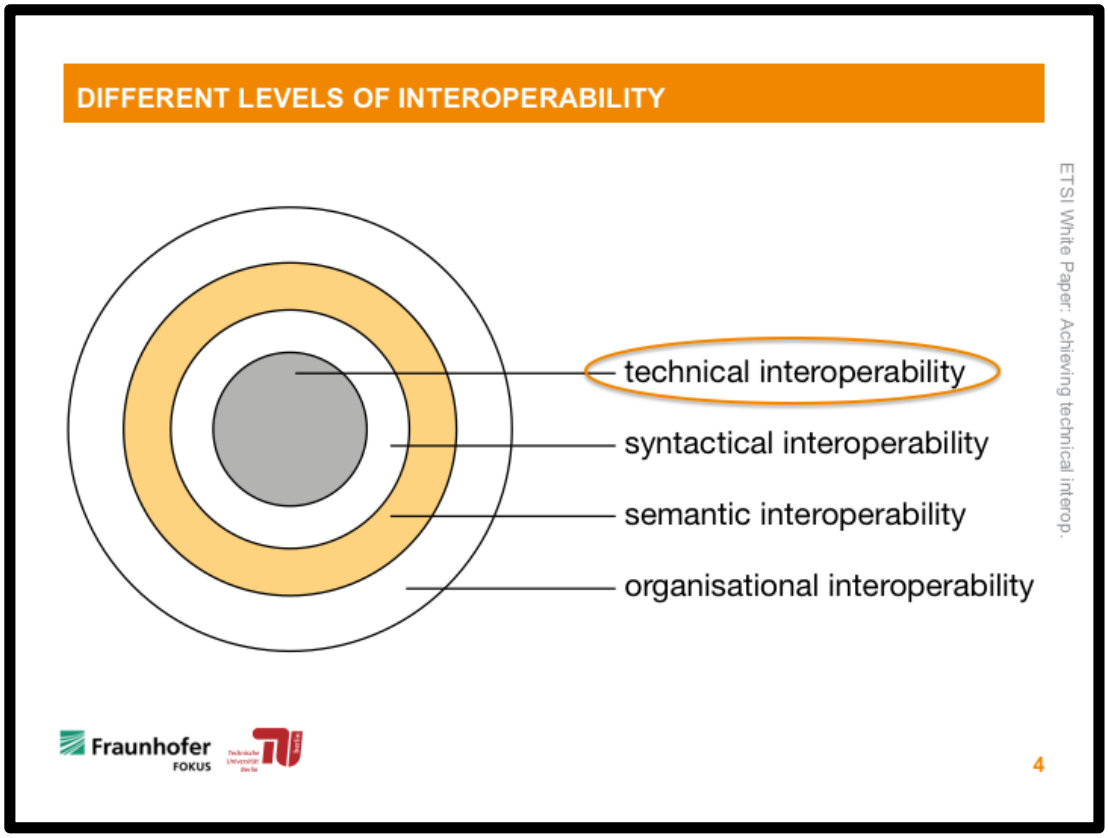

INDUSTRIAL INTERNET OF THINGS (IIOT)
PART 5: SEMANTIC INTEROPERABILITY / DIGITAL TWINS
AV Lecture in Summer Term 2018
Dr.-Ing. Alexander **Willner**
Fraunhofer
FOKUS

