Exposing Real World Information for the Web Of Things

Alexandra Moraru
Matevž Vučnik
Maria Porcius
Carolina Fortuna
Mihael Mohorčič
Dunja Mladenič
Introduction

- Try to imagine a "world littered with trillions" of wireless sensors. Now try to imagine the problems getting even a few thousand of them to work together in any kind of intelligible way… [1]

- We want a way of doing sensing that can make the data available to any application that needs that specific data [1]

Introduction

- **Internet of Things** – world-wide network of heterogeneous smart objects
  - sensors, actuators, RFIDs, MEMS
  - based on standard communication protocols
  - focused on establishing connectivity

- **Web of Things** – integrating smart objects into the Web
  - a.k.a Sensor Web, Physical Web
  - based on standards like HTML, XML, RSS
  - focused on application layer

- **The “Things”**
  - a set of sensor nodes and/or embedded device + physical things which are abstracted as one “thing” (large water tank + set of sensor nodes monitoring water level, temperature and purity)

- **Wireless Sensor Network - WSN:**
  - wireless network of spatially distributed nodes, which jointly observe certain phenomena
  - Traditionally: low complexity, low power, small size/weight, long life, autonomous, short range, low cost
Introduction

Web Of Things use cases

- Motivated by an increased interest in automatic management of large systems
- Commercial use cases (non-exhaustive list):
  - Power grids
  - Transport systems
  - Water distribution
  - Logistics
  - Industrial automation
  - Agriculture
- Academic
  - Distributed sensing infrastructure
    - Microclimate monitoring
    - Volcano monitoring
    - Psychology of masses
Outline

- Exposing Real World Data
- SemSense Architecture
  - Data Collection
  - Storage
  - Semantic Enrichment
  - Data Publishing
- Conclusions
Exposing Real World Data

- Web 1.0
  - Static data
  - Read-Only

- Web 2.0
  - User generated data
    - blogs, socializing sites
  - Read-Write

- Web 3.0
  - Semantic Web, Web of Data
  - RDF, OWL, RDFa for describing things instead of documents
Exposing Real World Data

- Philosophy behind Web 3.0
  - Provide machine understandable representation of data
  - Link these data for discovery and reasoning

- Linked Data
  - method of exposing, sharing, and connecting data via dereferenceable URIs on the Web.
    - URI for the real-world object itself.
    - URI for a related information resource that describes the real-world object and has an HTML representation.
    - URI for a related information resource that describes the real-world object and has an RDF/XML representation.
SemSense Architecture

Implementation Scenario

- **Data Collection**
  - Versatile Sensor Nodes deployed in an outdoor testbed
  - observed properties: temperature, humidity, light and pressure
  - Two protocols for meta-data and measurements collection

- **Storage component**
  - Database schema for separation of data
  - Running on MySQL server

- **Semantic Enrichment**
  - Semantic Sensor Network (SSN) ontology – W3C standardization
  - Mapping rules between the database and vocabulary

- **Publishing Component**
  - D2R Server exposes enriched data
  - According to LOD principles
SemSense Architecture
SemSense Architecture

Information Collection

- Crowdsourcing
  - depends on participants willingness for providing accurate and complete descriptions
  - large amount of data (Pachube >9000 sources)

- Automatic collection
  - Implementation of an **identification protocol** - SIDP

![Diagram of SemSense architecture](attachment:semSense.png)
SemSense Architecture

Data Collection

Data Storage

Semantic Enrichment

Data Publishing

Mappings
Ontologies

Geo-tagging with mobile GW

Internet

SIDP

Web Services
SPARQL Clients
Web Browsers
SemSense Architecture

Data Storage

- Database Management Systems
  - Abstraction levels
  - Analysis and querying
  - Large amounts of data

- Distributed storage on the sensor network level (i.e. TinyDB)
  - Data retrieved directly from the sensor

- Centralized storage on the middle level (MySQL)
  - Storage of both meta-data and measurements
  - Automatic data insertion by data collection server
  - Database design closely related to hardware design, where a sensor node features a set of sensors
SemSense Architecture
Data Storage

- Database schema
  - Meta-data: physical devices and phenomena observed
  - Measurements: timestamp, value, sensor id
  - Separation between meta-data and measurements.
- A sensor node can have several sensors attached to it
  - Our testbed: each node has six sensors
  - Same type of sensors on a sensor node
SemSense Architecture
SemSense Architecture
Semantic Enrichment

- Semantic Vocabulary
  - SSN ontology
    - Result of W3C Semantic Sensor Network Incubator Group
    - Aligned to DOLCE Ultra Lite
  - Subset of concepts and relationships
SemSense Architecture
Semantic Enrichment

- Semantic Vocabulary
  - Basic GeoWGS84 vocabulary
    - Geographical location of platforms
    - namespace for representing the coordinates
  - GeoNames
    - Geographical region names
    - findNearbyPlaceName web services
  - `based_near` predicate from FOAF
SemSense Architecture

Semantic Enrichment

Between the database content and semantic vocabulary, based on the D2RQ language.
SemSense Architecture
SemSenseArchitecture
Data Publishing

- Publishing methods
  - standardized web services – OGC’s SOS
  - application specific: Pachube, Sensorpedia
  - Linked Sensor Data

- Publishing tool used
  - D2R Server – generates RDF and HTML descriptions of database content based on the mapping rules
    - no replication of the database
    - can be accessed by SPARQL clients and browsed using HTML interface
SemSense Architecture
Data Publishing

http://sensorlab.ijs.si:2020/

This is a database published with D2R Server. It can be accessed using:
1. your plain old web browser
2. Semantic Web browsers
3. SPARQL clients.

1. HTML View
You can use the navigation links at the top of this page to explore the database.

2. RDF View
You can also explore this database with Semantic Web browsers like Tabulator or Disco. To start browsing, open this entry point URL in your Semantic Web browser:

   http://sensorlab.ijs.si:2020/all

3. SPARQL Endpoint
SPARQL clients can query the database at this SPARQL endpoint:

   http://sensorlab.ijs.si:2020/sparql

The database can also be explored using this AJAX-based SPARQL Explorer.
Which are the sensors measuring temperature located in the Vič region of the city of Ljubljana?

```
SELECT DISTINCT ?s WHERE {
  ?sn ssn:hasSubSystem ?s.
}
```

- [Sensor Device #3](http://sensorlab.ijs.si:2020/resource/sensor_devices/403AB8FC-3)

```
<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ssn:hasSubSystem</td>
<td><a href="http://sensorlab.ijs.si:2020/resource/sensor-nodes/403AB8FC-3">http://sensorlab.ijs.si:2020/resource/sensor-nodes/403AB8FC-3</a></td>
</tr>
<tr>
<td>rdfs:label</td>
<td>sensor device #3</td>
</tr>
<tr>
<td>rdfs:label</td>
<td>sensor device #3</td>
</tr>
</tbody>
</table>
```

Conclusions

- It is important to make sensor data available
- SemSense architecture for collecting and exposing real world data to the Web
  - Data collection, storage, semantic enrichment, publishing according to LOD principles

- Future Work
  - Extend SemSense for multiple heterogeneous sensor networks
  - Automatically generate mapping rules
  - Comparative analysis with other similar systems
THE END!