Dynamic Semantics for the Internet of Things

Payam Barnaghi
Institute for Communication Systems (ICS)
University of Surrey
Guildford, United Kingdom

Things, Devices, Data, and lots of it

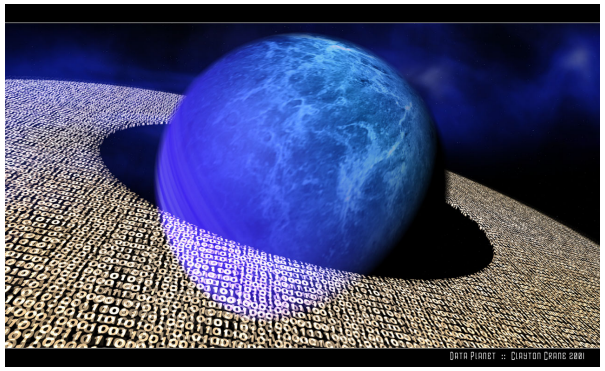


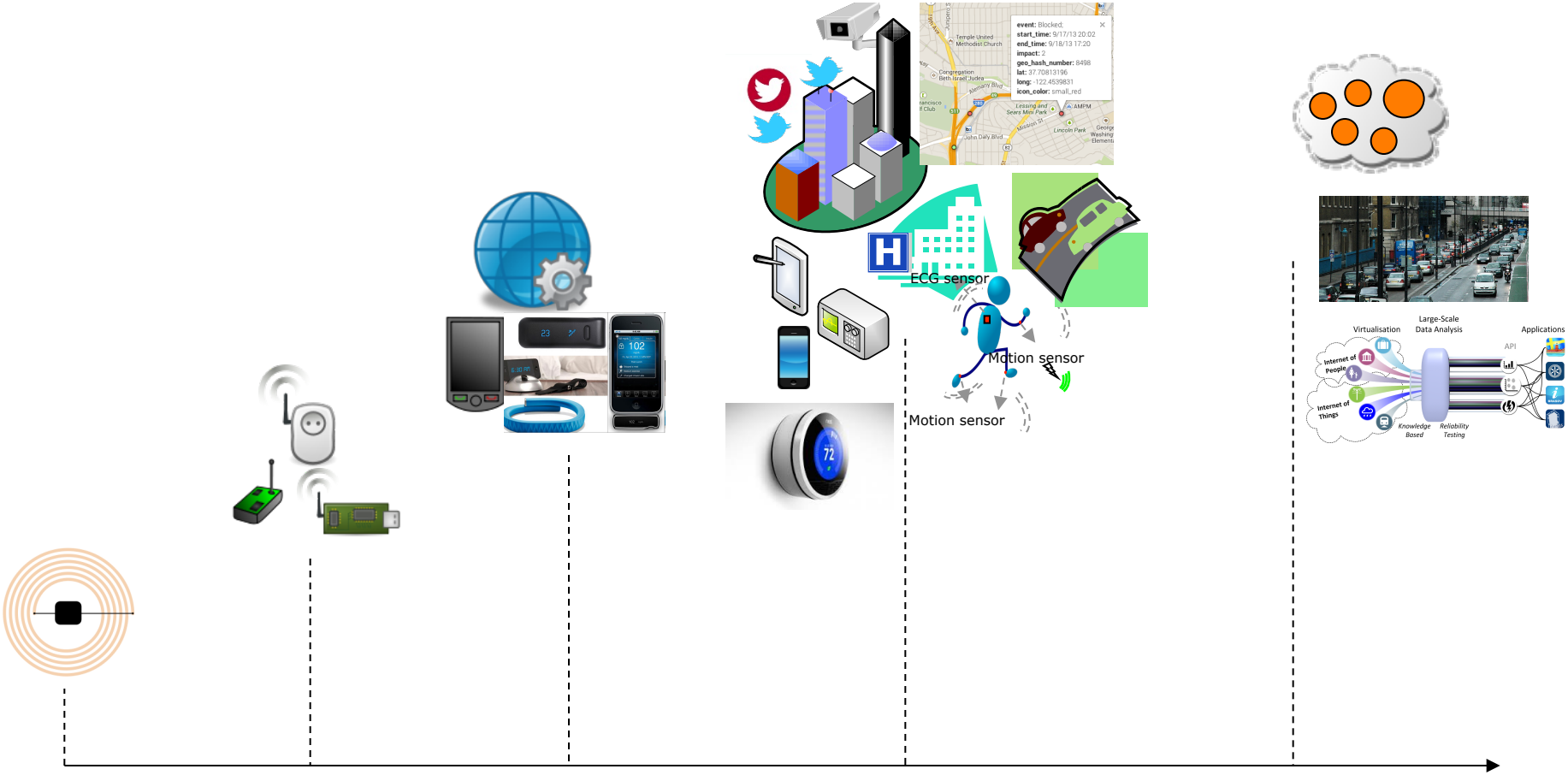
image courtesy: Smarter Data - I.03_C by Gwen Vanhee

Data in the IoT

- Data is collected by sensory devices and also crowd sensing sources.
- It is time and location dependent.
- It can be noisy and the quality can vary.
- It is often continuous - streaming data.

- There are other important issues such as:
  - Device/network management
  - Actuation and feedback (command and control)
- Service and entity descriptions are also important.

Internet of Things: The story so far



RFID based solutions

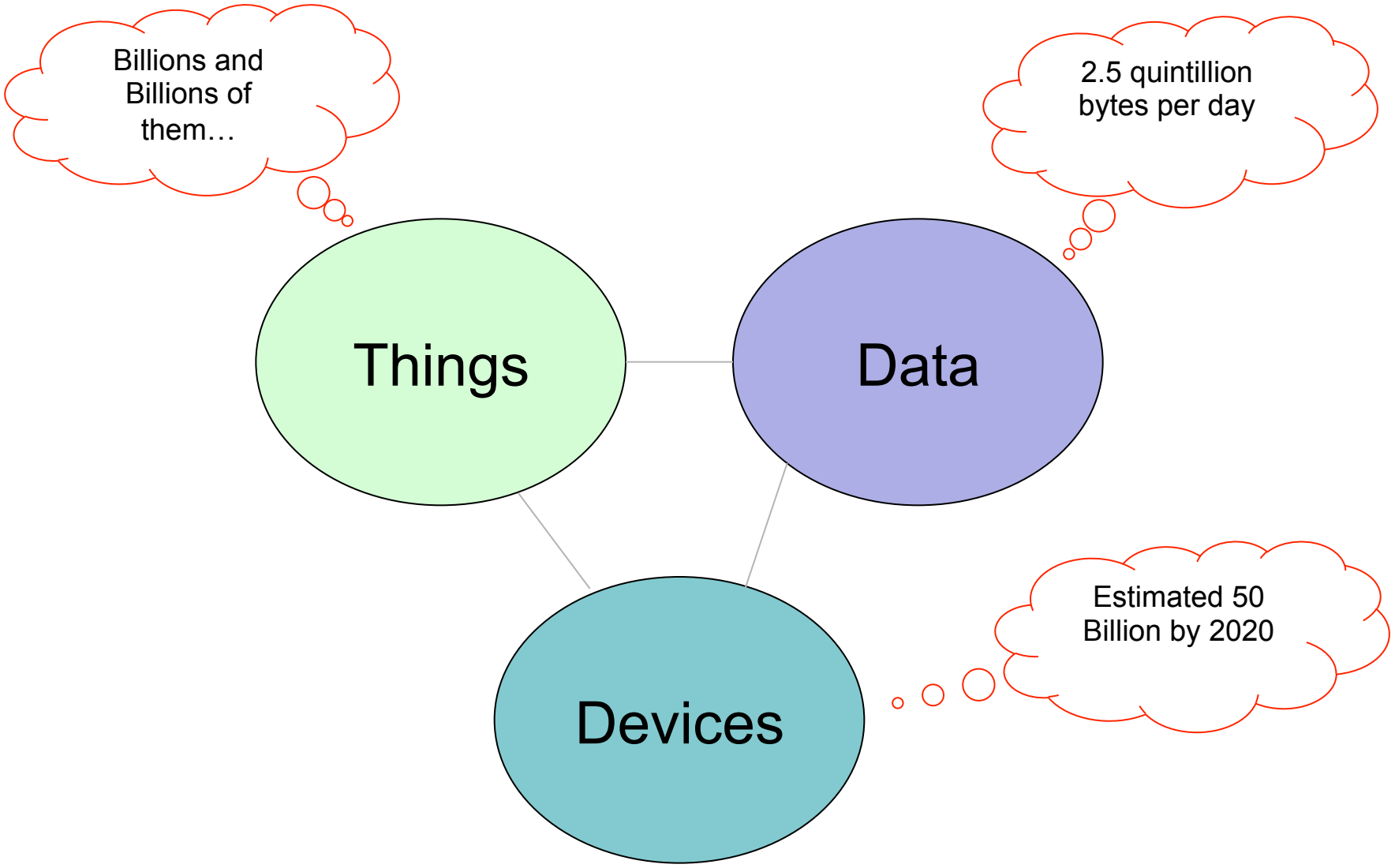
Wireless Sensor and Actuator networks , solutions for communication technologies, energy efficiency, routing, …