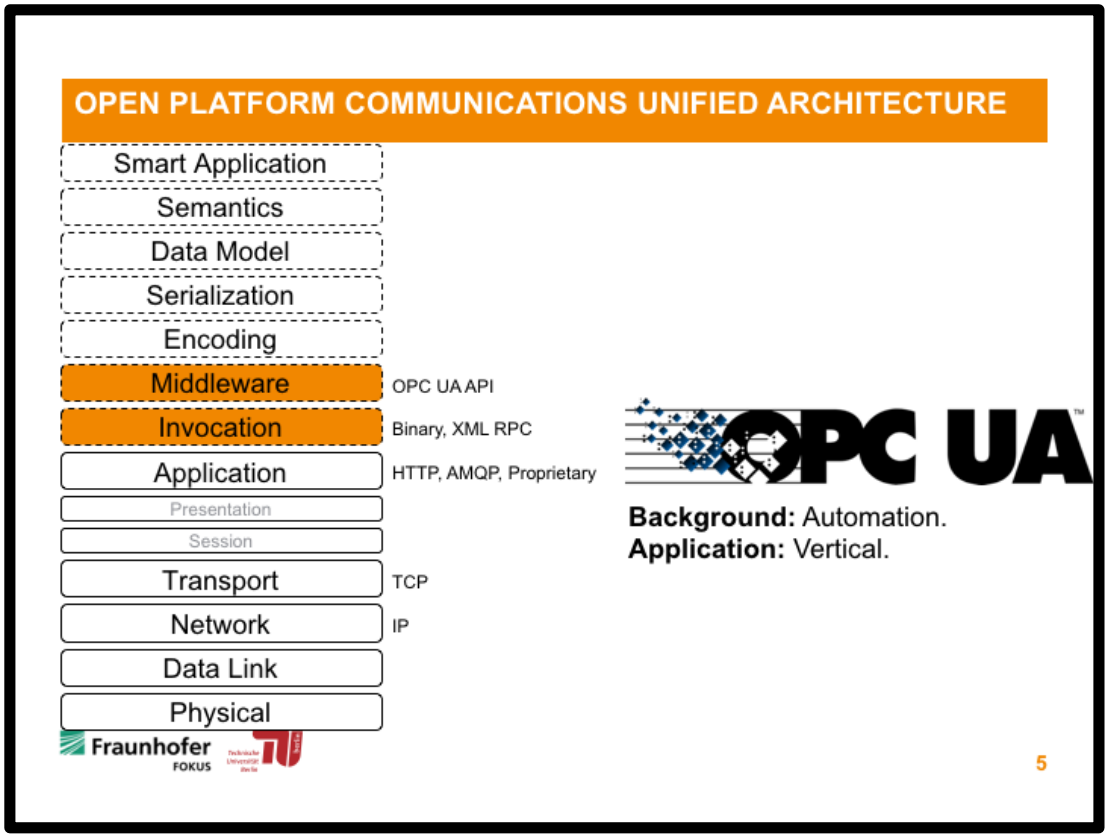
THE LAST LECTURE

5 minutes

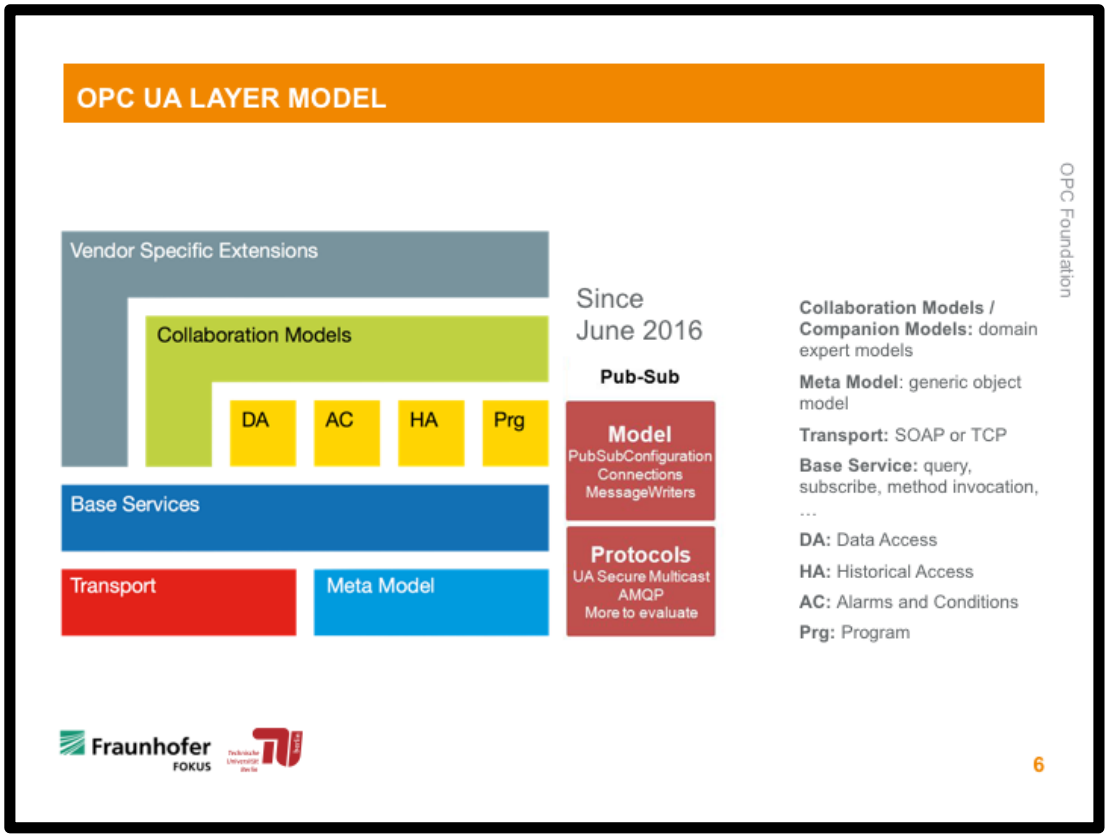
- Last lecture we talked about aspects 2 and 3



