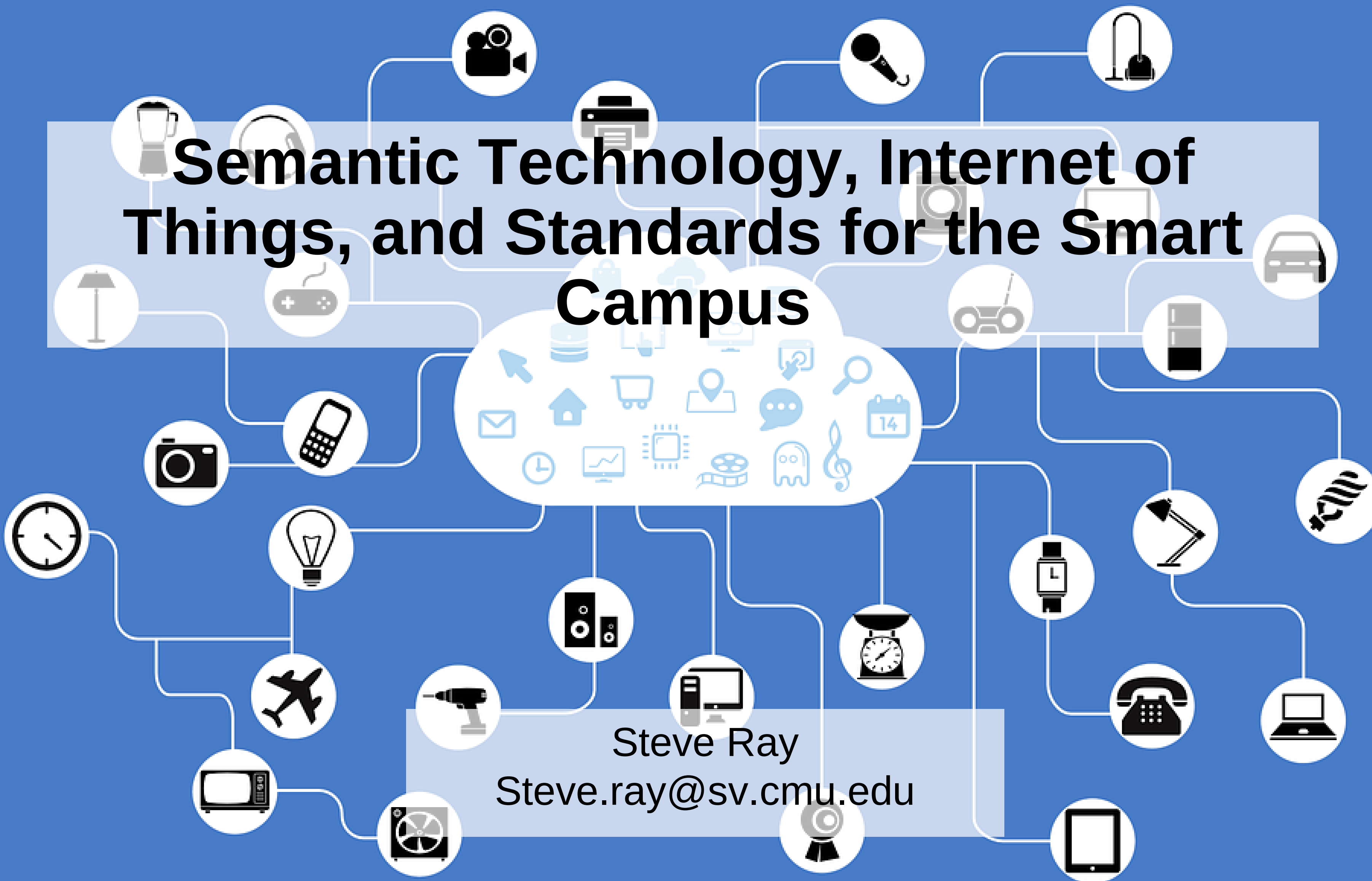
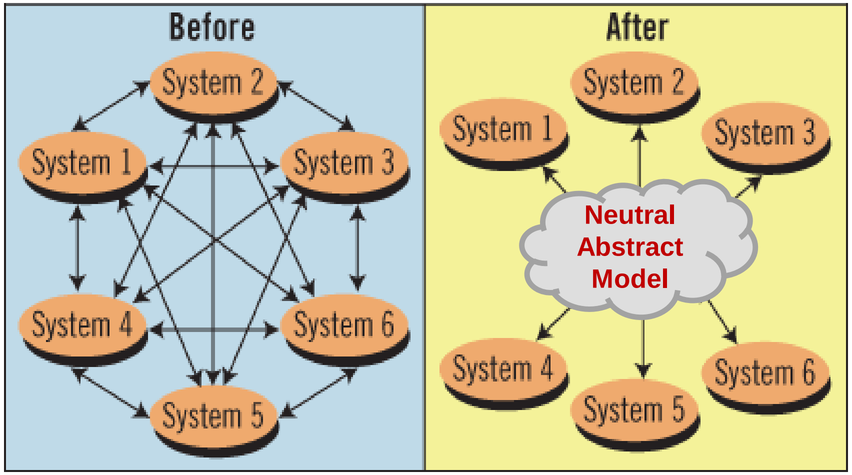


Semantic Technology, Internet of Things, and Standards for the Smart Campus

Steve Ray
Steve.ray@sv.cmu.edu

Carnegie Mellon University
Silicon Valley

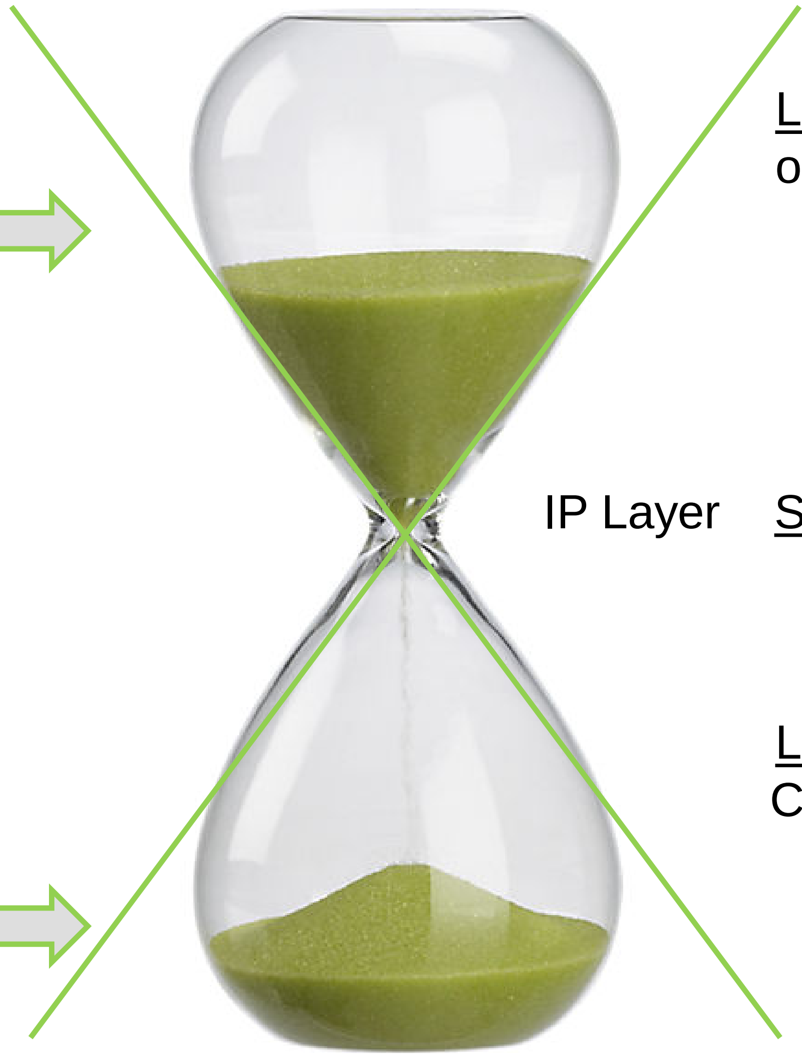
Smart Campus
Smart Community
Smart City
…



- We need to share information across different contexts (need information about the information)
- Information needs to be defined in a way that machines can reason over the definitions
- A neutral abstract model and semantic mappings give us a way to manage complexity
- Semantic web technology is well-suited to loosely-coupled, distributed, linked communities of systems

Connectivity ≠ Interoperability

Not so much attention



Lots of attention

<u>Lots of unproven choices</u> of Information content

IP Layer   <u>Strong agreement</u> here

<u>Lots of proven choices</u> of Communication Protocols

- Zigbee
- Wifi
- Power line
- Cellular

- “*Need smart data so you don’t need such smart software*”
  - Krzysztof Janowicz , UCSB

What's the issue?

- Misunderstanding of information flowing between systems
- "Set Thruster to 324.59"

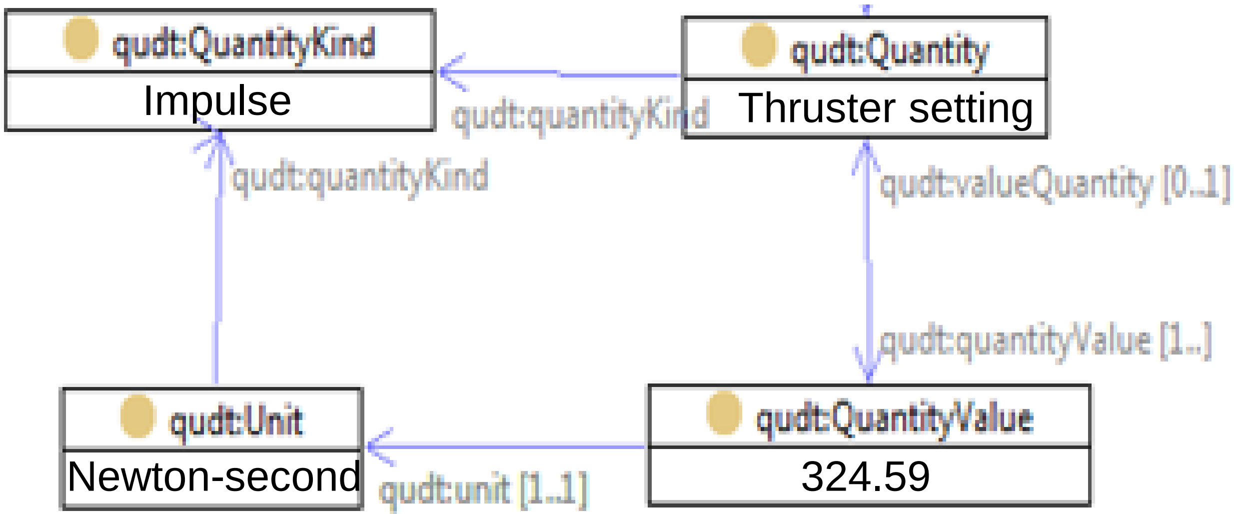
By Robin Lloyd
CNN Interactive Senior Writer

(CNN) -- NASA lost a $125 million Mars orbiter because a Lockheed Martin engineering team used English units of measurement while the agency's team used the more conventional metric system for a key spacecraft operation, according to a review finding released Thursday.