Smart Devices/ Web-enabled Apps/ Services, initial products, vertical applications, early concepts and demos, …

Physical-Cyber-Social Systems, Linked-data, semantics, M2M, More products, more heterogeneity, solutions for control and monitoring, …

Future: Cloud, Big (IoT) Data Analytics, Interoperability, Enhanced Cellular/Wireless Com. for IoT, Real-world operational use-cases and Industry and B2B services/applications, more Standards…

Scale of the problem



Heterogeneity, multi-modality and volume are among the key issues.

We need interoperable and machine-interpretable solutions…

Human Brain and (Sensory) Big Data

- Collecting the data is done by human senses but encoding and retrieving it is a bigger challenge.
- The two key properties of the human brain and its design are Richness and Associative Access*.
- Associative access enables us to access our thoughts in different ways by semantic or perceptual associations.
- Brian can process these data and provide actionable-knowledge.

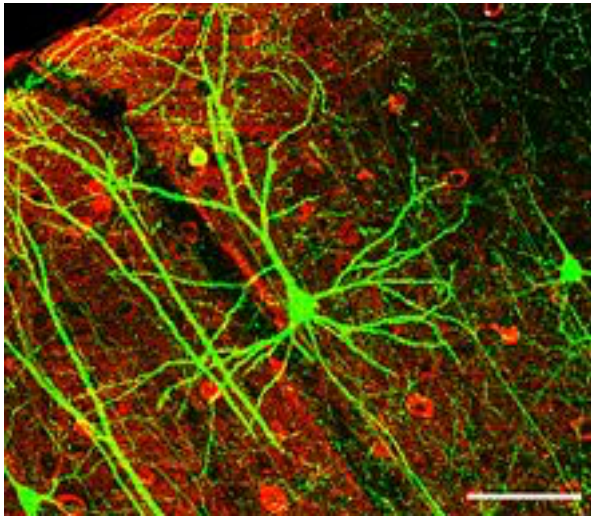


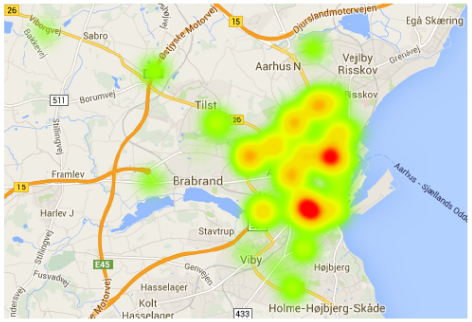
Image source: Wikipedia

* The organised Mind, Daniel J. Levitin, Penguin Books.

IoT and and (Sensory) Big Data

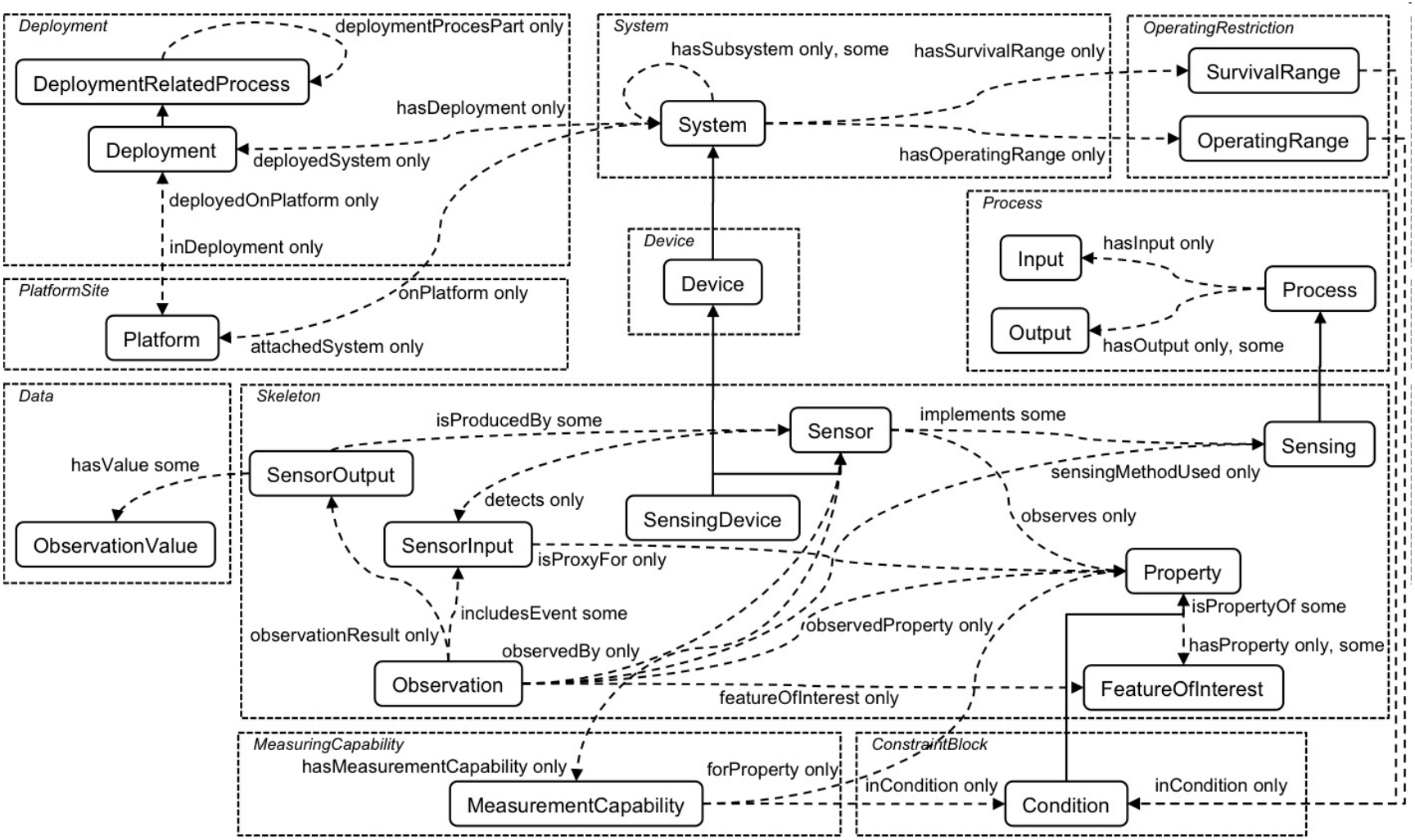
- Collecting data is not the most difficult challenge (of course we still need better devices, more energy efficient devices/way of collecting data, intelligent networks and better telecom)
- The biggest challenge is to organise and access/ retrieve data more efficiently and by using different (high-level) associations.
- We need to integrate different sources and process/analyse them to extract actionable-information from the raw data.