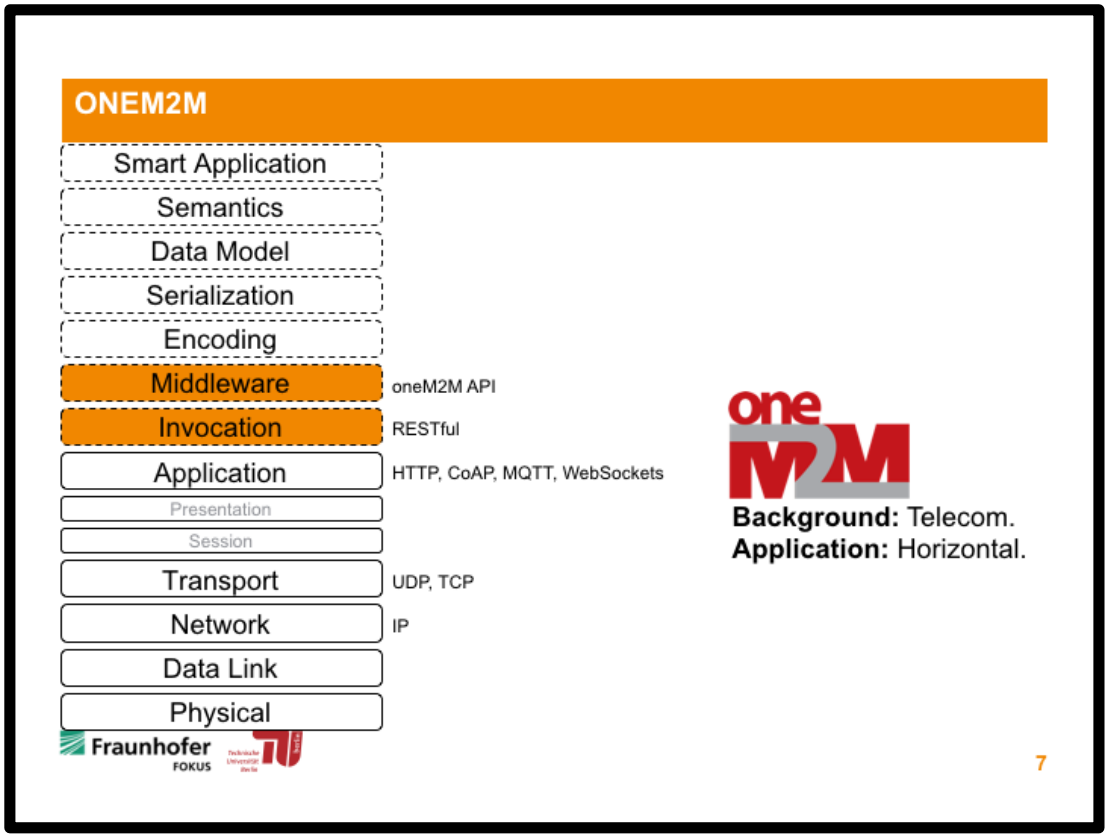
- We finished the part about technical interoperability



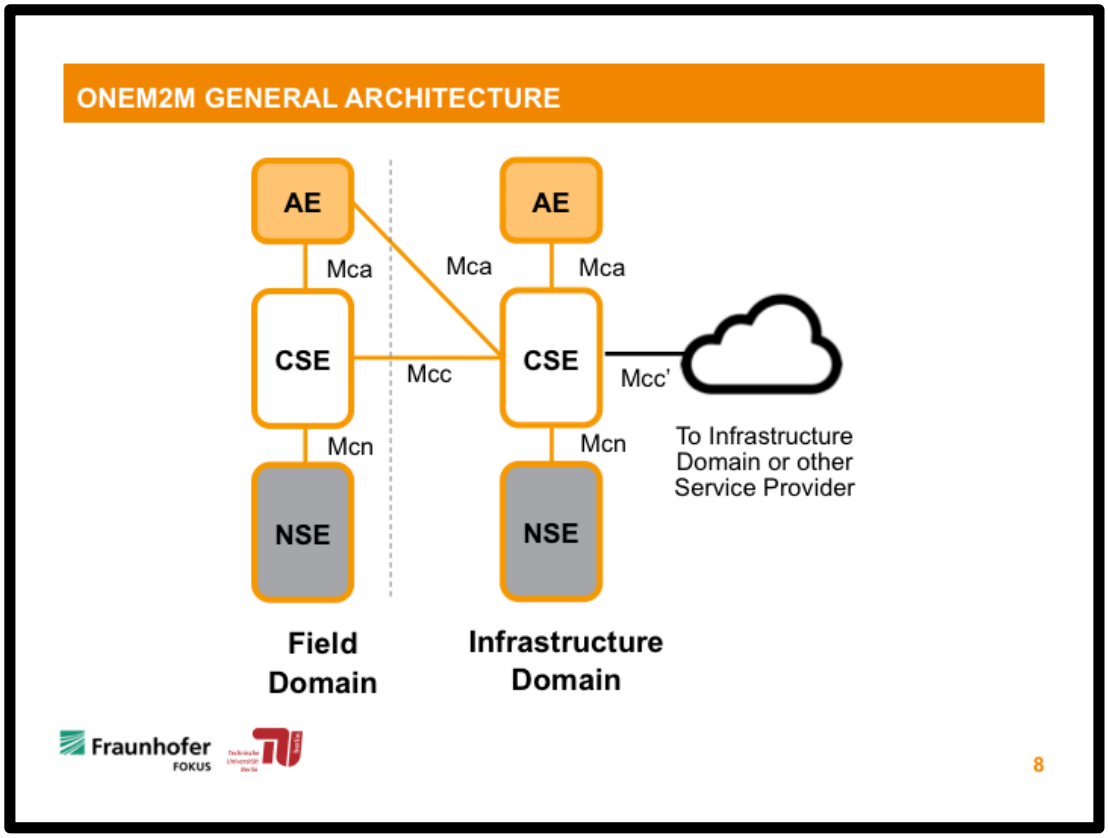
We talked about the most important middleware in the manufacturing context: OPC UA