September, 1999

Need context (a.k.a. metadata)

“Thruster setting to Impulse with value 324.59 Newton-seconds”

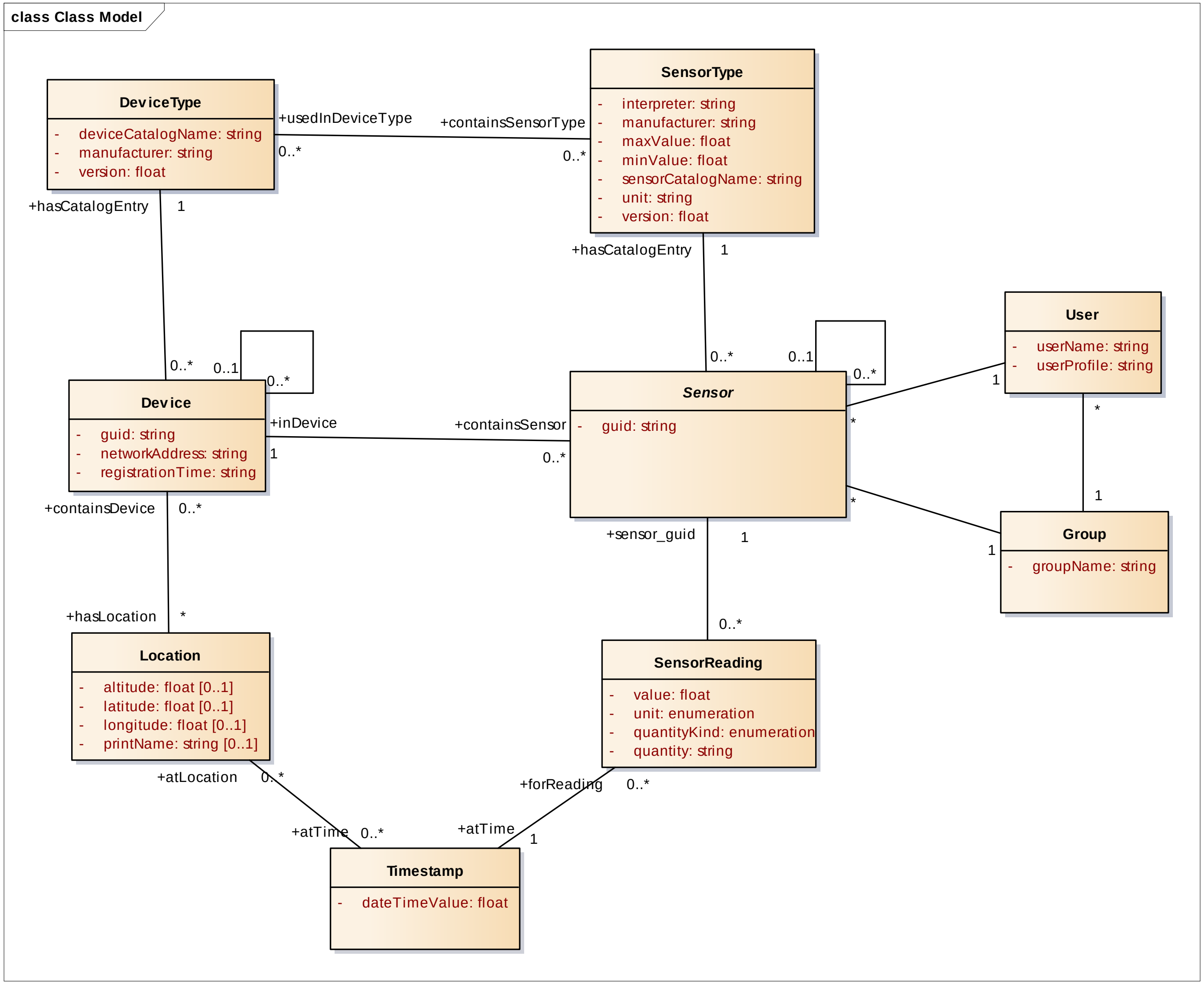


Whenever you see something like this:

GET SENSOR READING FROM A SENSOR(SPECIFIED BY DEVICE URI AND SENSOR TYPE NAME) AT A TIMESTAMP

- **Purpose**: Query sensor reading for a specific sensor at a specific time point.
- **Method**: GET
- **URL**: http://einstein.sv.cmu.edu:9000/getSensorReading/<"deviceUri">/<"sensorTypeName">/<"timestamp">/<"resultFormat">
- **Semantics**:
  - **deviceUri**: Existing device uri.
  - **sensorTypeName**: Existing sensor type name.
  - **timestamp**: Time of the readings to query.
  - **resultFormat**: Either JSON or CSV.

You should have access to something like this:



Carnegie Mellon University Silicon Valley

Context

☹ Many standards and API specs provide context through supporting documentation for human implementers to read – and possibly misinterpret.

☺ A good information model (or information standard) explicitly captures metadata (context)

– Information model, or	UML
– Semantic model, or	OWL/SHACL
– Formal logic model	First order logic

What are some minimal requirements needed to start improving the quality of existing information standards?

	Most standards today			Semantic Technology	
Capability	**Paper, .doc, .pdf**	**XML/XSD**	**UML**	**OWL/SHACL**	**FOL**
Machine readable		✓	✓	✓	✓
Queryable				✓	✓
Reasoning				✓	✓
Proofs					✓

Old-style (most common) standards specifications: (e.g. ISO 14258, Requirements for enterprise-reference architectures and methodologies)

"3.6.1.1 Time representation

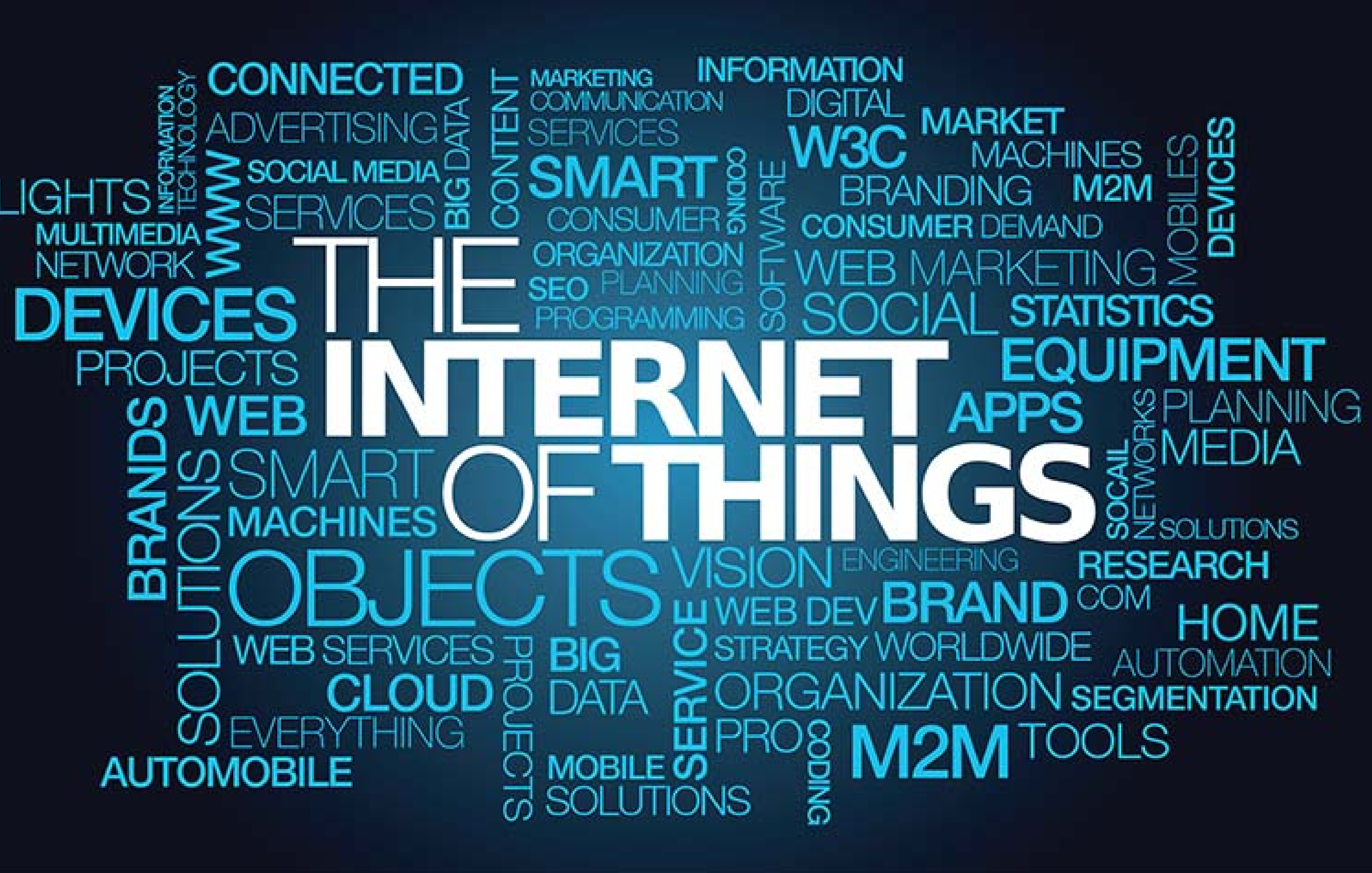
If an individual element of the enterprise system has to be traced then properties of time need to be modeled to describe short-term changes. If the property time is introduced in terms of duration, it provides the base to do further analyses (e.g., process time). There are two kinds of behavior description relative to time: static and dynamic."

Data-model standards (e.g. ISO 10303-41, Product Description and Support)

```
ENTITY product_context
  SUBTYPE OF (application_context_element);
  discipline_type : label;
END_ENTITY;
```

Semantic-model standards (e.g. ISO 18629-11, PSL Core)

```
(forall (?t1 ?t2 ?t3)
    (=>   (and   (before ?t1 ?t2)
                 (before ?t2 ?t3))
          (before ?t1 ?t3)))
```


THE INTERNET OF THINGS
CONNECTED
INFORMATION TECHNOLOGY
ADVERTISING
SOCIAL MEDIA
SERVICES
WWW
BIG DATA
CONTENT
MARKETING
COMMUNICATION
SERVICES
SMART
CONSUMER
ORGANIZATION
SEO
PLANNING
PROGRAMMING
CODING
SOFTWARE
INFORMATION
DIGITAL
W3C
MARKET
MACHINES
BRANDING
M2M
CONSUMER DEMAND
WEB MARKETING
SOCIAL
STATISTICS
MOBILES
DEVICES
LIGHTS
MULTIMEDIA
NETWORK
DEVICES
PROJECTS
WEB
BRANDS
SOLUTIONS
SMART
MACHINES
OBJECTS
EQUIPMENT
APPS
PLANNING
MEDIA
SOCIAL NETWORKS
SOLUTIONS
VISION
ENGINEERING
RESEARCH
WEB DEV
BRAND
COM
HOME
STRATEGY
WORLDWIDE
AUTOMATION
WEB SERVICES
PROJECTS
BIG
DATA
SERVICE
ORGANIZATION
SEGMENTATION
CLOUD
EVERYTHING
PRO
CODING
M2M
TOOLS
AUTOMOBILE
MOBILE
SOLUTIONS

Current's CityIQ @ Work

San Diego:
3,200 intelligent sensor nodes
14,000 new LED light fixtures expected to save $2.4 million in annual energy costs

Interoperability Standards Battles for the Internet of Things

Industrial Internet Consortium
AT&T, Cisco, IBM, Intel…
plus ~75 others

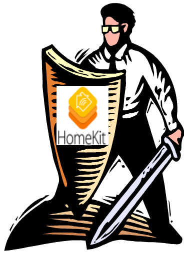
Open Interconnect Consortium
Intel, Samsung, Dell…
plus a few others

Thread Group
Google (Nest), ARM,
Samsung appliance…
plus a few others

?