- Semantic technologies and rich metadata seem to be the way forward.



But why don’t we still have fully integrated semantic solutions in the IoT?

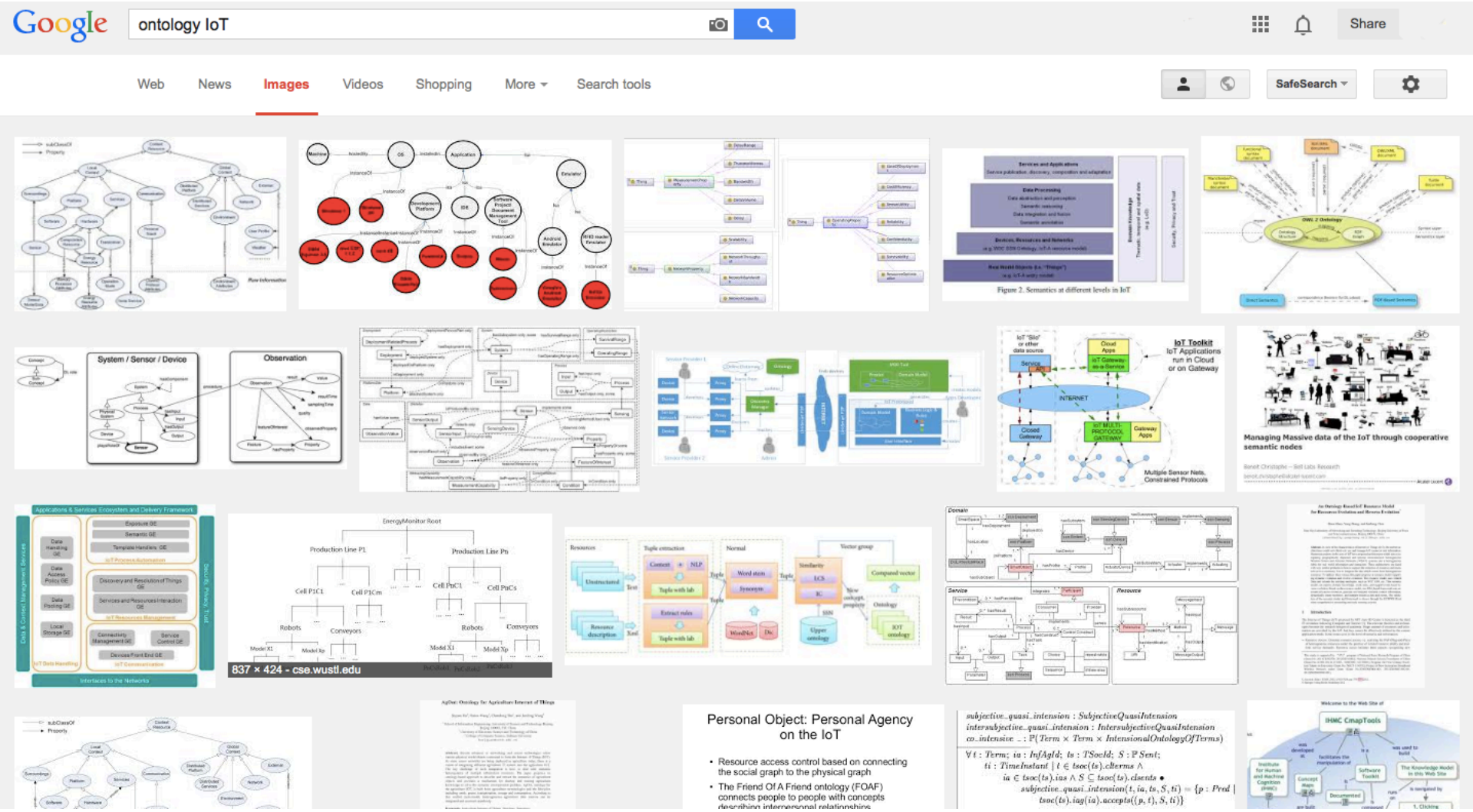
Some good existing models: SSN Ontology



Ontology Link: http://www.w3.org/2005/Incubator/ssn/ssnx/ssn

M. Compton et al, "The SSN Ontology of the W3C Semantic Sensor Network Incubator Group", Journal of Web Semantics, 2012.

Several ontologies and description models



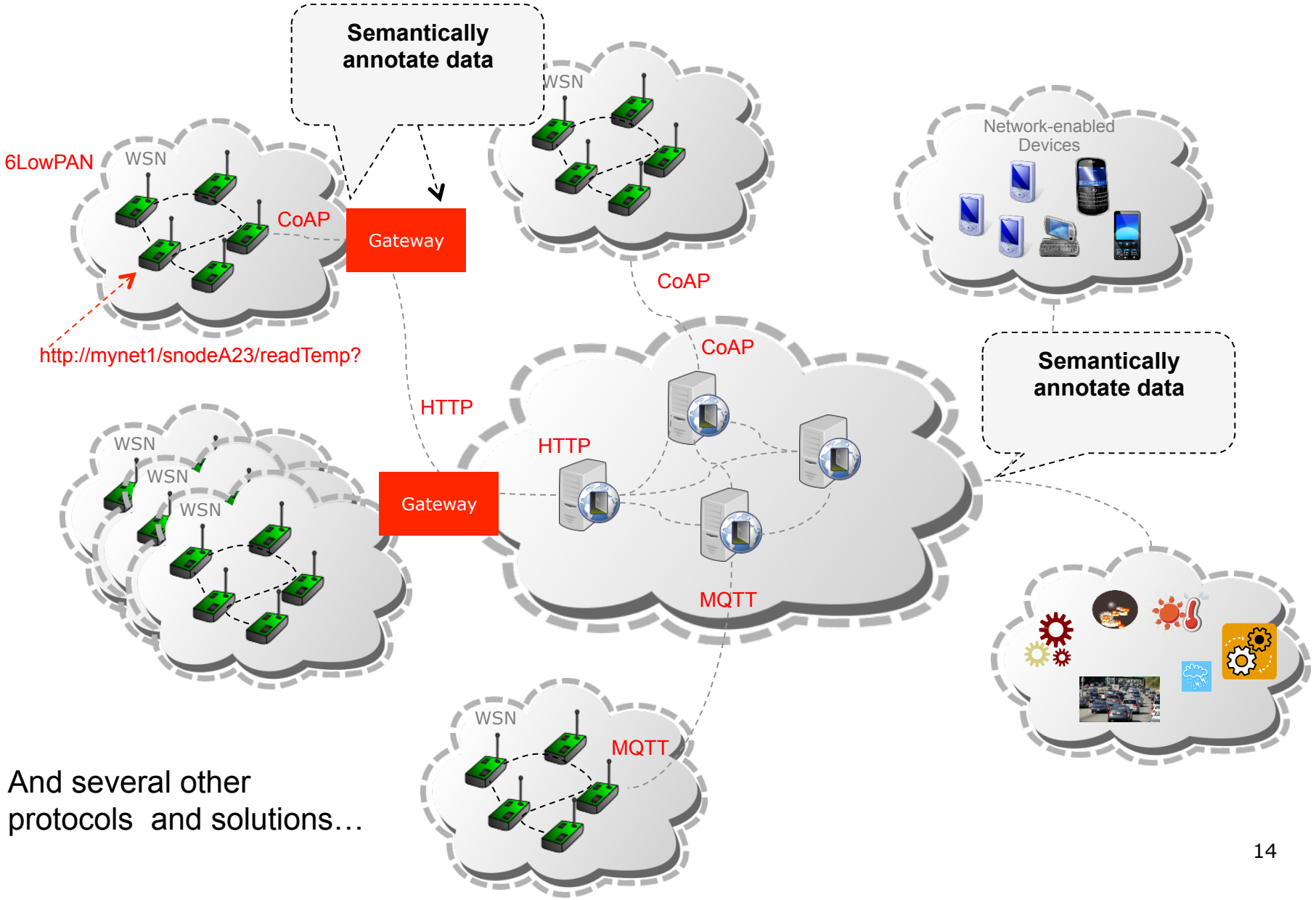
We have good models and description frameworks;