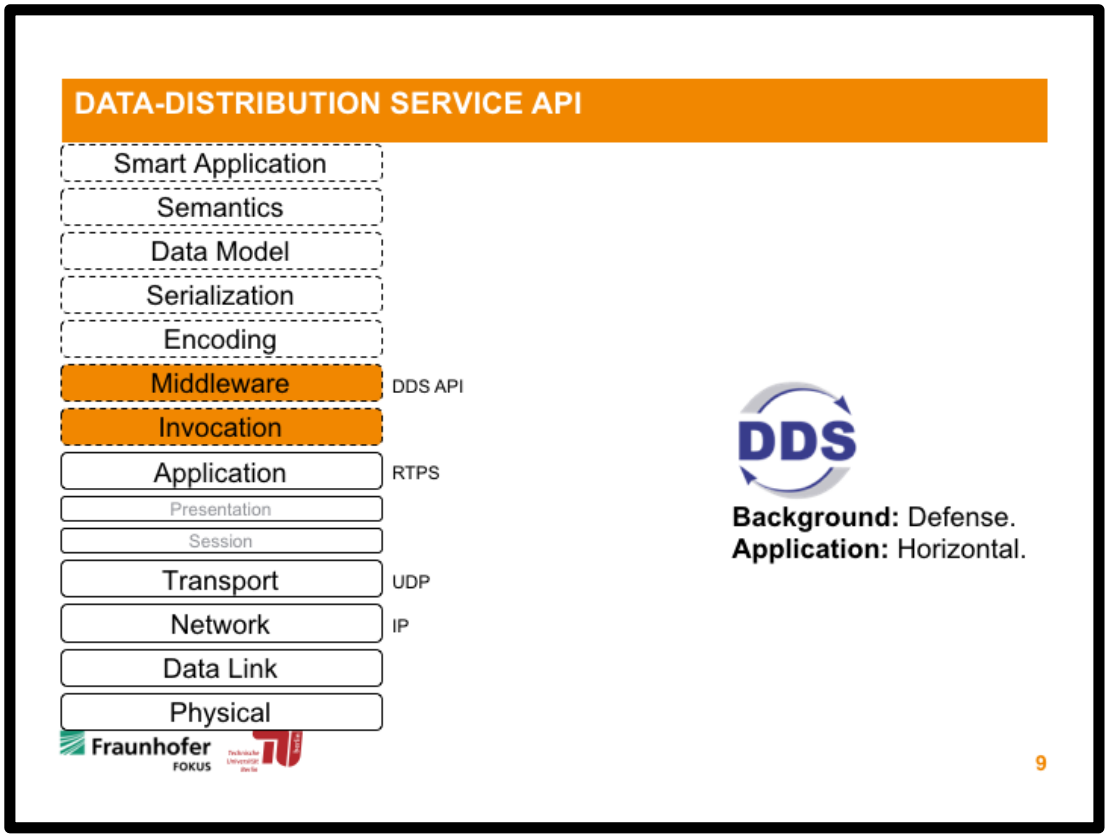
The overall architecture of OPC UA



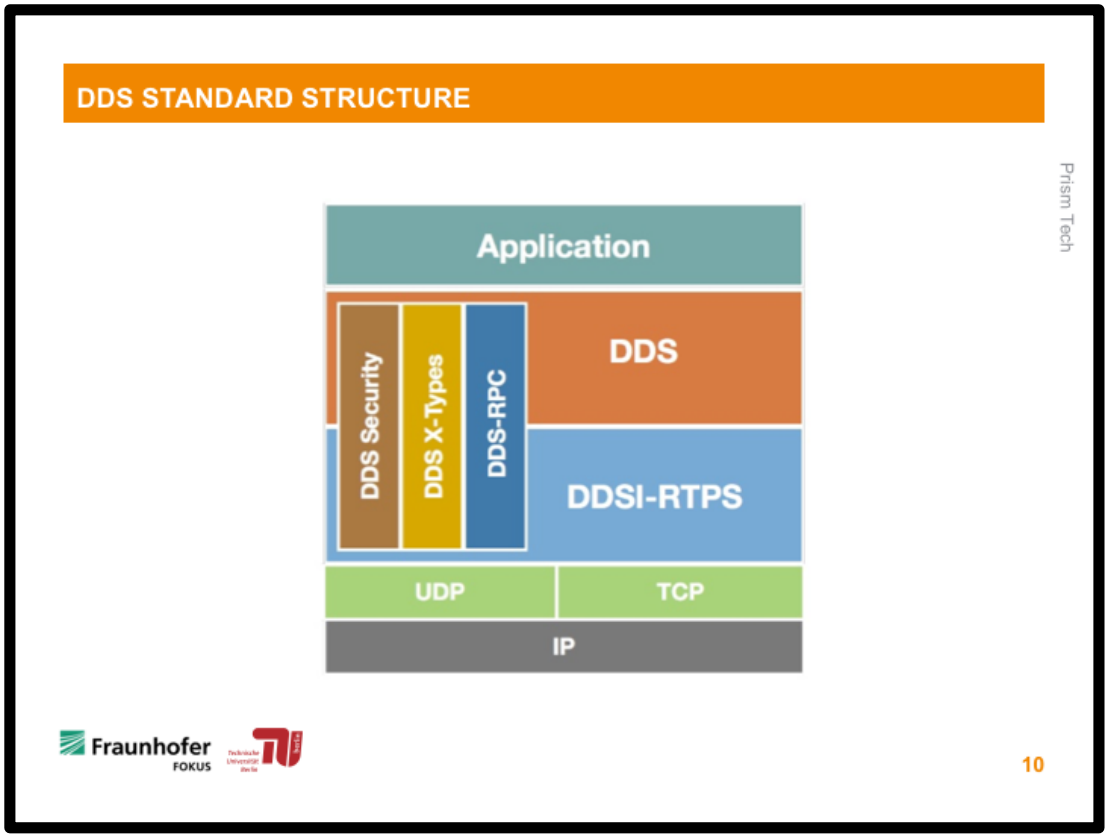
- We talked about an important middleware for horizontal integration