AllSeen Alliance
Microsoft, Cisco,
Qualcomm, LG…
plus ~60 others

The Open Group



Apple Homekit

Interoperability Standards Battles for the Internet of Things

industrial internet® CONSORTIUM

AT&T, Cisco, IBM, Intel…
plus ~165 others

oneM2M

IoT-A Internet of Things - Architecture

SIoT Open Source Internet of Things

OPEN INTERCONNECT CONSORTIUM℠

Intel, Samsung, GE, Cisco, Dell…
plus ~50 others

INTERNET OF THINGS IoT-GSI GLOBAL STANDARDS INITIATIVE ITU-T

Thread Group
Google (Nest), ARM,
Samsung appliance…
plus ~120 others

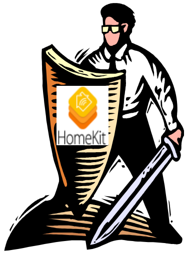
?

ALLSEEN ALLIANCE

Microsoft, Cisco,
Qualcomm, Sony, LG…
plus ~140 others

THE Open GROUP

Hewlett-Packard, Cisco,
IBM, Oracle, Philips
Plus ~400 others



Apple Homekit
“*dozens of partners*”
as of May 14, 2015

IPSO Alliance

OASIS
Advancing open standards for the information society

Interoperability Standards Battles for the Internet of Things

industrial internet® CONSORTIUM

AT&T, Cisco, IBM, Intel…
plus ~250 others

Thread Group
Google (Nest), ARM,
Samsung appliance…
plus ~240 others

THE Open GROUP
Hewlett-Packard, Cisco,
IBM, Oracle, Philips
Plus ~400 others

oneM2M

OGC® Making location count.

IoT-A Internet of Things - Architecture

SIoT Open Source Internet of Things

OPEN CONNECTIVITY FOUNDATION™
Intel, Qualcomm, Microsoft, Samsung,
GE, Cisco, Dell, ARRIS, Electrolux…
plus ~400 others

INTERNET OF THINGS IoT-GSI GLOBAL STANDARDS INITIATIVE ITU-T

ALLSEEN ALLIANCE
Microsoft, Cisco,
Qualcomm, Sony, LG…
plus ~140 others

Apple Homekit
139 product lines
as of Sep. 2017

IPSO Alliance

OASIS Advancing open standards for the information society

Industrial ⟷ **Residential**

Interoperability Standards Battles for the Internet of Things

industrial internet® CONSORTIUM
AT&T, Cisco, IBM, Intel…
plus ~250 others

oneM2M

OGC® Making location count.

IoT-A Internet of Things - Architecture

SIoT Open Source Internet of Things

OPEN CONNECTIVITY FOUNDATION™
Intel, Qualcomm, Microsoft, Samsung, GE, Cisco, Dell, ARRIS, Electrolux…
plus ~400 others

INTERNET OF THINGS IoT-GSI GLOBAL STANDARDS INITIATIVE ITU-T

Thread Group
Google (Nest), ARM, Samsung appliance…
plus ~240 others

THE Open GROUP
Hewlett-Packard, Cisco, IBM, Oracle, Philips
Plus ~400 others

ALLSEEN ALLIANCE
Microsoft, Cisco, Qualcomm, Sony, LG…
plus ~140 others

Apple Homekit
139 product lines as of Sep. 2017

IPSO Alliance

OASIS
Advancing open standards for the information society

Industrial ⟷ **Residential**

Challenge

- How to quickly and easily integrate a new data source into a generalized system?
  - Possibly without any model for the new data
  - New data may use foreign terminology
  - Any implicit model of the data may not align with the receiving system

Semantic data integration for Internet of Things

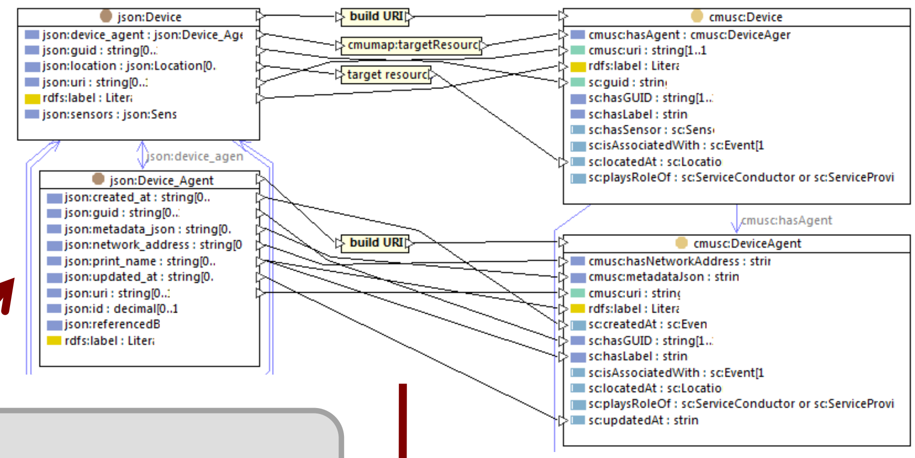
CMUSV Sensor Data

NASA Sensor Data

NICT (Japan) Sensor Data

⋮

- Retrieve data
- Import or synthesize ontologies
- Map to abstract ontology



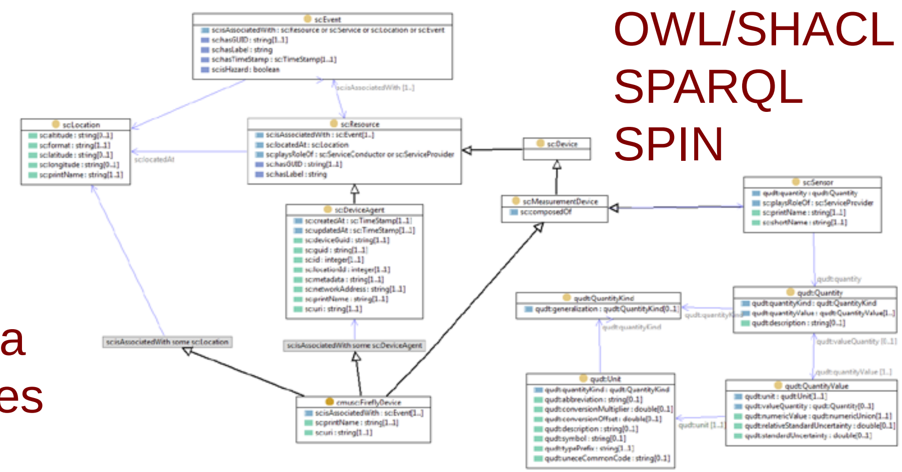
Technologies

RDF

OWL/SHACL

SPARQL

SPIN



Also Enables

Inferencing over data

Federated data stores

Distributed queries

Sample input data

- {"uri":"23420ca4e4830bee","deviceUserDefinedFields":"For test","location":{"longitude":10.123,"latitude":10.123,"altitude":10.123,"representation":"test location description"},"sensorNames":["fireImpXAccelerometer23420ca4e4830bee","fireImpYAccelerometer23420ca4e4830bee","fireImpZAccelerometer23420ca4e4830bee","fireImpMotion23420ca4e4830bee","fireImpLight23420ca4e4830bee","fireImpPressure23420ca4e4830bee","fireImpHumidity23420ca4e4830bee","fireImpDigitalTemperature23420ca4e4830bee"],"deviceTypeName":"FireImp V1C","manufacturer":"FireImp","version":"1C","deviceTypeUserDefinedFields":"","sensorTypeNames":["fireImpXAccelerometer","fireImpYAccelerometer","fireImpZAccelerometer","fireImpMotion","fireImpDigitalTemperature","fireImpLight","fireImpPressure","fireImpHumidity"]},

Interpreted as JSON Classes

- json:Device (10)
- json:DeviceType (10)
- json:Location (10)
- json:Reading (11)
- json:Sensor (25)
- json:SensorType (16)

- json:altitude
- json:deviceTypeName
- json:deviceTypeUserDefinedFields
- json:deviceUri
- json:deviceUserDefinedFields
- json:interpreter
- json:isIndoor
- json:latitude
- json:location
- json:longitude
- json:manufacturer
- json:maximumValue
- json:minimumValue
- json:purpose
- json:referencedBy
- json:representation
- json:sensorCategoryName
- json:sensorName
- json:sensorNames
- json:sensorTypeName
- json:sensorTypeNames
- json:sensorTypeUserDefinedFields
- json:sensorUserDefinedFields
- json:timeStamp
- json:unit
- json:uri
- json:value
- json:version

