Using Semantic Annotation for Knowledge Extraction from Geographically Distributed and Heterogeneous Sensor Data

Alexandra Moraru
Carolina Fortuna
Dunja Mladenić

Jozef Stefan Institute
Slovenia
Outline

• Introduction

• System architecture

• Case Study: Participatory Sensing

• Conclusions
Introduction

Internet of Things
- Scalability
- Mobility
- Interoperability

Sensor Web
- Web accessible sensor network and archive sensor data

Participatory sensing
- Access to various types of data
- Problems may appear in understanding the data
- Solution: providing semantic context
Semantic Annotation

• Annotating sensor descriptions with concepts from an ontology

• Machine understanding of the sensors descriptions and data streams

• Enables reasoning mechanism for selecting streams for processing or monitoring
System Architecture

Knowledge Base
Ontology
Logic Rules

Inference Engine

Semantic Sensor Web

Semantic Annotations

Publisher

Sensor Descriptions

Sensor Data
Case Study: Participatory Sensing

• Pachube
  – platform that supports storing and sharing sensor data (stream of measurements).
  – structured metadata describing the sensor data streams (including natural language description and tags).
    www.pachube.com

• Cyc
  – general ontology and a knowledge base for representing common sense knowledge
  – organized by contexts (microtheories)
    www.opencyc.org
Case Study: Participatory Sensing

---

**Individual: IJSSensor**

isa: Sensor

hasDataStream: IJSSensor-Data1

hasDomain: Physical

hasExposure: Indoor

latitude: (Degree-UnitOfAngularMeasure 46.0425085163033)

longitude: (Degree-UnitOfAngularMeasure 14.4882792234421)

---

**Individual: IJSSensor-Data1**

isa: DataStream

hasUnitOfMeasurement: DegreeCelsius

measures: Temperature

---

```xml
<title> IJSSensor</title>
<status>live</status>
<location domain="physical" exposure="indoor">
  <lat>46.0425085163033</lat>
  <lon>14.4882792234421</lon>
</location>
<data id="0">
  <tag>Temperature</tag>
  <unit type="basicSI" symbol="°C">Celsius</unit>
</data>
```
Case Study: Participatory Sensing

<title> IJSSensor</title>
<status>live</status>
<location domain="physical" exposure="indoor">
  <lat>46.0425085163033</lat>
  <lon>14.4882792234421</lon>
</location>
<data id="0">
  <tag>Temperature</tag>
  <unit type="basicSI" symbol="°C">Celsius</unit>
</data>

Individual: IJSSensor
isa: Sensor
hasDataStream: IJSSensor-Data1
hasDomain: Physical
hasExposure: Indoor
latitude: (Degree(UnitOfAngularMeasure 46.0425085163033))
longitude: (Degree(UnitOfAngularMeasure 14.4882792234421))

Individual: IJSSensor-Data1
isa: DataStream
hasUnitOfMeasurement: DegreeCelsius
measures: Temperature
Case Study: Participatory Sensing

Frequent tags for data streams descriptions in Pachube

<table>
<thead>
<tr>
<th>Domain</th>
<th>Tag</th>
<th>Number of occurrences</th>
<th>Cyc Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temperature related tags</strong></td>
<td>temperature</td>
<td>336</td>
<td>Fever</td>
</tr>
<tr>
<td></td>
<td>temp</td>
<td>32</td>
<td>Temporary Worker</td>
</tr>
<tr>
<td></td>
<td>celsius</td>
<td>293</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td><strong>Power consumption related tags</strong></td>
<td>electricity</td>
<td>389</td>
<td>Electricity</td>
</tr>
<tr>
<td></td>
<td>watts</td>
<td>34</td>
<td>Watt</td>
</tr>
<tr>
<td><strong>Distinct tags</strong></td>
<td>Distinct tags</td>
<td>2238</td>
<td>Data streams</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>2238</td>
<td>9466</td>
<td></td>
</tr>
</tbody>
</table>
Searching for Sensors

• Which are the sensors that measure temperature in Ljubljana?

  (and
    (isa ?X Sensor)
    (hasDataStream ?X ?DS)
    (measures ?DS Temperature)
    (distanceBetween ?X CityOfLjubljanaSlovenia (Kilometer ?DIST))
    (lessThan ?DIST 10))
Reasoning with Sensor Data

• Detection of anomalous data measurements
  • data streams measuring temperature
  • Mediterranean region
  • Summer time

• Temperature measurements below a 10 °C are considered anomalous
  • for an outdoor exposure of the sensing device
Reasoning with Sensor Data

(implies
 (and
  (isa ?SENSOR Sensor)
  (sensorMeasurementsInterval ?SENSOR ?INT)
  (temporalBoundsContain ?SEASON ?INT)
  (isa ?SEASON SummerSeason)
  (hasRegionLocation ?SENSOR ?REGION)
  (hasClimateType ?REGION MediterraneanClimateCycle)
  (hasExposure ?SENSOR Outdoor)
  (hasDataStream ?SENSOR ?DS)
  (measures ?DS Temperature)
  (valueOf ?DS (DegreeCelsius ?C))
  (lessThan ?C 10))
 (anomalousMeasurements ?SENSOR ?DS))
Conclusions & Future Work

• Semantic annotations can provide context for the sensor measurements and observations

• We proposed and discussed a system architecture for automatic annotation
  – a too general ontology will not be able to successfully annotate all sensor descriptions

• Future Work
  – provide more specific context for the concepts used in sensor annotation
  – virtual sensor composition