The problem is that having good models and developing ontologies is not enough.

Semantic descriptions are intermediary solutions, not the end product.

They should be transparent to the end-user and probably to the data producer as well.

A WoT/IoT Framework



And several other protocols and solutions…

Publishing Semantic annotations

- We need a model (ontology) – this is often the easy part for a single application.
- Interoperability between the models is a big issue.
- Express-ability vs Complexity is a challenge.

- How and where to add the semantics
- Where to publish and store them
- Semantic descriptions for data, streams, devices (resources) and entities that are represented by the devices, and description of the services.

Simplicity can be very useful…

Hyper/CAT

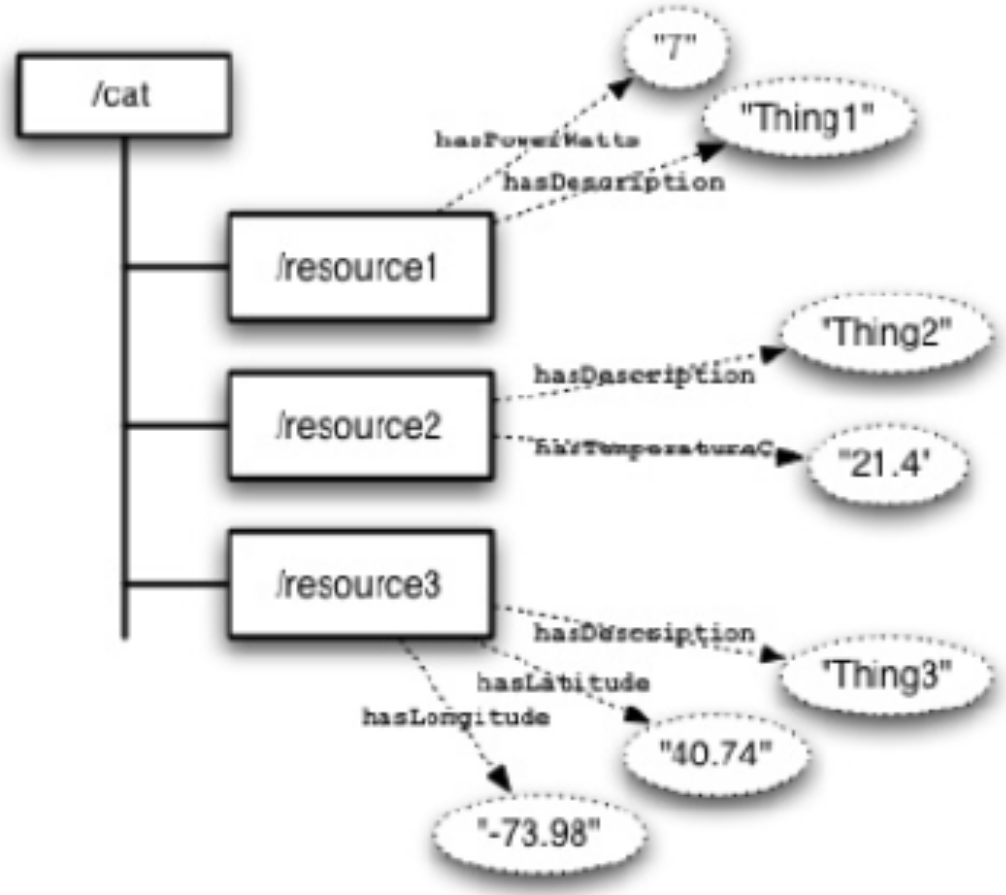
- Servers provide catalogues of resources to clients.

- A catalogue is an array of URIs.

- Each resource in the catalogue is annotated with metadata (RDF-like triples).

Source: Toby Jaffey, HyperCat Consortium, http://www.hypercat.io/standard.html

Hyper/CAT model



Source: Toby Jaffey, HyperCat Consortium, http://www.hypercat.io/standard.html

Complex models are (sometimes) good for publishing research papers….

But they are often difficult to implement and use in real world products.

What happens afterwards is more important

- How to index and query the annotated data
- How to make the publication suitable for constrained environments and/or allow them to scale
- How to query them (considering the fact that here we are dealing with live data and often reducing the processing time and latency is crucial)
- Linking to other sources