Class Form

Name: json:Reading

Annotations

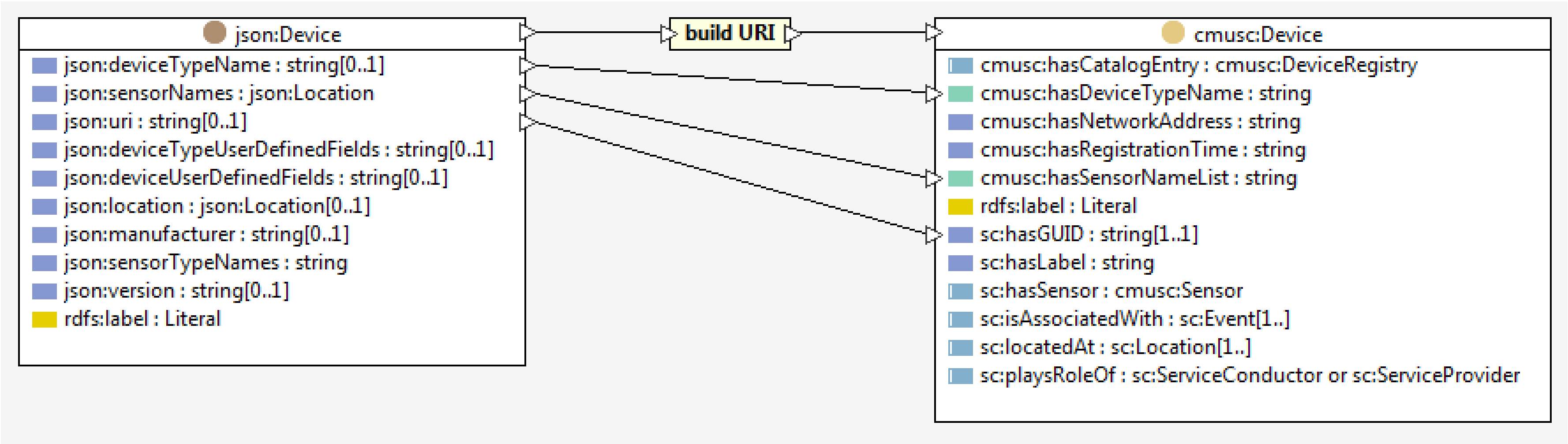
Class Axioms

rdfs:subClassOf

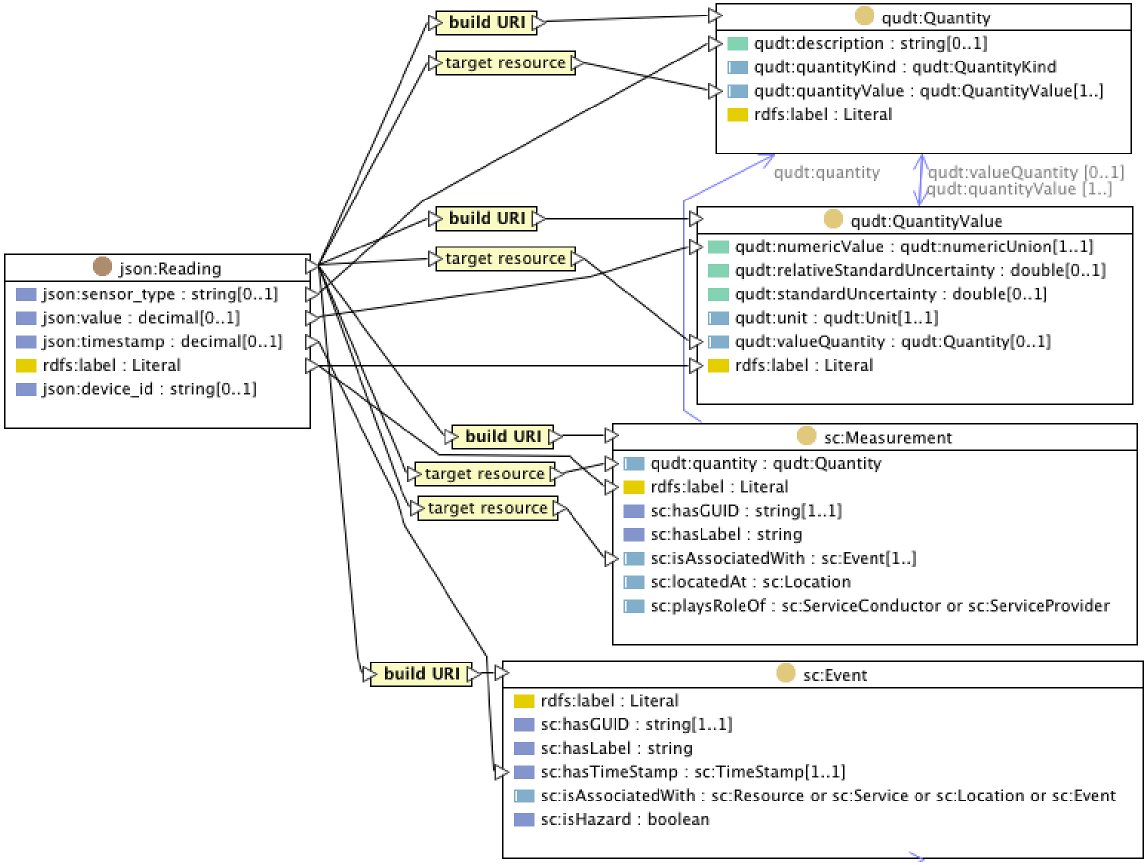
- owl:Thing
- json:isIndoor **max** 1
- json:isIndoor **only** xsd:boolean
- json:sensorName **max** 1
- json:sensorName **only** xsd:string
- json:timeStamp **max** 1
- json:timeStamp **only** xsd:string
- json:value **max** 1
- json:value **only** xsd:string

Manually Map JSON Entities to Target Ontology (the one manual step)

Sometimes trivial…

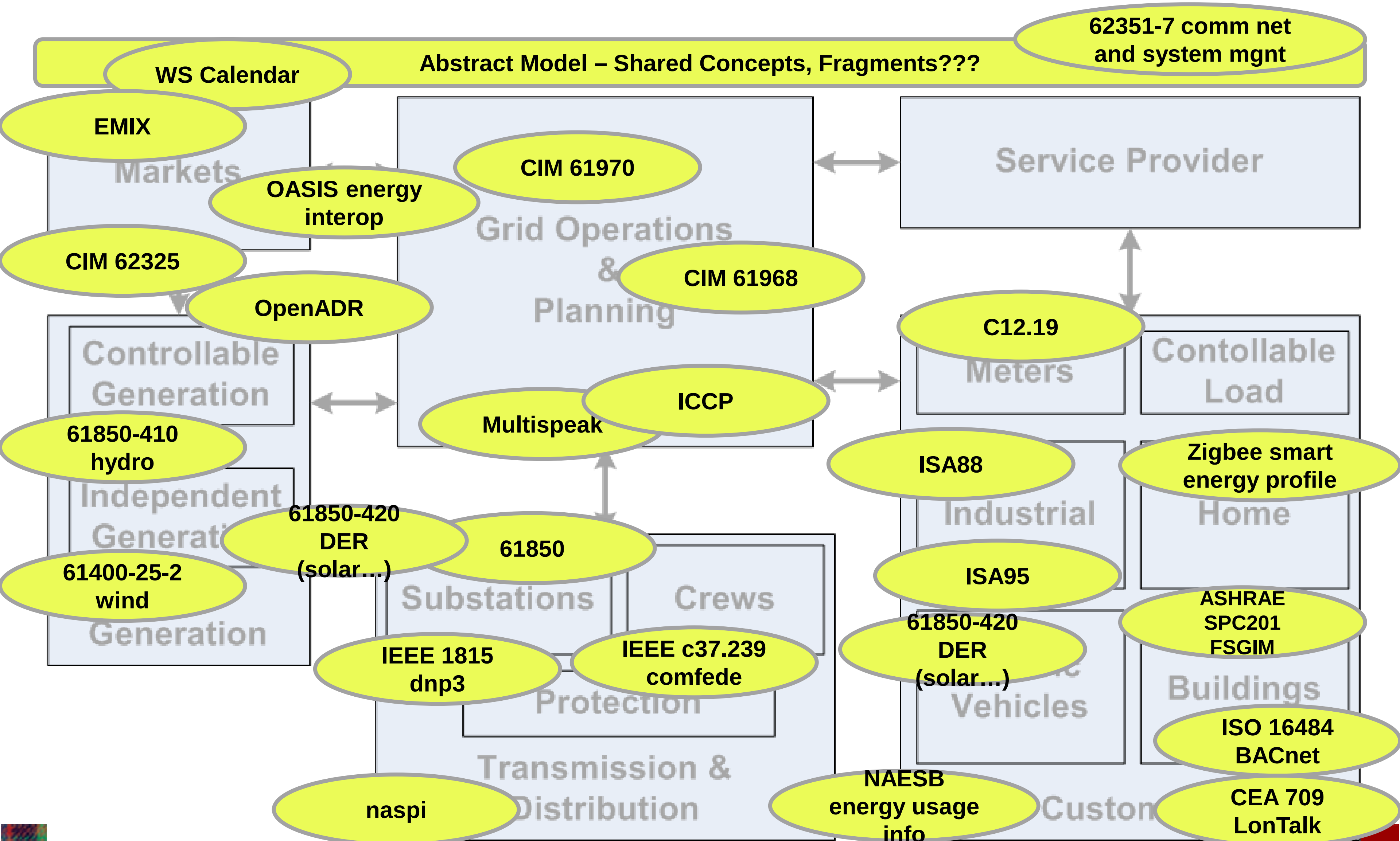


…sometimes more complex



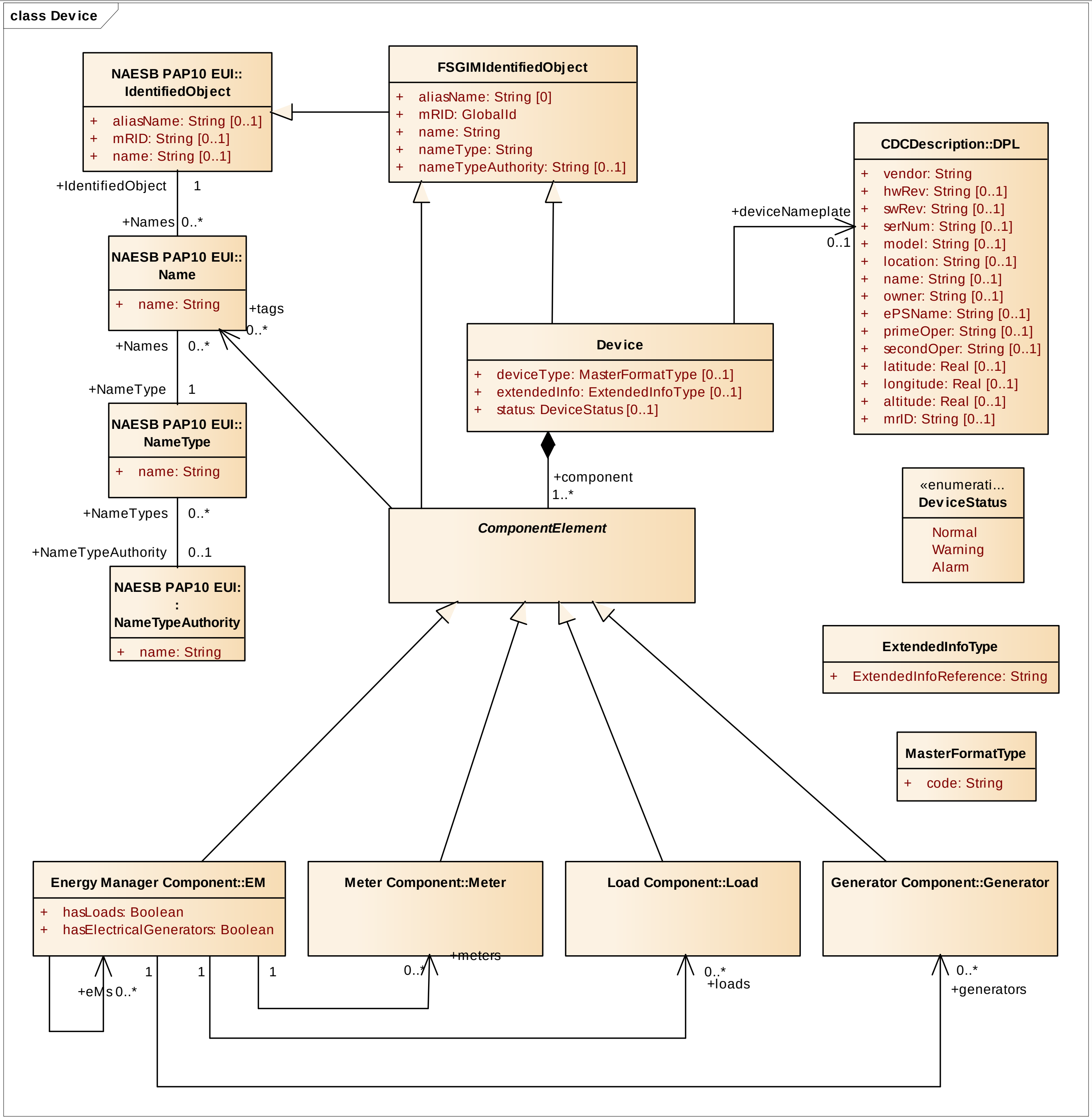
Smart Grid

Some Smart Grid Standards



UML Diagram:

Part of ASHRAE Standard SPC201 “Facility Smart Grid Information Model” (FSGIM)



A Power Aggregation Rule

Ruleset

DemandRuleset

name = DemandRuleset
nameType = Standard Rulesets
nameTypeAuthority = ASHRAE 201 Standard

notes

Calculate “metered demand”:

- For each direct subordinate that is an instance of the ElectricMeter Class where *only*

1. direct subordinate instances of the Load Class (including instances of the CurtailableLoad Class) and/or
2. direct subordinate instances of the EM Class where EM.hasElectricalGenerators is False

are connected to its output ConnectionPoint and where ElectricMeter.powerReading is available, sum ElectricMeter.powerReading.

Calculate “demand from unmetered loads”:

- For each direct subordinate that is an instance of the Load Class (including instances of the CurtailableLoad Class) that is not included in "metered demand" and where Load.actualDemand is available, sum Load.actualDemand.

Calculate “demand from energy managers”:

- For each direct subordinate that is an instance of the EM Class that is not included in "metered demand" and where EMPresentData.presentAggregateDemand is available, sum the subordinate EMPresentData.presentAggregateDemand.aggregateQuantity.

Calculate final result:

- Sum “metered demand”, “demand from unmetered loads”, and “demand from energy managers.”