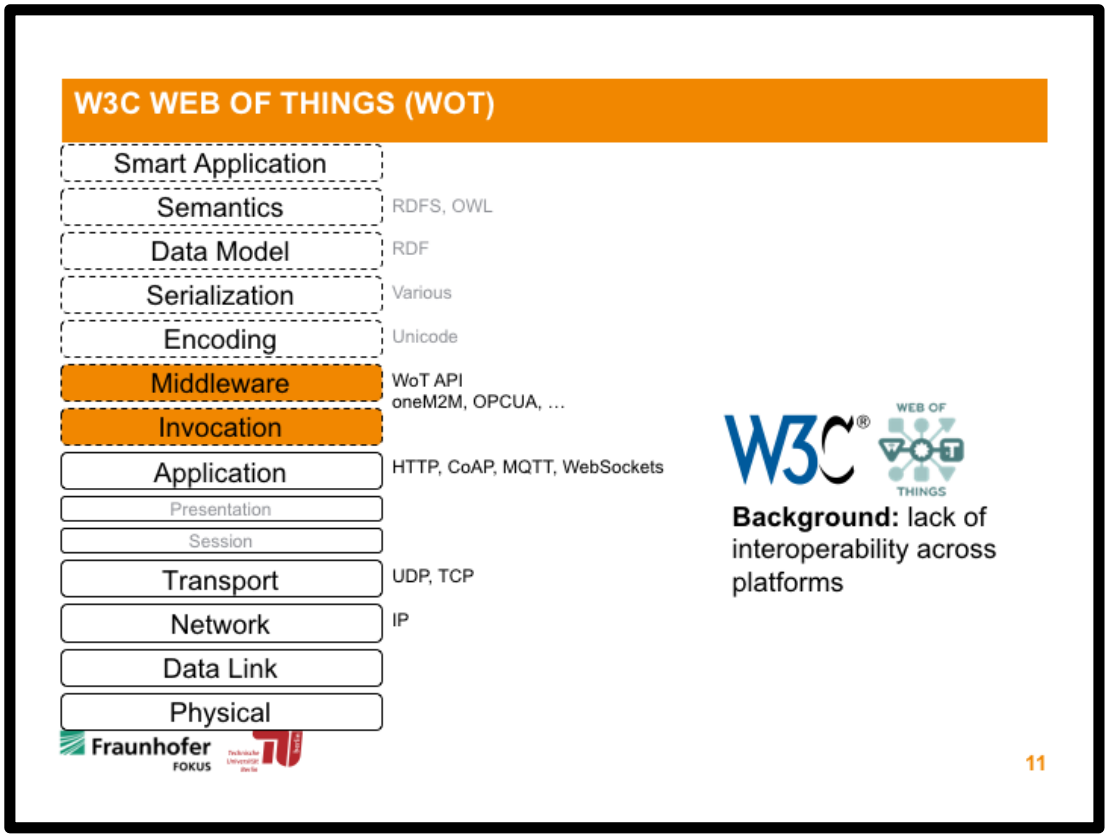
- The overall architecture of oneM2M



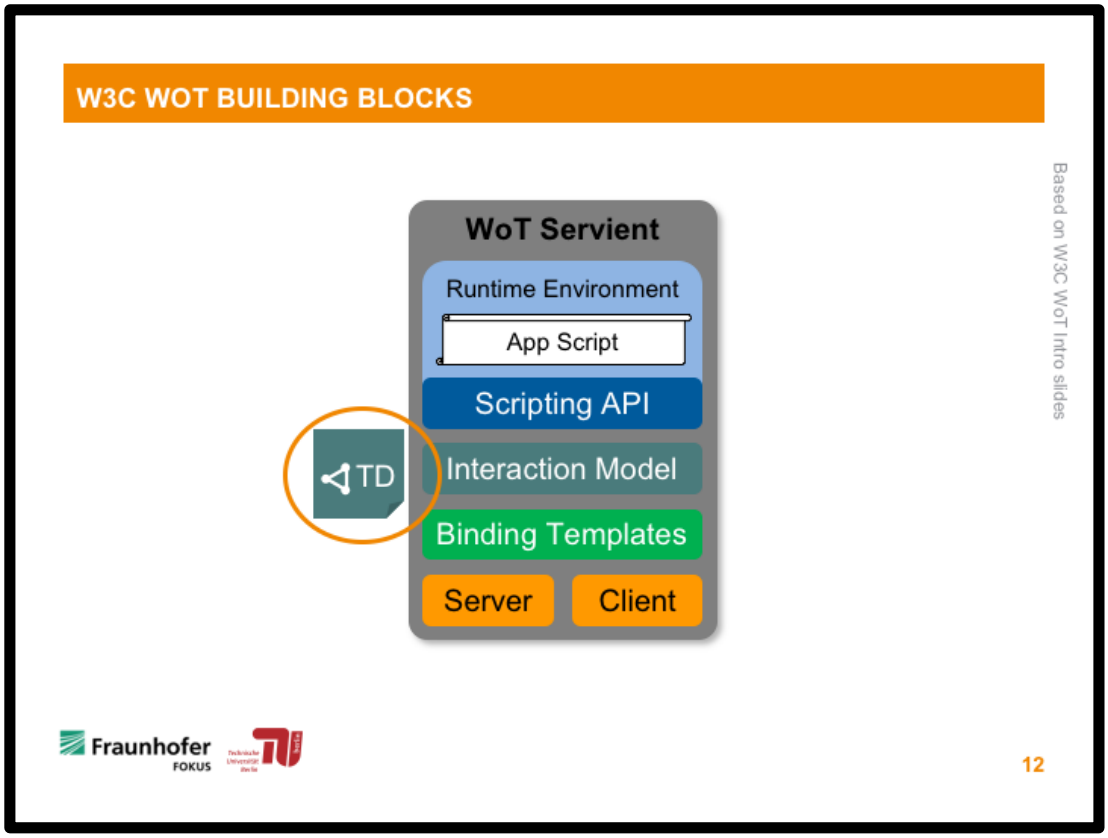
- We talked about DDS, which has