The IoT is a dynamic, online and rapidly changing world

UNIVERSITY OF SURREY

Annotation for the (Semantic) Web

isPartOf

Annotation for the IoT

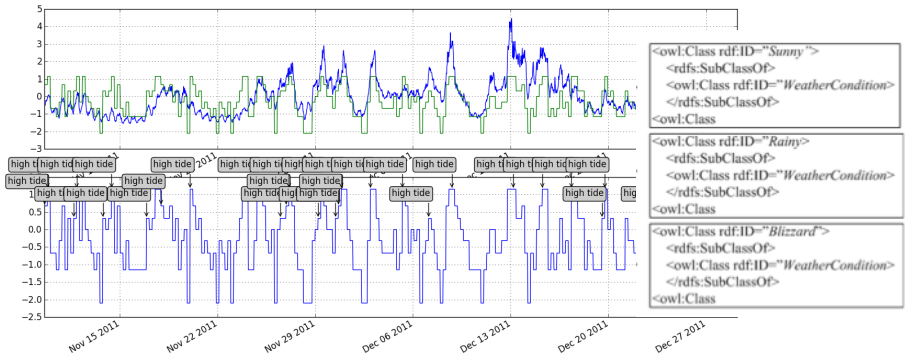
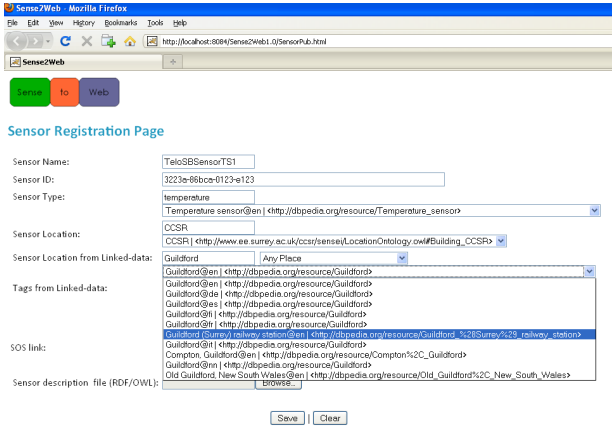
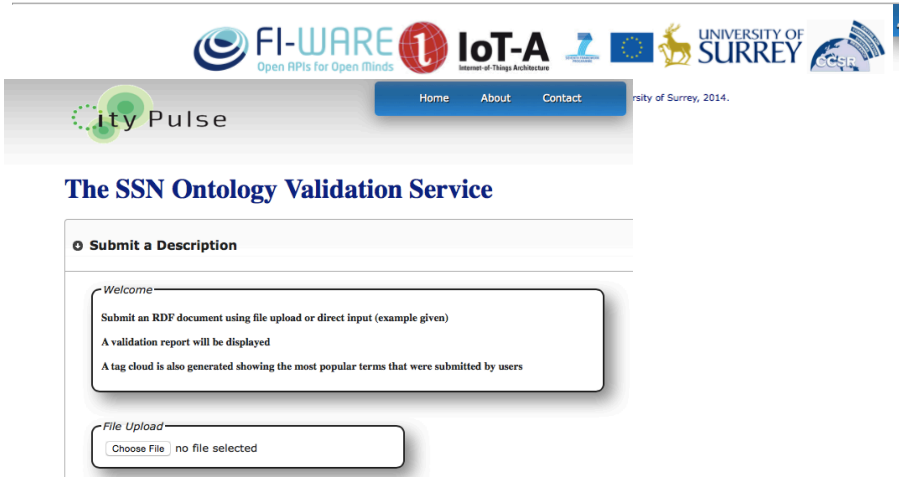
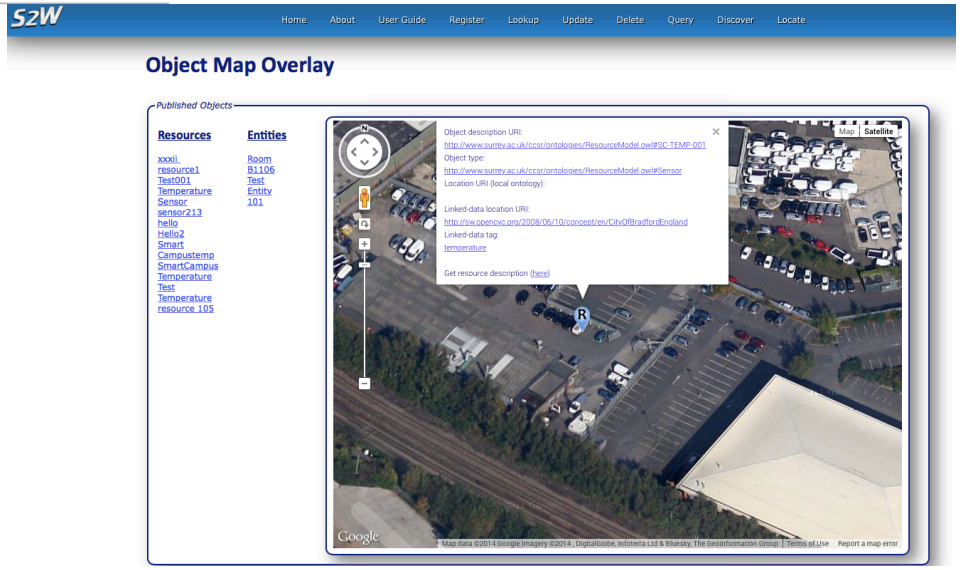


Image sources: ABC Australia and 2dolphins.com

Tools and APIs







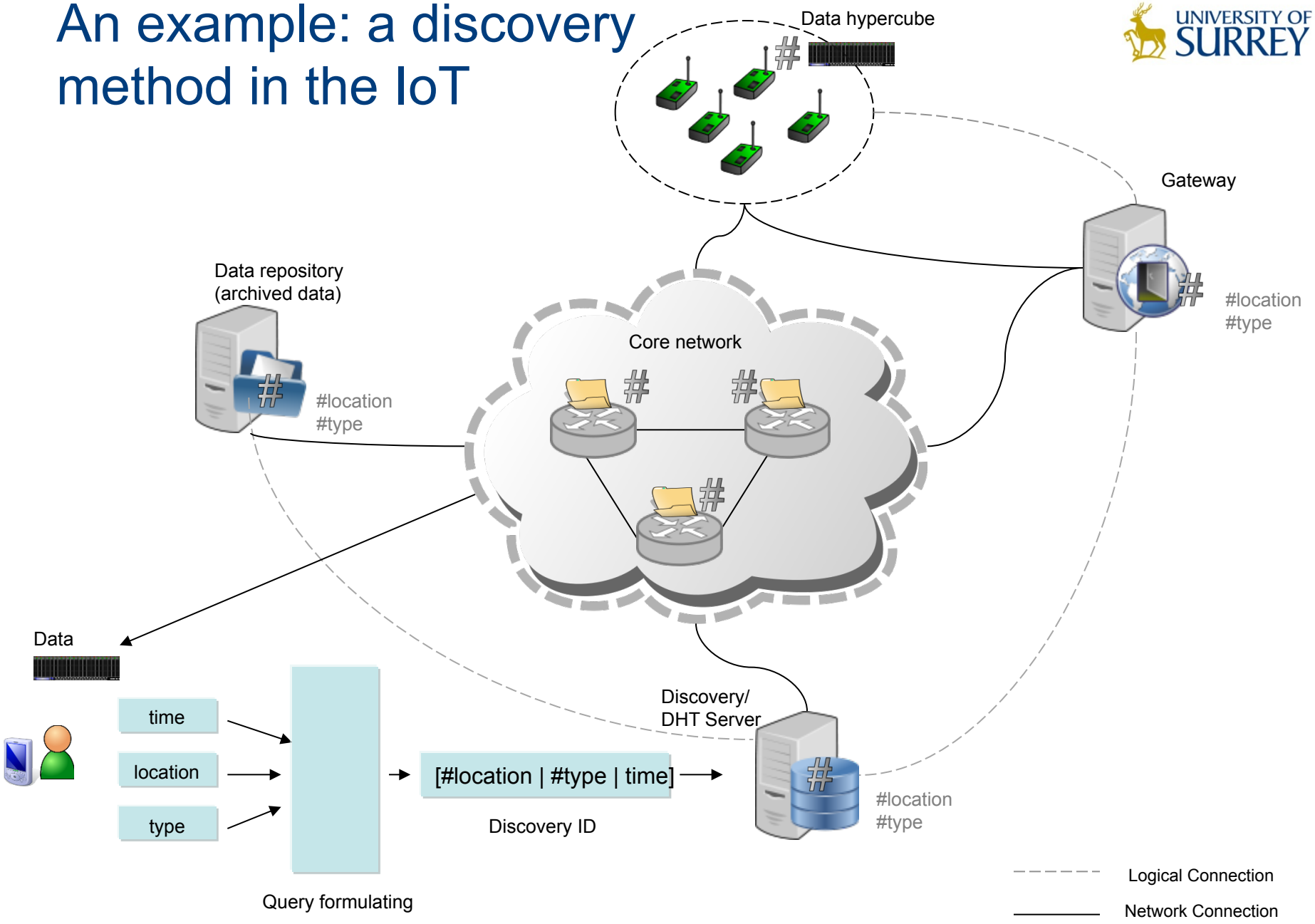
http://iot3.ee.surrey.ac.uk/s2w/

Creating common vocabularies and taxonomies are also equally important e.g. event taxonomies.