Class Form

Name: collections_rulesets_and_aggregations---energy_manager_component:Aggregation

Annotations

rdfs:comment

An aggregation allows the assembly of statistics of interest arrived at through defined algorithms applied to other parts of the model. The description of an aggregation includes all of the inputs that comprise the algorithms as well as the outputs of the algorithm. An aggregation instance represents one or more aggregateQuantities that are computed by applying aggregation rules for each quantity against a Collection or FilteredCollection of Components.

rdfs:label

Class Axioms

rdfs:subClassOf

- device---device_and_model_components:FSGIMIdentifiedObject
- collections_rulesets_and_aggregations---energy_manager_component:hasAggregateQuantityRef **exactly** 1
- collections_rulesets_and_aggregations---energy_manager_component:hasAggregateQuantityRef **only** measurements---common_primitive_types_classes_and_enumerations:Measurement
- collections_rulesets_and_aggregations---energy_manager_component:hasAggregationRulesetRef **exactly** 1
- collections_rulesets_and_aggregations---energy_manager_component:hasAggregationRulesetRef **only** collections_rulesets_and_aggregations---energy_manager_component:Ruleset
- collections_rulesets_and_aggregations---energy_manager_component:hasCollectionRef **min** 0
- collections_rulesets_and_aggregations---energy_manager_component:hasCollectionRef **only** collections_rulesets_and_aggregations---energy_manager_component:Collection

Other Properties

spin:ruleWave1

```
# 001 Assign NameType of "Standard Aggregations"
CONSTRUCT {
  ?this device---device_and_model_components:hasNameType "Standard Aggregations" .
}
WHERE {
  FILTER NOT EXISTS {
    ?this device---device_and_model_components:hasNameType ?val .
  } .
}
```

```
# 002 Copy hasAggregateQuantityRef from the redefined property to the overwritten property to avoid a constraint violation
CONSTRUCT {
  ?this collections_rulesets_and_aggregations---energy_manager_component:hasAggregateQuantityRef ?value .
}
WHERE {
  NOT EXISTS {
    ?this collections_rulesets_and_aggregations---energy_manager_component:hasAggregateQuantityRef ?existingVal .
  } .
  ?this a ?type .
  ?type rdfs:subClassOf ?restriction .
```

Ruleset

DemandRuleset

name = DemandRuleset
nameType = Standard Rulesets
nameTypeAuthority = ASHRAE 201 Standard

notes

Calculate "metered demand":

- For each direct subordinate that is an instance of the ElectricMeter Class where *only*
  1. direct subordinate instances of the Load Class (including instances of the CurtailableLoad Class) and/or
  2. direct subordinate instances of the EM Class where EM.hasElectricalGenerators is False

are connected to its output ConnectionPoint and where ElectricMeter.powerReading is available, sum ElectricMeter.powerReading.

Calculate "demand from unmetered loads":

- For each direct subordinate that is an instance of the Load Class (including instances of the CurtailableLoad Class) that is not included in "metered demand" and where Load.actualDemand is available, sum Load.actualDemand.

Calculate "demand from energy managers":

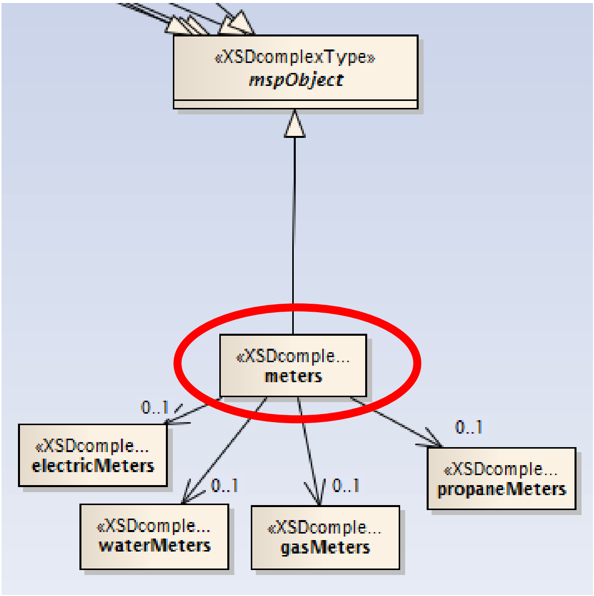
- For each direct subordinate that is an instance of the EM Class that is not included in "metered demand" and where EMPresentData.presentAggregateDemand is available, sum the subordinate EMPresentData.presentAggregateDemand.aggregateQuantity.

Calculate final result:

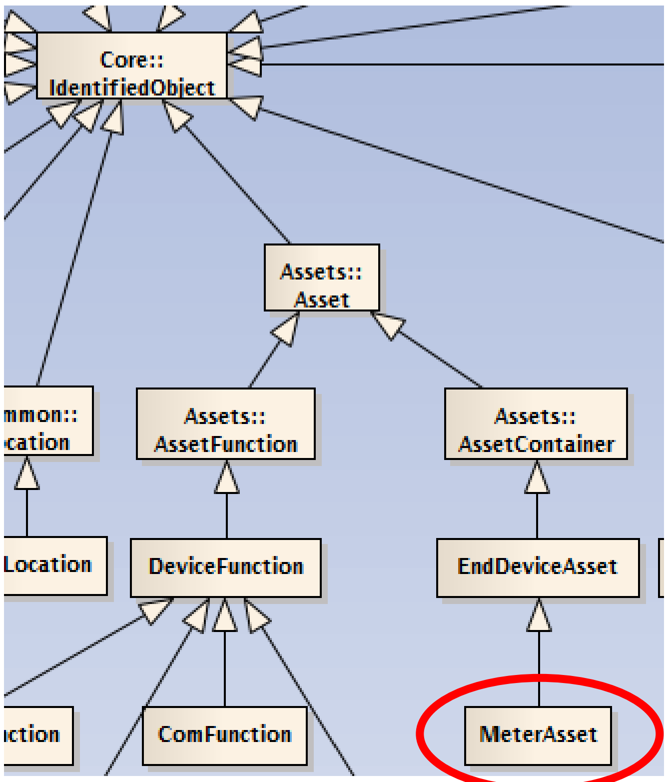
- Sum "metered demand", "demand from unmetered loads", and "demand from energy managers."

"Meter"
Do they all really mean the same thing?

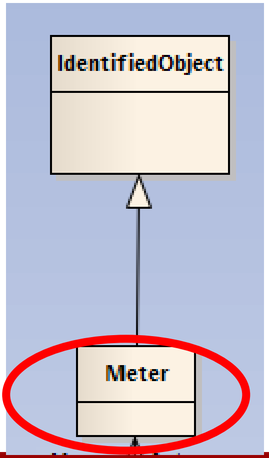
Multispeak V4.1



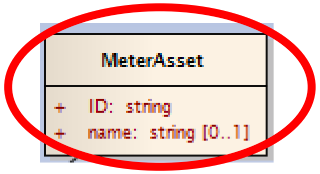
IEC 61968



IEC 61970



NAESB PAP10



Carnegie Mellon University Silicon Valley

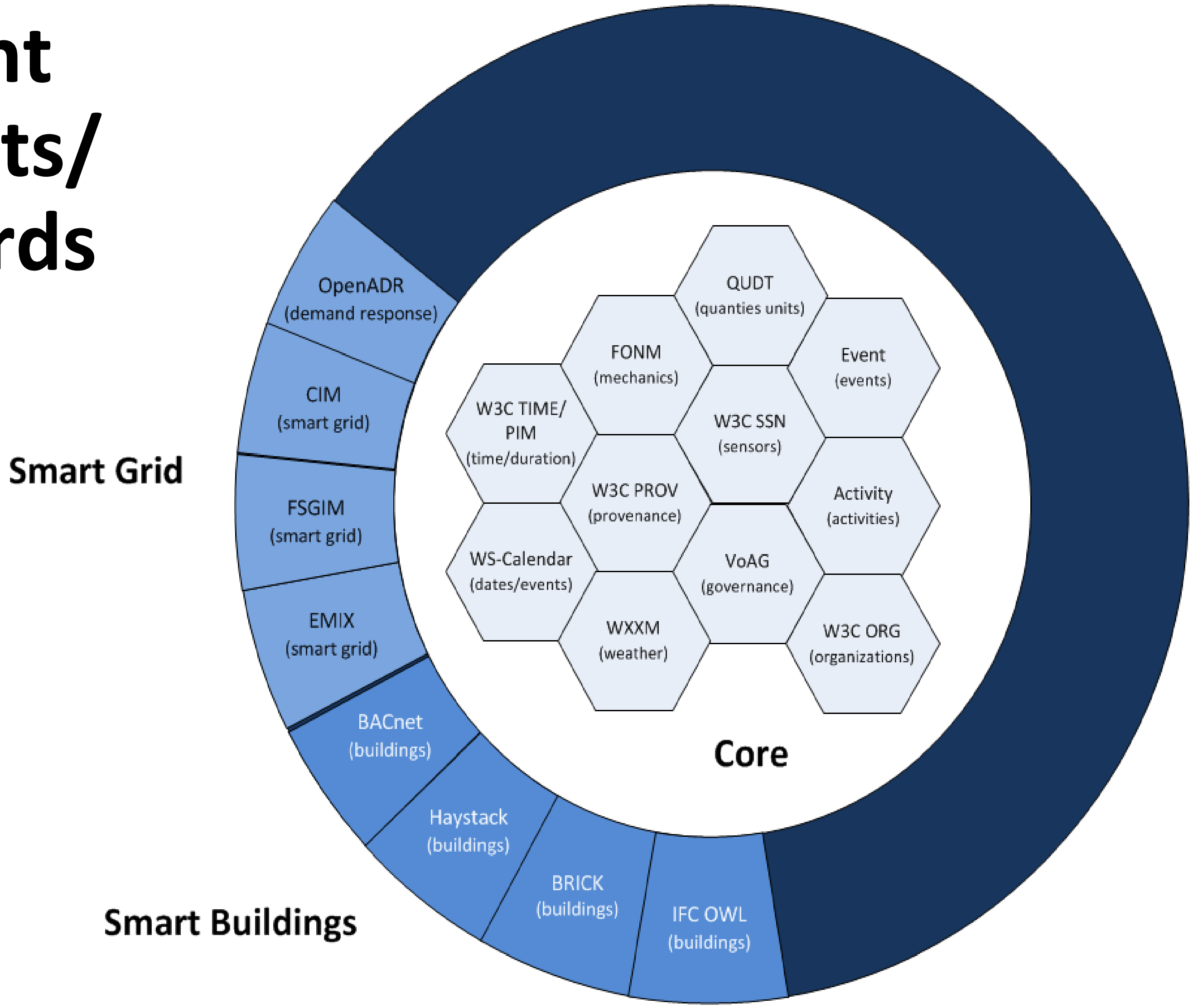
What to do?