been standardized by OMG, which has a close relation to the IIC



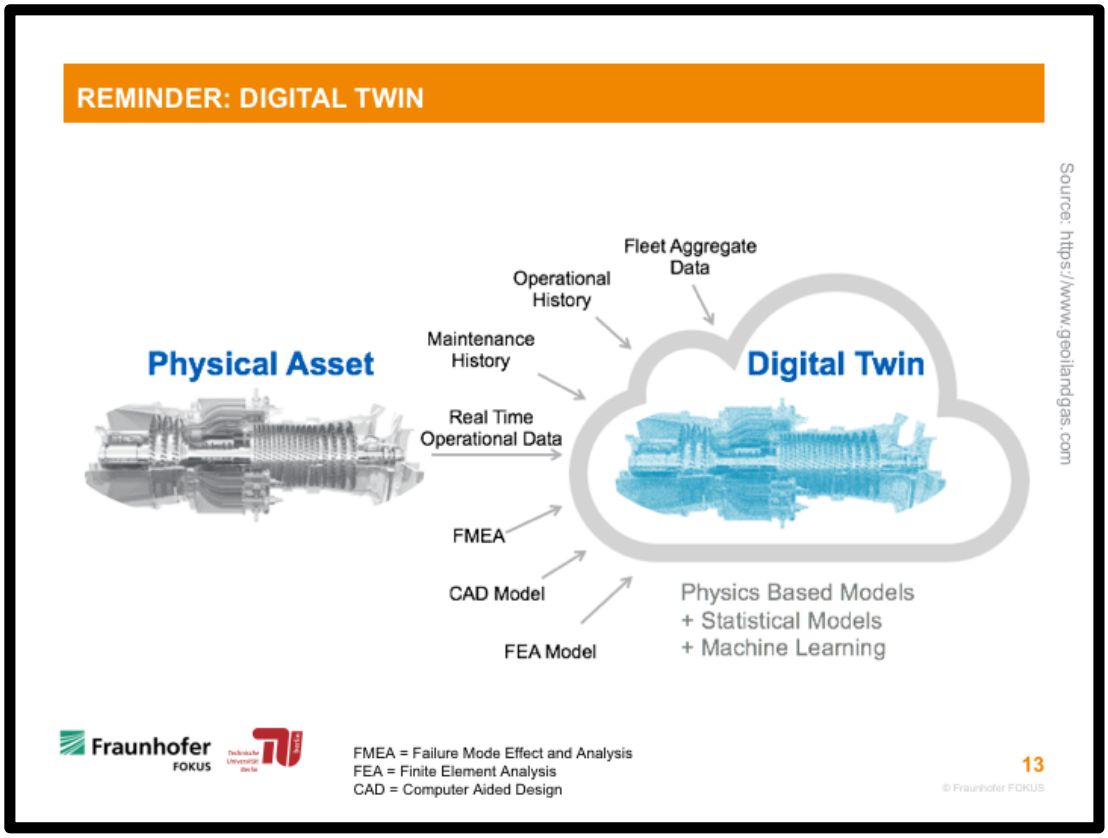
- The overall architecture of a DDS node