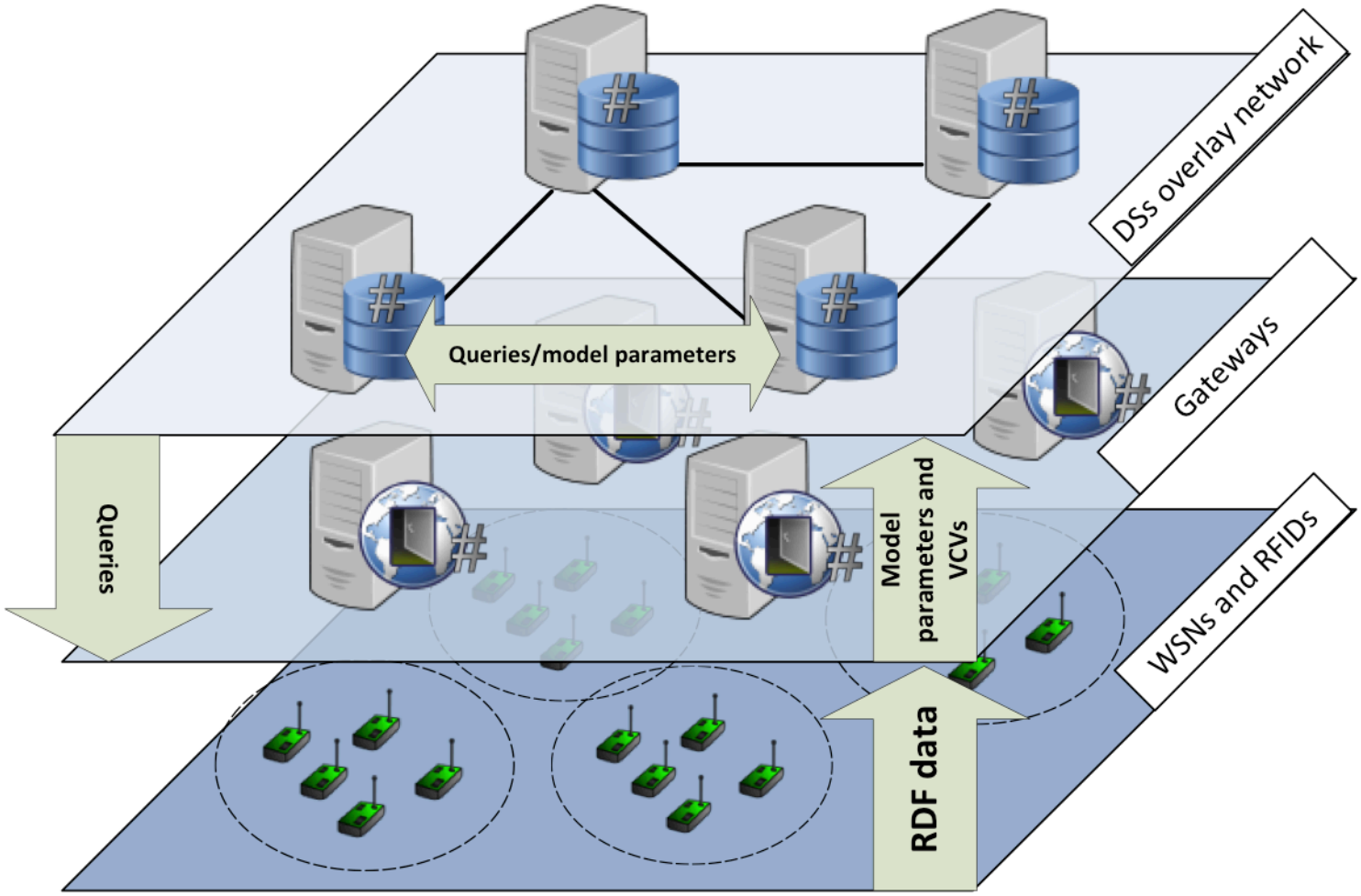
We should accept the fact that sometimes we do not need (full) semantic descriptions.

Think of the applications and use-cases before starting to annotate the data.

An example: a discovery method in the IoT

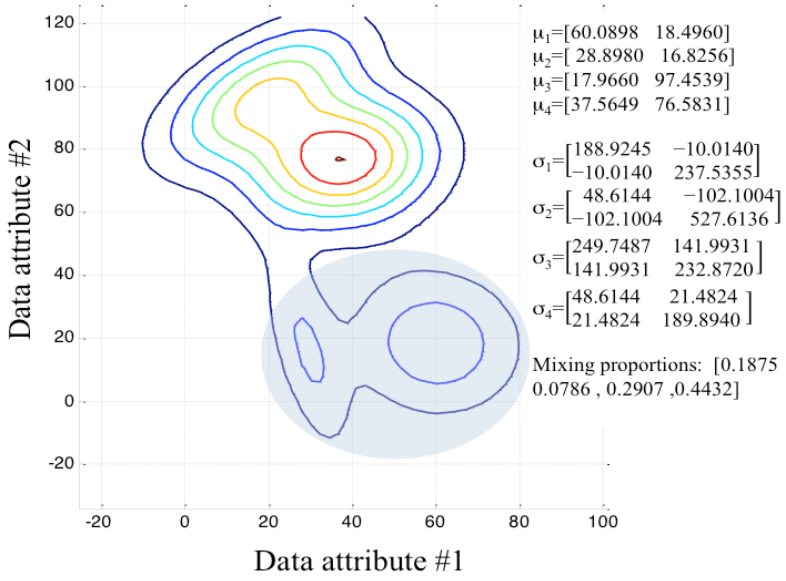


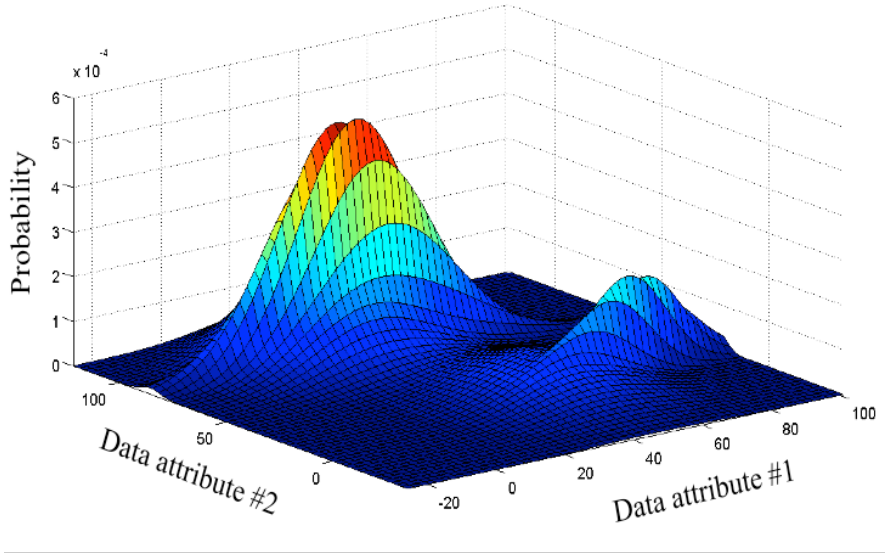
An example: a discovery method in the IoT



S. A. Hoseinitabatabaei, P. Barnaghi, C. Wang, R. Tafazolli, L. Dong, "A Distributed Data Discovery Mechanism for the Internet of Things", 2014.