✗ Require all standards to use the same universal definitions for all terms

✓ Require all standards to explicitly define all terms (in machine-readable form)

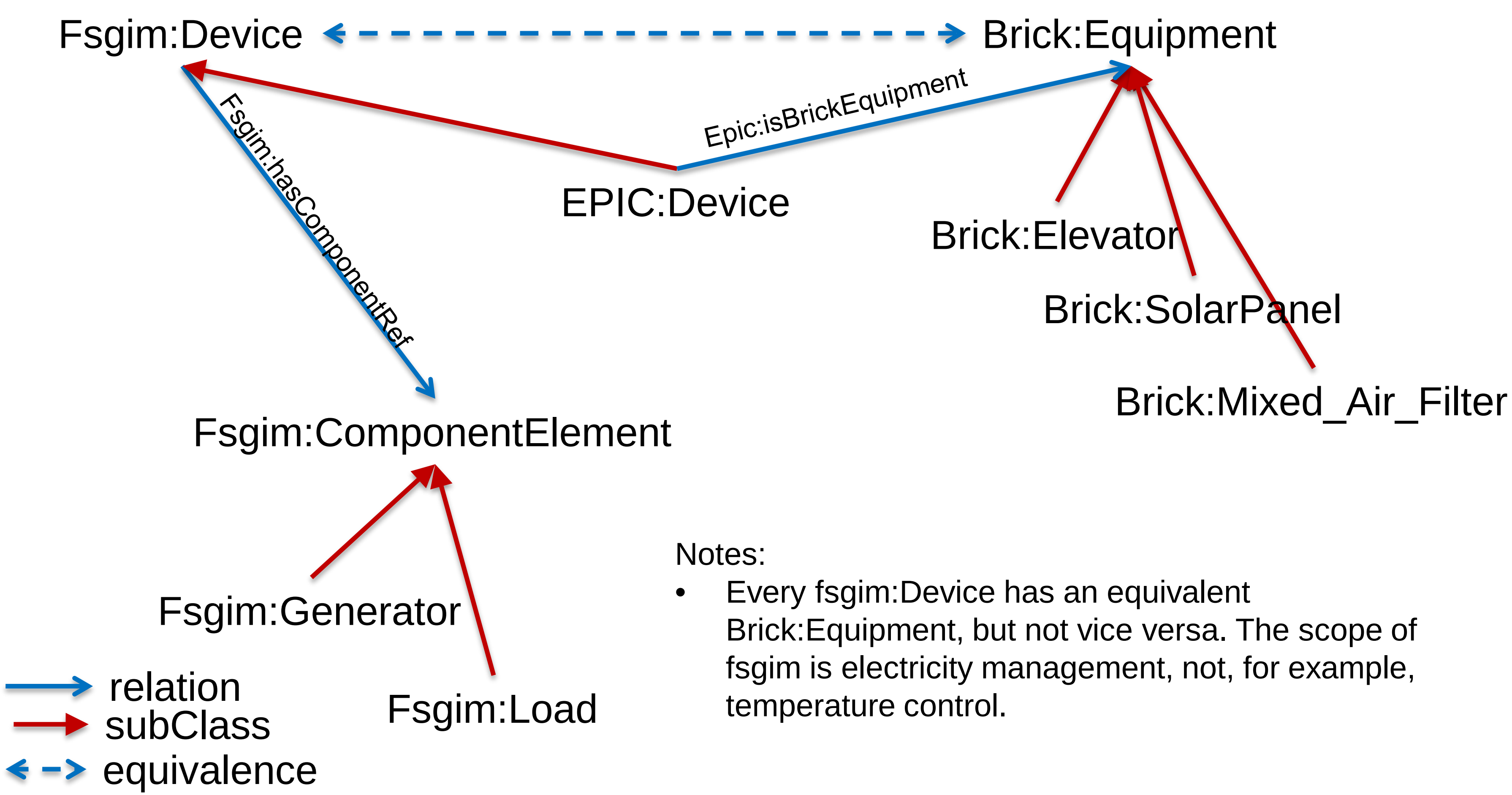
✓ Use technology to automatically map between terms in one standard to terms in another

Relevant Concepts/ Standards



Carnegie Mellon University Silicon Valley

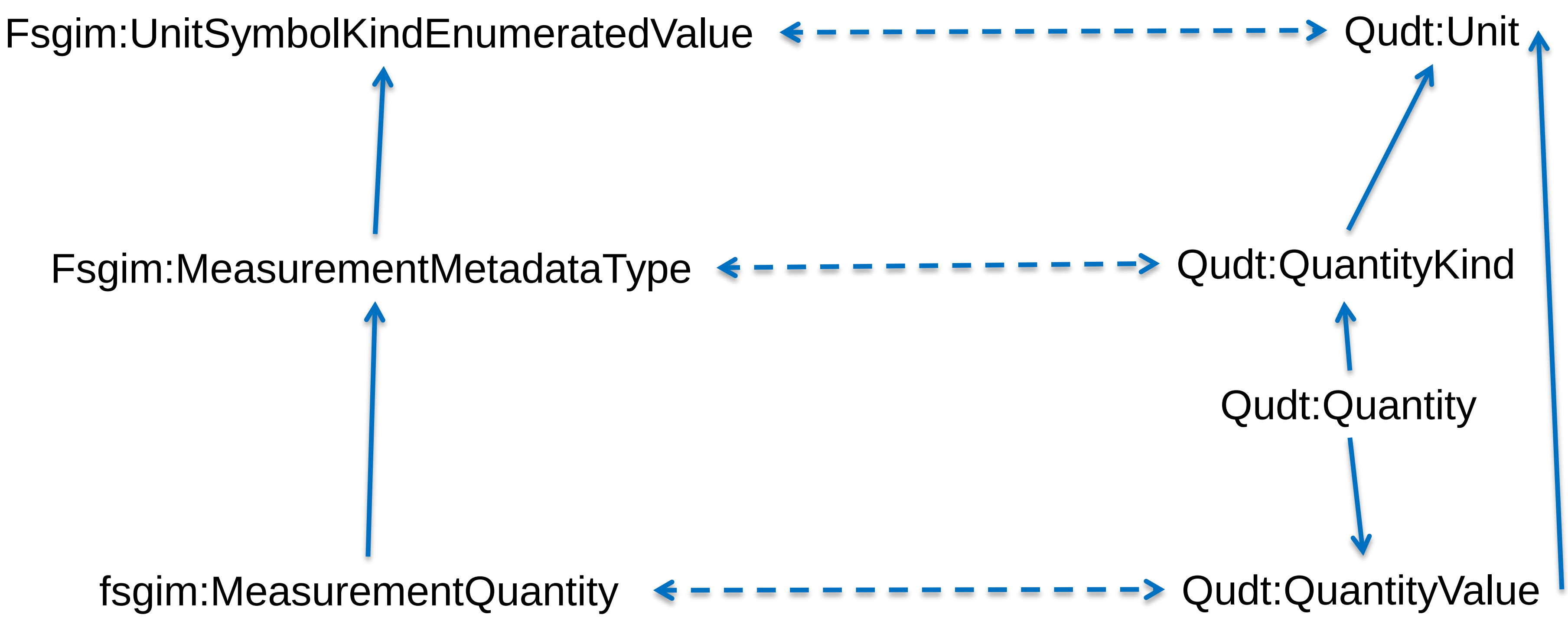
Implementation of FSGIM/BRICK Equivalencies



Notes:

- Every fsgim:Device has an equivalent Brick:Equipment, but not vice versa. The scope of fsgim is electricity management, not, for example, temperature control.

Measurement Model Equivalence



EPIC/FSGIM/QUDT/SSN Measurement Model Equivalence (Real power instances)

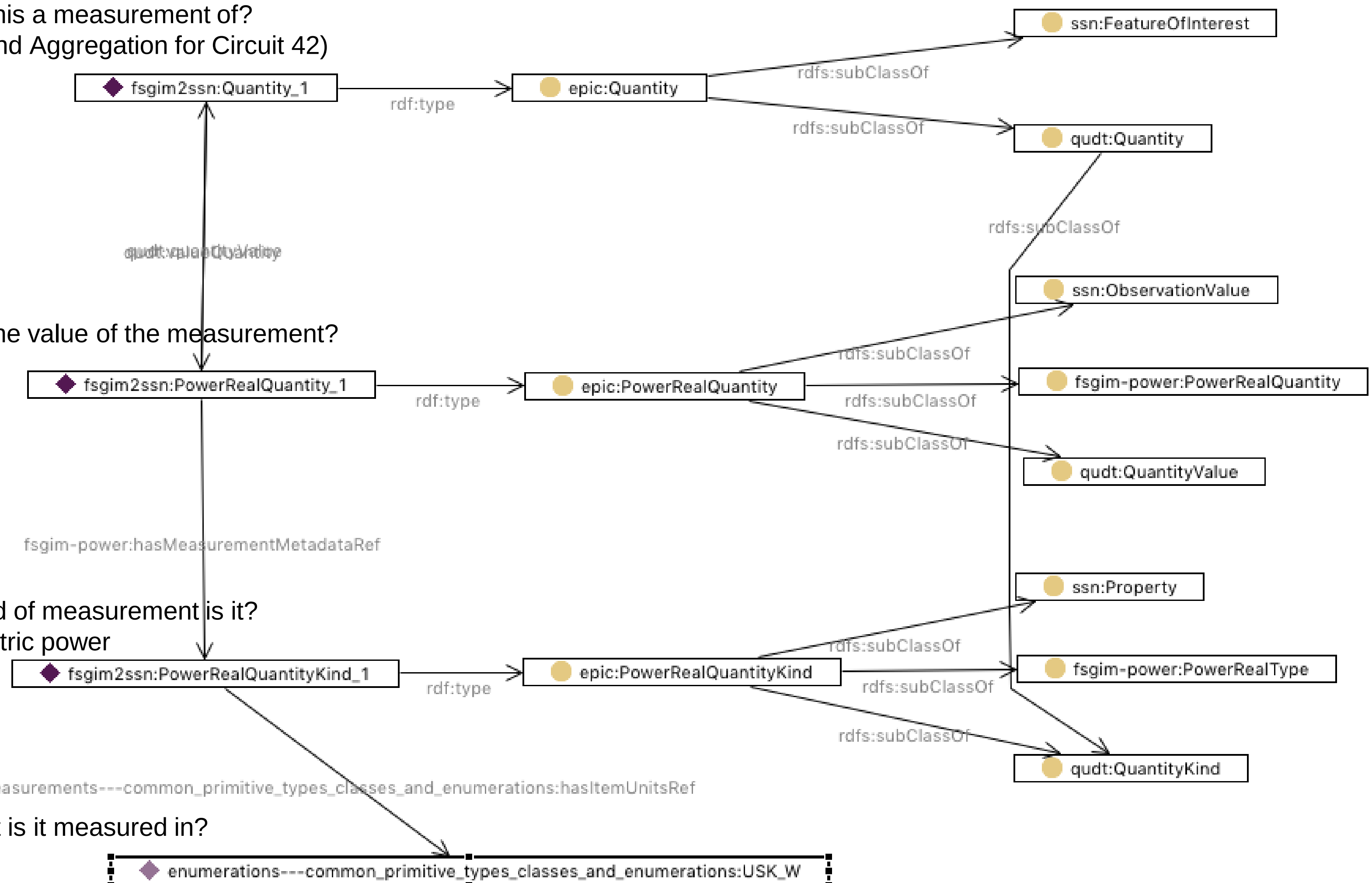
(QUDT Rev. 1.2)

What is this a measurement of?
(A Demand Aggregation for Circuit 42)

What is the value of the measurement?
4.3

What kind of measurement is it?
Real electric power

What unit is it measured in?
Watts



Where do we go from here?