- We talked about the W3C WoT approach to enable interoperability between the middleware systems



- The architecture of a W3C WoT node
- Things Description (TD) as the core interoperability aspect



- Semantically, distributed description of digital twin might be the core of future interoperability
- Extends the concept for initial simulation towards a fully interconnected graph of knowledge

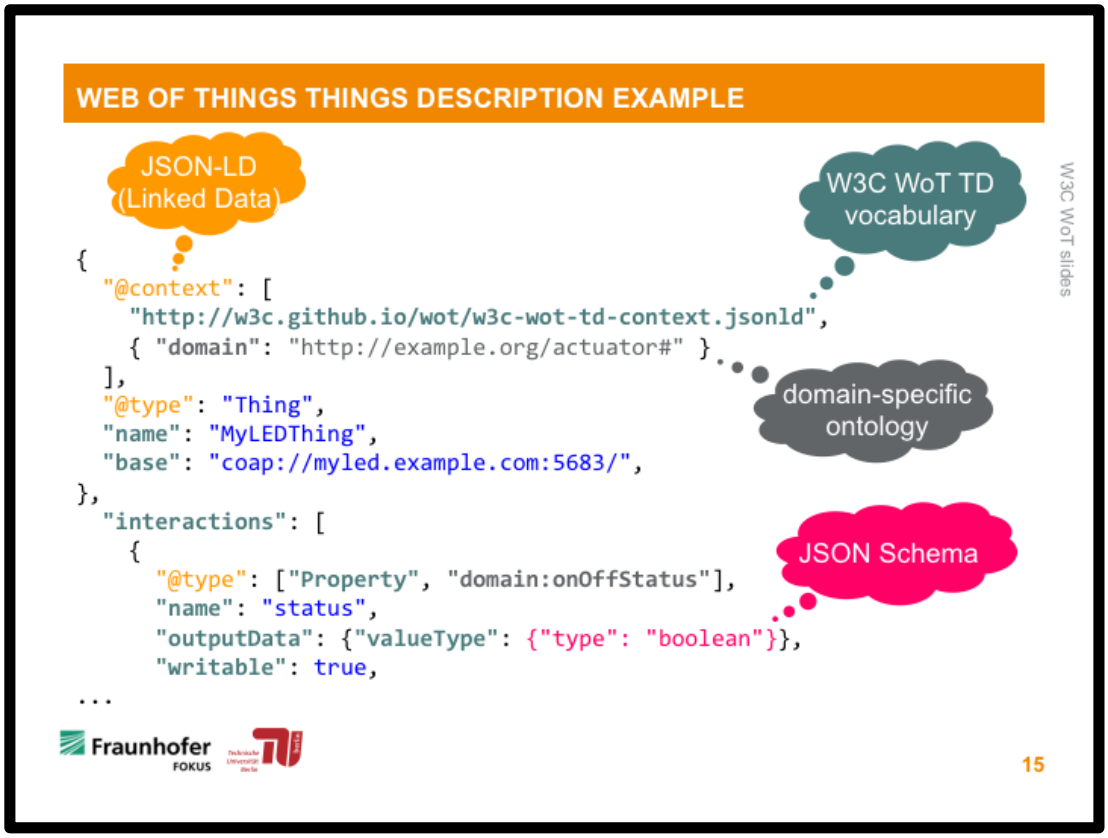
NEW LANGUAGE FOR MACHINES (1 MIN)

