relevant individual
Drechslera tritici-repentis

Temporal location
Sun May 25 2014 00:00:00 GMT+0300 (EEST)
Fri May 30 2014 00:00:00 GMT+0300 (EEST)

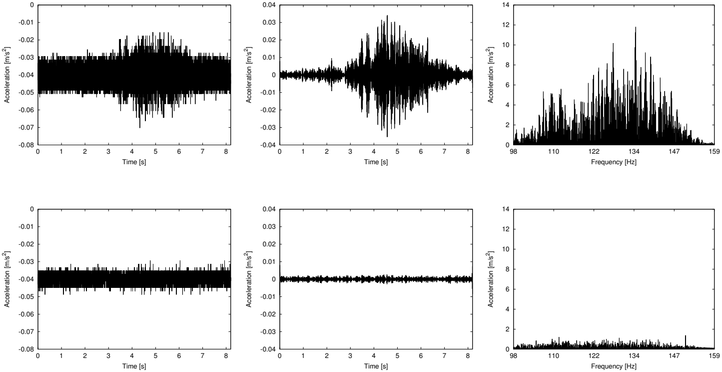
Spatial location
Pelto C

Polarity
True



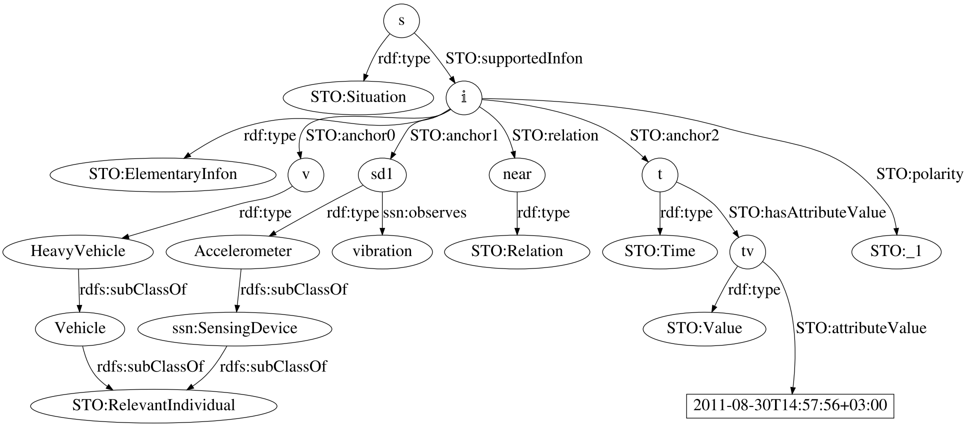



CARPOOL IS 3 OR MORE PERSONS PER VEHICLE


Acceleration [m/s^2]
Time [s]
Acceleration [m/s^2]
Time [s]
Acceleration [m/s^2]
Frequency [Hz]
Acceleration [m/s^2]
Time [s]
Acceleration [m/s^2]
Time [s]
Acceleration [m/s^2]
Frequency [Hz]

Situations

- Structured parts of reality
- Formalized in situation theory (Barwise, Perry, Devlin)
- Situation s is said to support ($\models$) infons
- Infon σ is a tuple consisting of
- Relation R; Objects $a_1, \ldots, a_m$; Polarity $1/0$
- Objects can be physical entities in the environment, or ...
- Temporal and spatial locations, values, situations
- If polarity is 1, objects stand in the relation R


s
rdf:type
STO:supportedInfon
STO:Situation
i
rdf:type
STO:anchor0
STO:anchor1
STO:relation
STO:anchor2
STO:polarity
STO:ElementaryInfon
v
sd1
near
t
rdf:type
rdf:type
ssn:observes
rdf:type
rdf:type
STO:hasAttributeValue
HeavyVehicle
Accelerometer
vibration
STO:Relation
STO:Time
tv
STO:_1
rdfs:subClassOf
rdfs:subClassOf
rdf:type
Vehicle
ssn:SensingDevice
STO:Value
STO:attributeValue
rdfs:subClassOf
rdfs:subClassOf
STO:RelevantIndividual
2011-08-30T14:57:56+03:00

Take aways

- Knowledge engineering relevant to environmental science
- Environmental research infrastructures
  - Interesting and challenging data-based systems
  - Perhaps even more as knowledge-based systems
  - Opportunities for computer science communities
- Knowledge-based ERI
  - Interesting application area for Semantic Web technologies
- Is it a luxury? Who is going to fund this?