- Systems engineering works well when you can specify all the components
  - But not so well otherwise

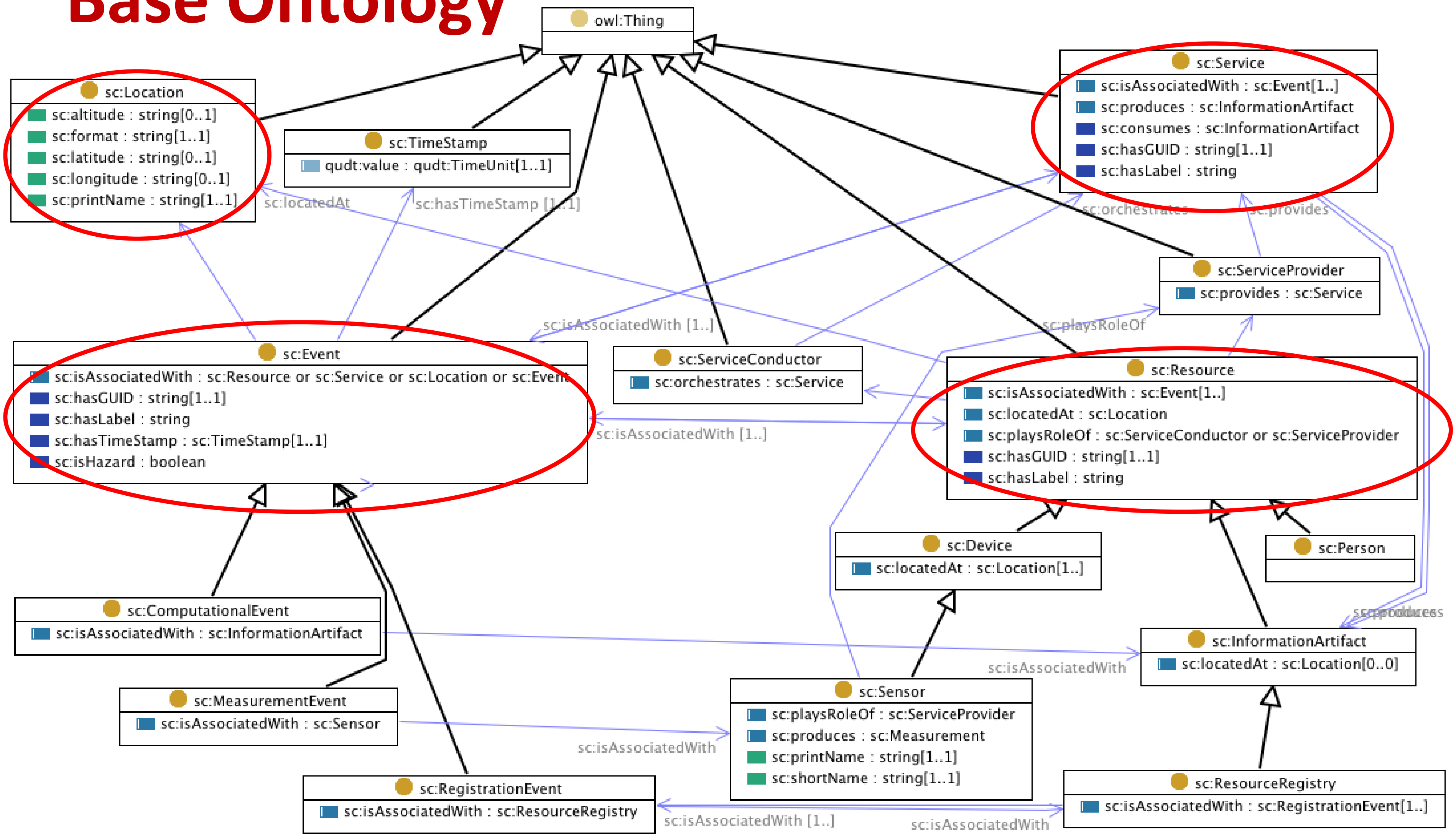
"Meta-systems engineering"

- Components that talk about themselves
- Who, what, why, where, when, how


models
standards
information
data

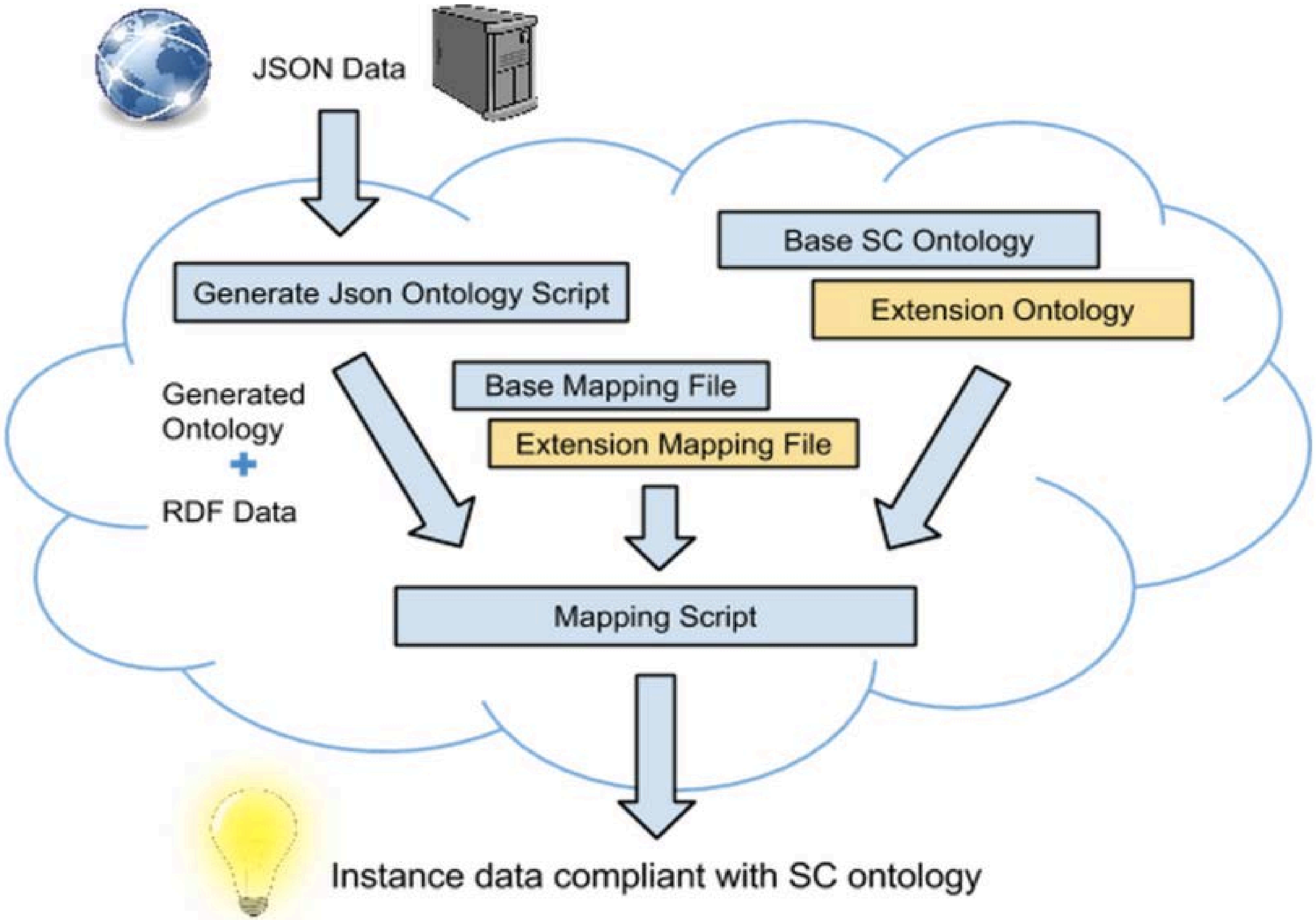
Backup Slides

Base Ontology

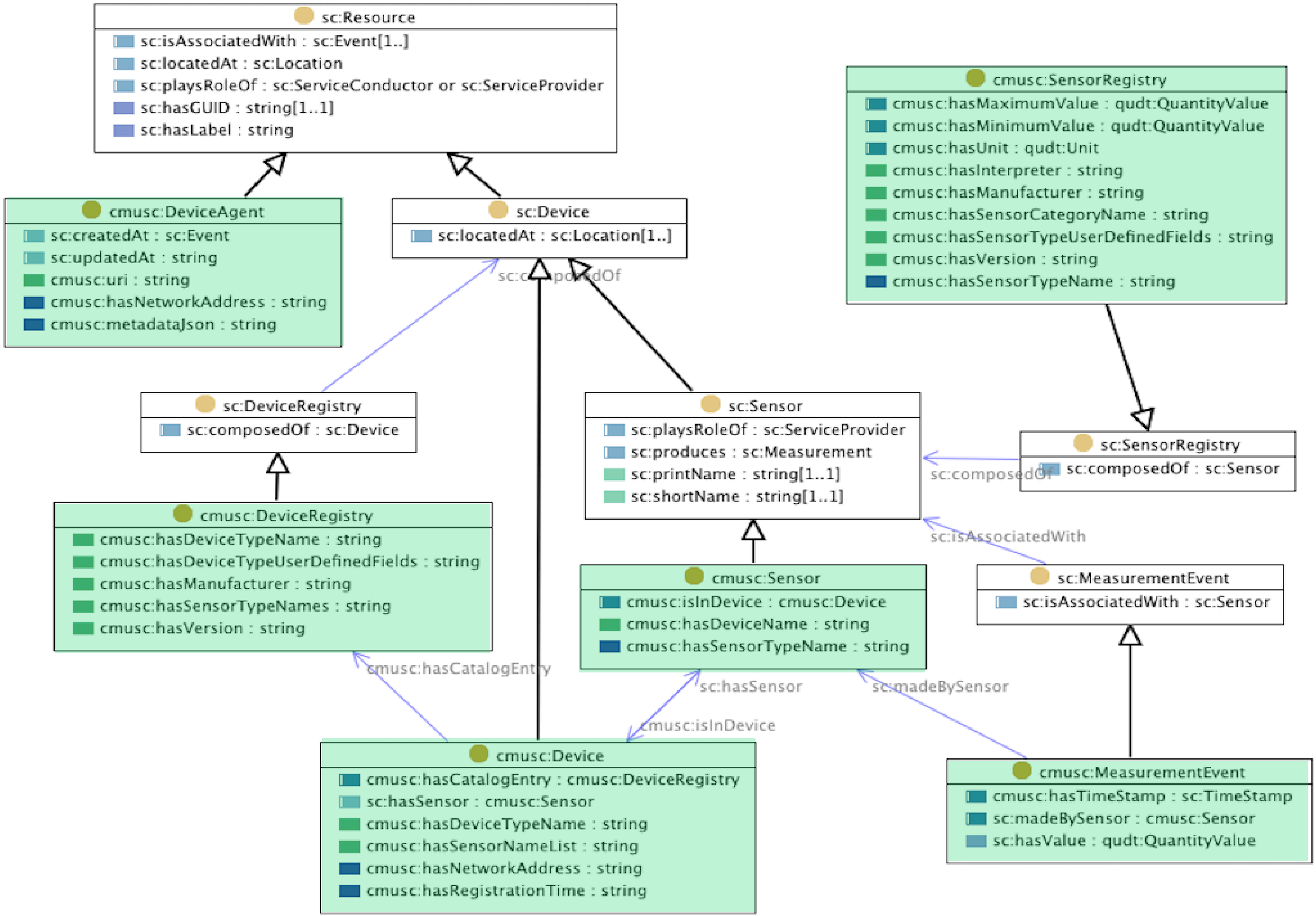


Version1 of Base Ontology available at https://sites.google.com/a/steveray.com/ontologies/home/repository/CoreSCModel.ttl?attredirects=0&d=1

High-Level System Design



Some Extension Classes for a Sensor Source



Why Consider RDF & OWL Semantic Web Technology?