An example: a discovery method in the IoT



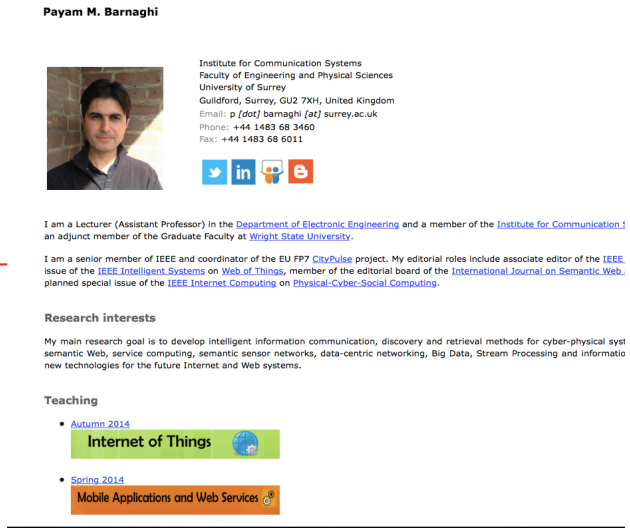


S. A. Hoseinitabatabaei, P. Barnaghi, C. Wang, R. Tafazolli, L. Dong, "A Distributed Data Discovery Mechanism for the Internet of Things", 2014.

Semantic descriptions can be fairly static on the Web;

In the IoT, the meaning of data and the annotations can change over time/ space…

Static Semantics



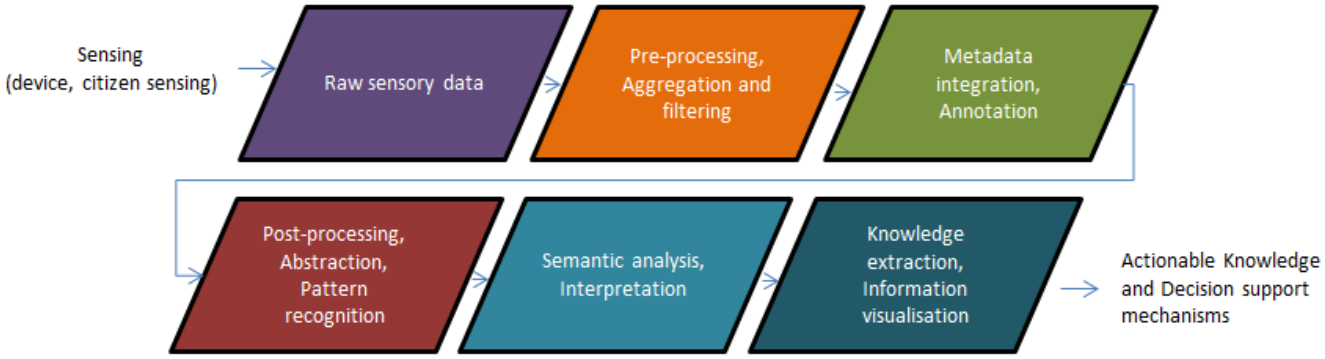

```
<rdf:RDF
      xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
      xmlns:foaf="http://xmlns.com/foaf/0.1/"
      xmlns:admin="http://webns.net/mvcb/">
<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:resource="#me"/>
  <foaf:primaryTopic rdf:resource="#me"/>
  <admin:generatorAgent rdf:resource="http://www.ldodds.com/foaf/foaf-a-matic"/>
  <admin:errorReportsTo rdf:resource="mailto:leigh@ldodds.com"/>
</foaf:PersonalProfileDocument>
<foaf:Person rdf:ID="me">
<foaf:name>Payam Barnaghi</foaf:name>
<foaf:givenname>Payam</foaf:givenname>
<foaf:family_name>Barnaghi</foaf:family_name>
<foaf:mbox_sha1sum>634c60aaa316463de0c93306ae21bb1d4d7ec556</foaf:mbox_sha1sum>
<foaf:depiction rdf:resource="http://personal.ee.surrey.ac.uk/Personal/P.Barnaghi/payam-
bw.jpg"/>
<foaf:phone rdf:resource="tel:+44-1483-68-9485"/>
<foaf:workplaceHomepage rdf:resource="http://personal.ee.surrey.ac.uk/Personal/P.Barnaghi/"/>
<foaf:workInfoHomepage rdf:resource="Researcher"/>
<foaf:schoolHomepage rdf:resource="http://www.ee.surrey.ac.uk/CCSR/"/></foaf:Person>
</rdf:RDF>
```

Dynamic Semantics

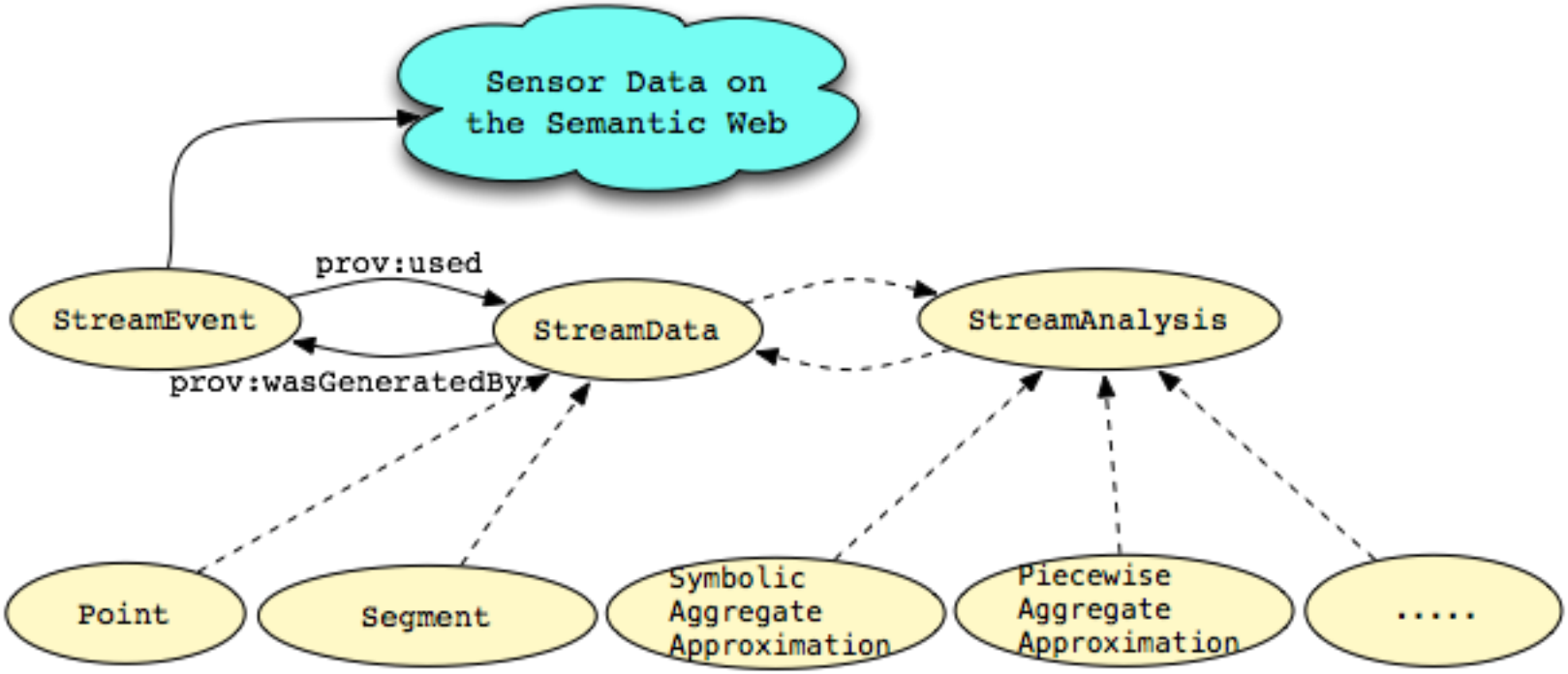
```
<iot:measurement>
<iot:type> temp</iot:type>
<iot:unit>Celsius</iot:unit>
<time>12:30:23UTC</time>
<iot:accuracy>80%</iot:accuracy>
<loc:long>51.2365<loc:lat>
<loc:lat>0.5703</loc:lat>
</iot:measurment>
```