RDF = Resource Description Framework
OWL = Web Ontology Language

1. Simple representation
   – Everything is a triple: <subject – predicate – object>
2. Self-describing models
   – Schemas and data coexist in data stores
3. Easy to interrogate
   – SPARQL queries (over schema and data)
4. Easy to validate
   – Supports automated reasoning
5. Easy to interoperate
   – Natively supports distributed data stores

Simple Representation

Everything is stored as triples:

<subject predicate object>

Self-Describing Models

- The schema (model) and the data is stored in the same place

- Schema:
  - Mammal subClassOf Animal
  - Human subClassOf Mammal

- Data:
  - george is-a Human
  - george marriedTo lisa

Easy to Interrogate

SPARQL[†] language to query an RDF database
(Just matches against patterns of triples)

```
SELECT ?x
WHERE {
   george  marriedTo  ?x .
}
```

Returns a table:

x
lisa

```
SELECT ?y
WHERE {
   y? subClassOf  Animal .
}
```

Returns a table:

y
Mammal

† SPARQL = SPARQL Protocol and RDF Query Language

Easy to Validate

SPARQL can be used

for reasoning,

not just interrogating

If

George	sonOf	Fred

and

Fred	siblingOf	Mary

Then

George	nephewOf	Mary

In SPARQL:

```
CONSTRUCT
{ ?a nephewOf ?c .}

WHERE
{
   ?a sonOf ?b .
   ?b siblingOf ?c .
}
```

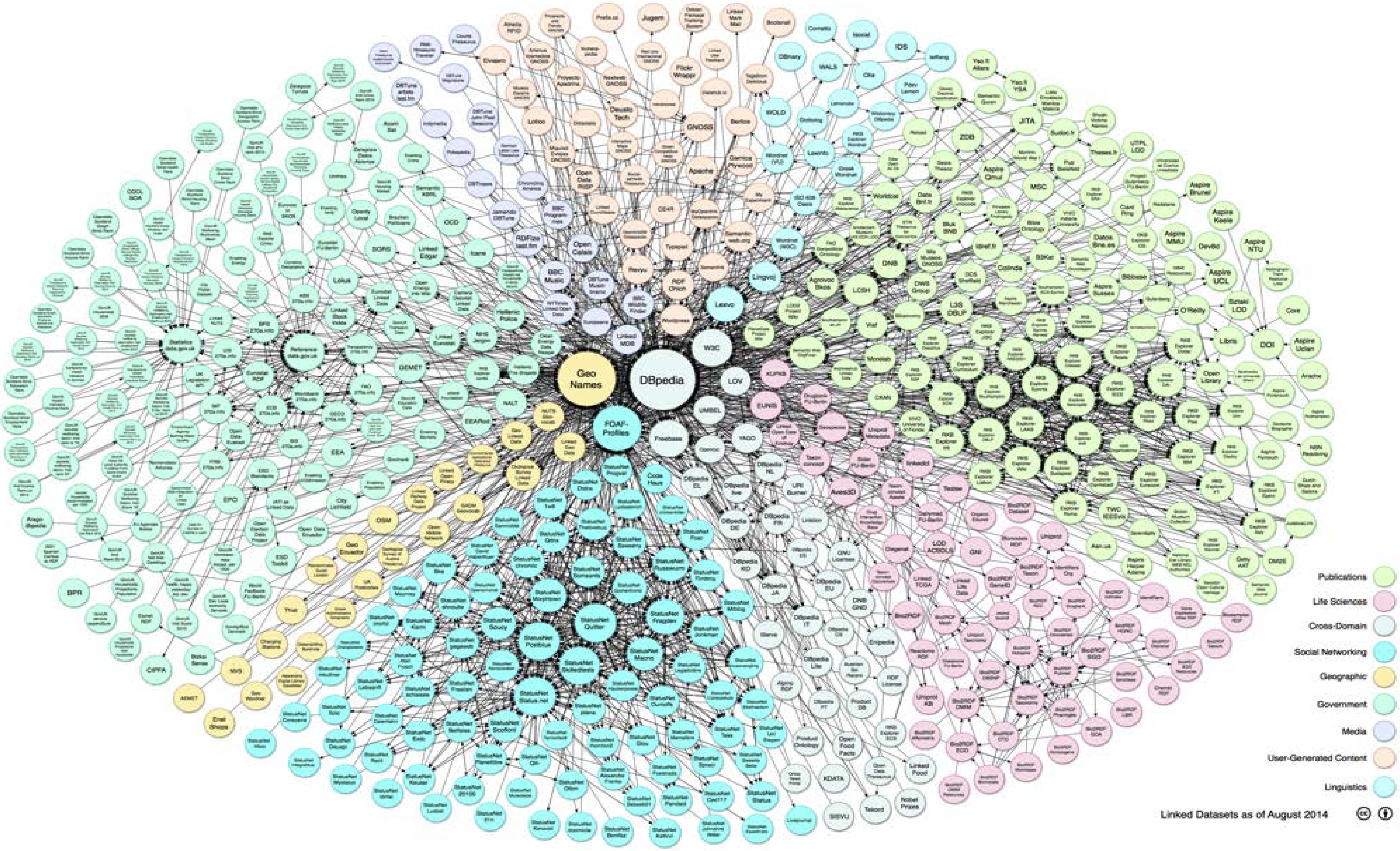
Easy to Interoperate

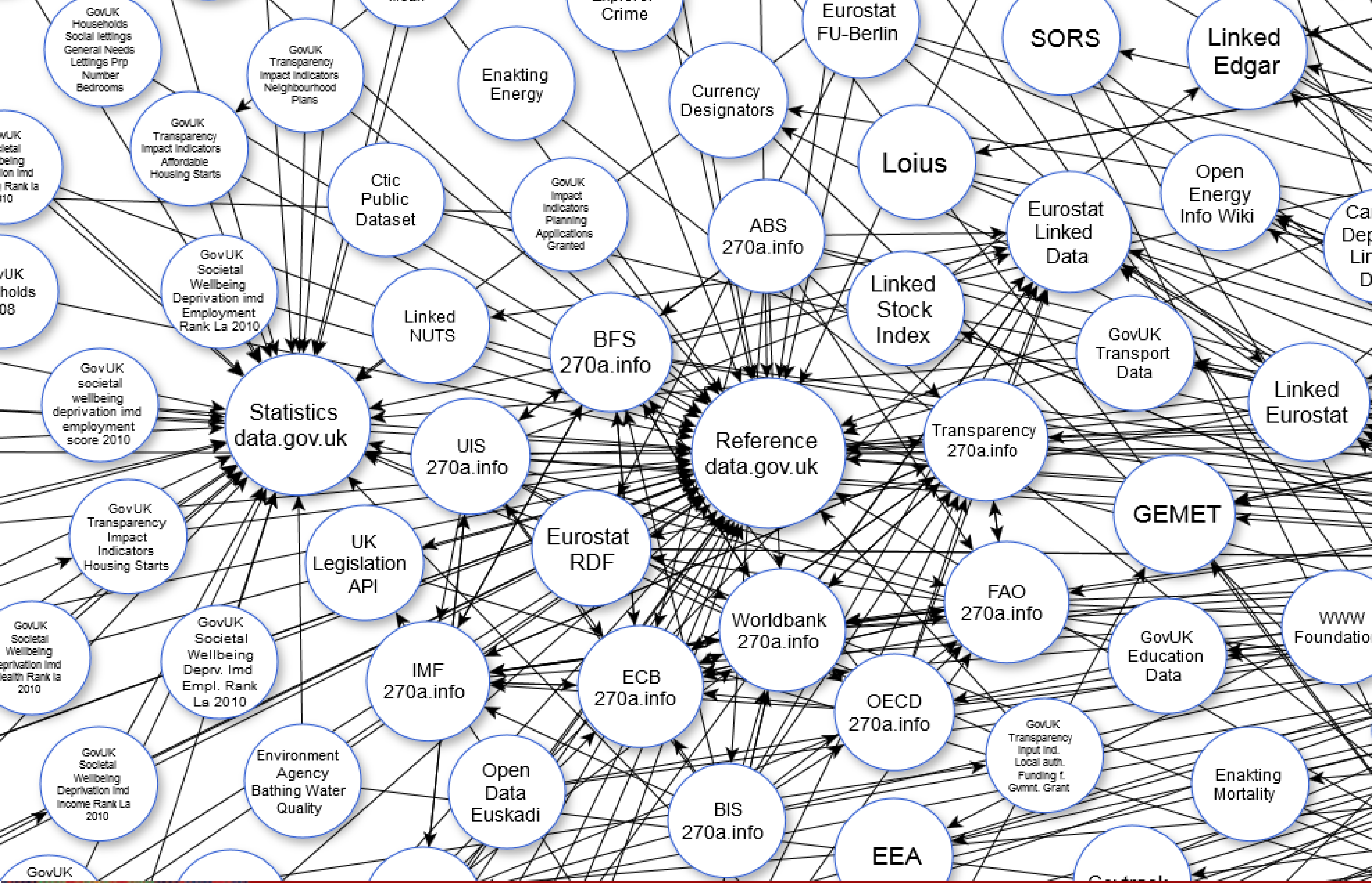
- A single query can interact with <u>more than one</u> RDF database
  - Linked Movie Database contains movies, actors
  - DBPedia contains people and birthdates
- Find the birthdates of all principal Star Trek movie actors
  - Answer does not exist in one source

```
PREFIX movie: <http://data.linkedmdb.org/resource/movie/>
SELECT DISTINCT ?actor_name_en ?birth_date
WHERE {
SERVICE <http://data.linkedmdb.org/sparql> {
?movie rdfs:label ?title .
FILTER (REGEX (?title, "Star Trek")) .
?movie movie:actor ?actor .
?actor movie:actor_name ?actor_name .
}
BIND (smf:setLanguage(?actor_name , "en") AS ?actor_name_en ) .
SERVICE <http://dbpedia.org/sparql> {
?actor2 foaf:name ?actor_name_en .
?actor2 a <http://umbel.org/umbel/rc/Actor> .
OPTIONAL {
?actor2 <http://dbpedia.org/property/birthDate> ?birth_date .
}
}
}
```

[actor_name_en]	birth_date
Alice Krige	1954-06-28
Armin Shimerman	1949-11-05
Brent Spiner	1949-02-02
Bryan Singer	1965-09-17
Daniel Hugh Kelly	1952-08-10
DeForest Kelley	
Dwight Schultz	1947-11-24
F. Murray Abraham	1939-10-24
Gates McFadden	1949-03-02
George Takei	1937-04-20
James Cromwell	1940-01-27
James Doohan	
James Sikking	1934-03-05
Jonathan Frakes	1952-08-19
Jude Ciccolella	1947-11-30
Kim Cattrall	1956-08-21
Kirstie Alley	1951-01-12
Kurtwood Smith	1943-07-03
LeVar Burton	1957-02-16
Leonard Nimoy	1931-03-26
Marina Sirtis	1955-03-29
Neal McDonough	1966-02-13
Nichelle Nichols	1932-12-28
Patrick Stewart	1940-07-13
Patti Yasutake	
Robert Picardo	1953-10-27
Ron Perlman	1950-04-13
Stephanie Niznik	1967-05-20
Whoopi Goldberg	1955-11-13
Wil Wheaton	1972-07-29
William Shatner	1931-03-22

Linked Open Data Cloud





Carnegie Mellon University
Silicon Valley

- Using semantic representations
  - expose inconsistencies in new standards
  - codify natural language rules
  - use automated reasoning for instantiation and explanation

- Exhaustively searches a standard to find errors that might escape human detection
  - Orphan definitions (defined but never used)
  - Opportunities for model refactoring (similar classes)
  - Disallowed changes to imported standards
  - Redundant classes and properties
  - Non-standard data type definitions

After Verification Testing

- Model Healing
  - Recommendations to correct errors
  - Automatic error correction for native OWL specifications
- Conformance Testing
  - Does a particular implementation properly represent the information according to the standard?
  - Generation of reference data sets
- Standards Harmonization
  - Checking for missing information
    - Information present in one standard but not in another
  - Mapping among different ways of modeling the same information