But this could be a function of time and location;

What would be the accuracy 5 seconds after the measurement?



Dynamic annotations for data in the process chain

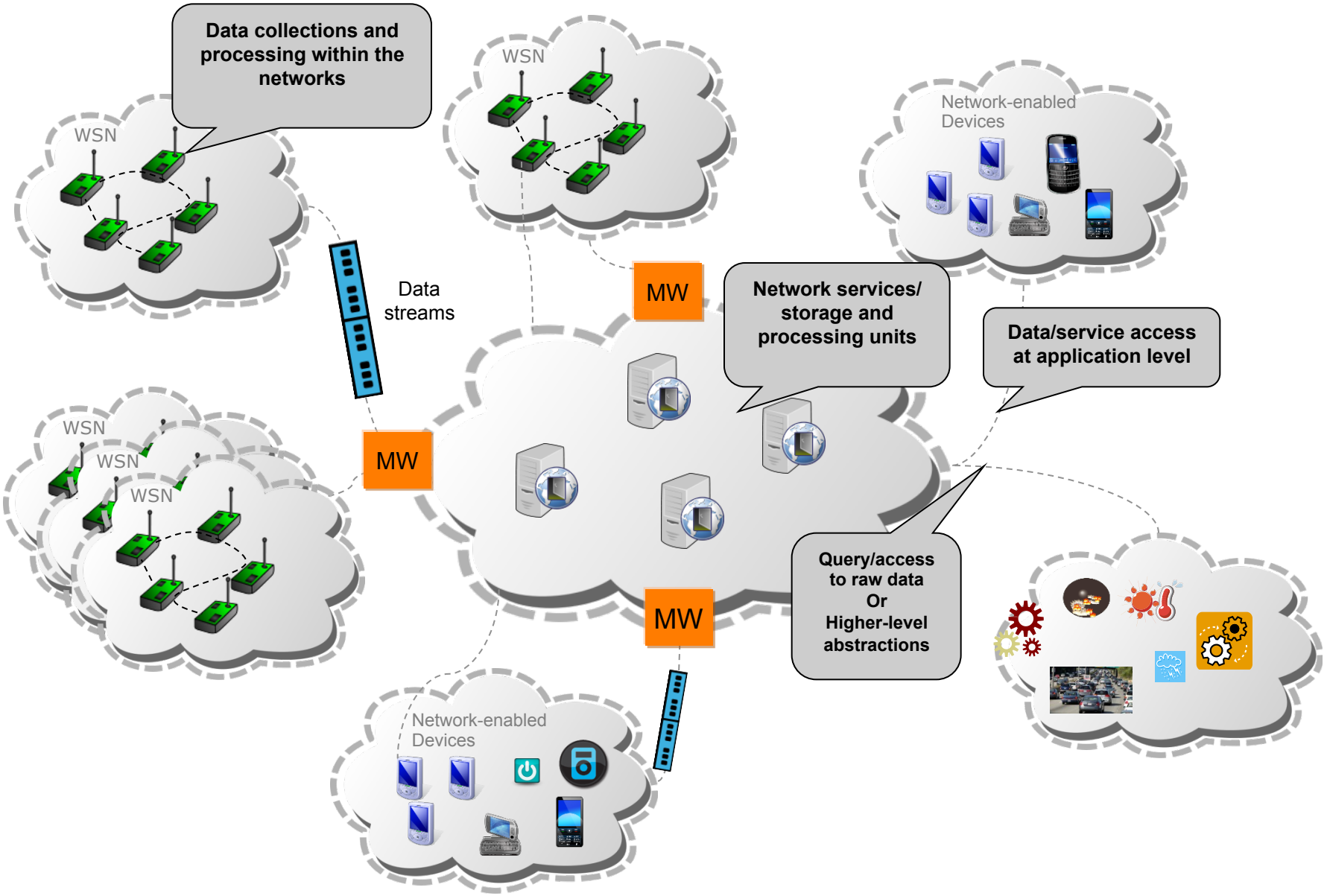


S. Kolozali et al, A Knowledge-based Approach for Real-Time IoT Data Stream Annotation and Processing", iThings 2014, 2014.

Overall, we need semantic technologies in the IoT and these play a key role in providing interoperability.

However, we should design and use the semantics carefully and consider the constraints and dynamicity of the IoT environments.

The IoT


Data collections and processing within the networks
WSN
WSN
Network-enabled Devices
Data streams
MW
Network services/ storage and processing units
Data/service access at application level
WSN
WSN
WSN
MW
MW
Query/access to raw data Or Higher-level abstractions
Network-enabled Devices

#1: Design for large-scale and provide tools and APIs.

#2: Think of who will use the semantics and how when you design your models.

#3: Provide means to update and change the semantic annotations.

#4: Create tools for validation and interoperability testing.

#5: Create taxonomies and vocabularies.

#6: Of course you can always create a better model, but try to re-use existing ones as much as you can.

#7: Link your data and descriptions to other existing resources.

#8: Define rules and/or best practices for providing the values for each attribute.

#9: Remember the widely used semantic descriptions on the Web are simple ones like FOAF.

In Conclusion

#10: Semantics are only one part of the solution and often not the end-product so the focus of the design should be on creating effective methods, tools and APIs to handle and process the semantics.

Query methods, machine learning, reasoning and data analysis techniques and methods should be able to effectively use these semantics.

Q&A

- Thank you.

- EU FP7 CityPulse Project:
  - http://www.ict-citypulse.eu/
  - @pbarnaghi
  - p.barnaghi@surrey.ac.uk

http://personal.ee.surrey.ac.uk/Personal/P.Barnaghi/