https://youtu.be/OvfU1NpCJQQ

Fraunhofer FOKUS

14

- Working on a language for machines in the Industrial IoT



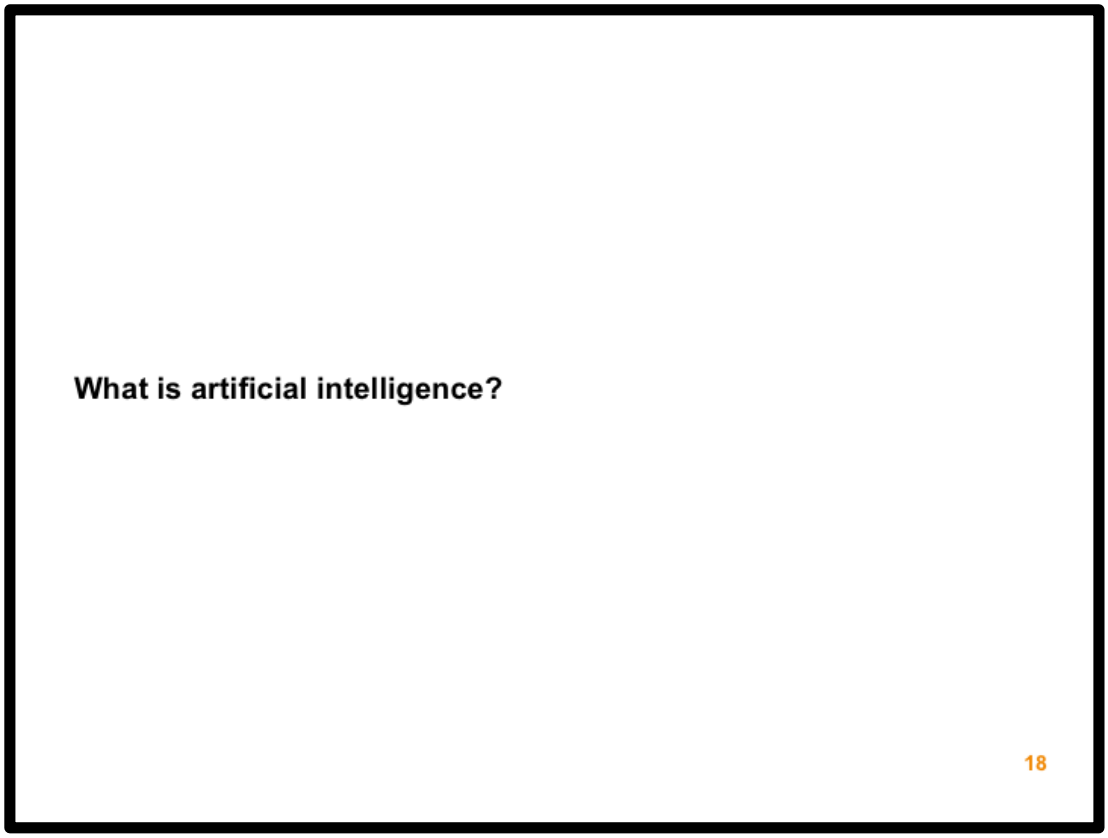
- A number of acronyms/concepts/technologies that we want to understand today

THE LAST LECTURE

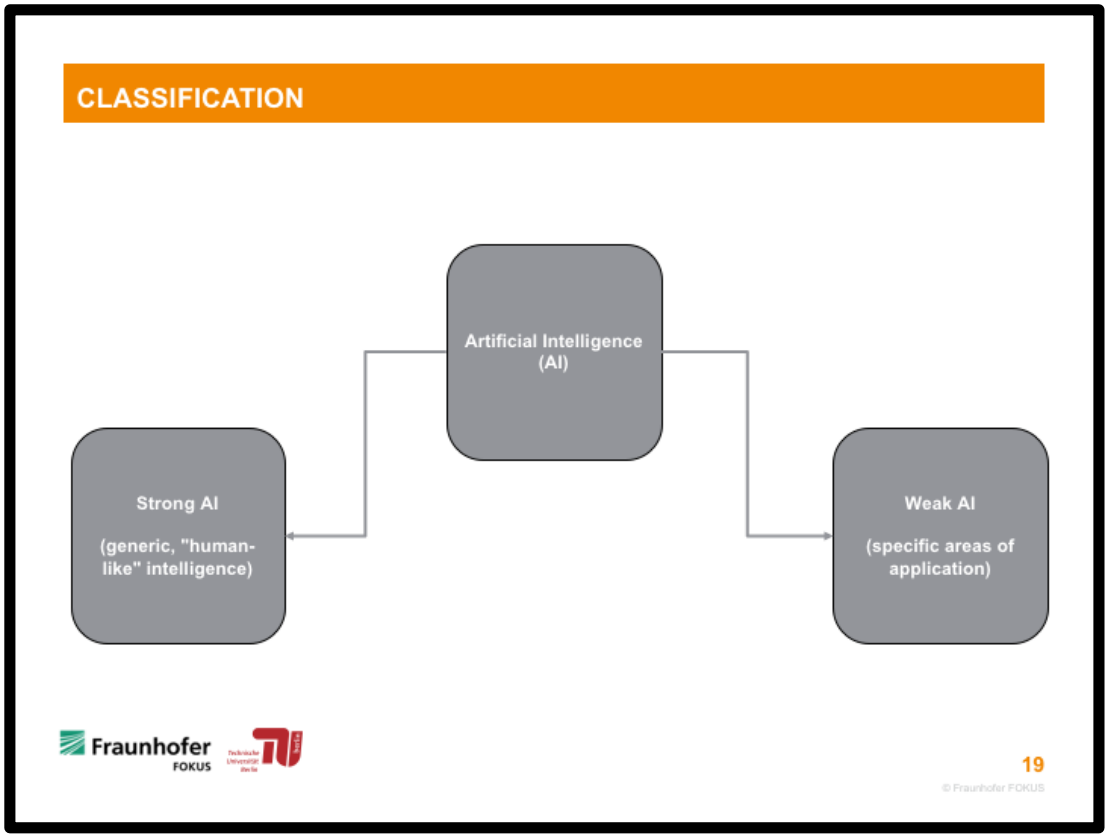
Questions?

MOTIVATION: SYMBOLIC ARTIFICIAL INTELLIGENCE

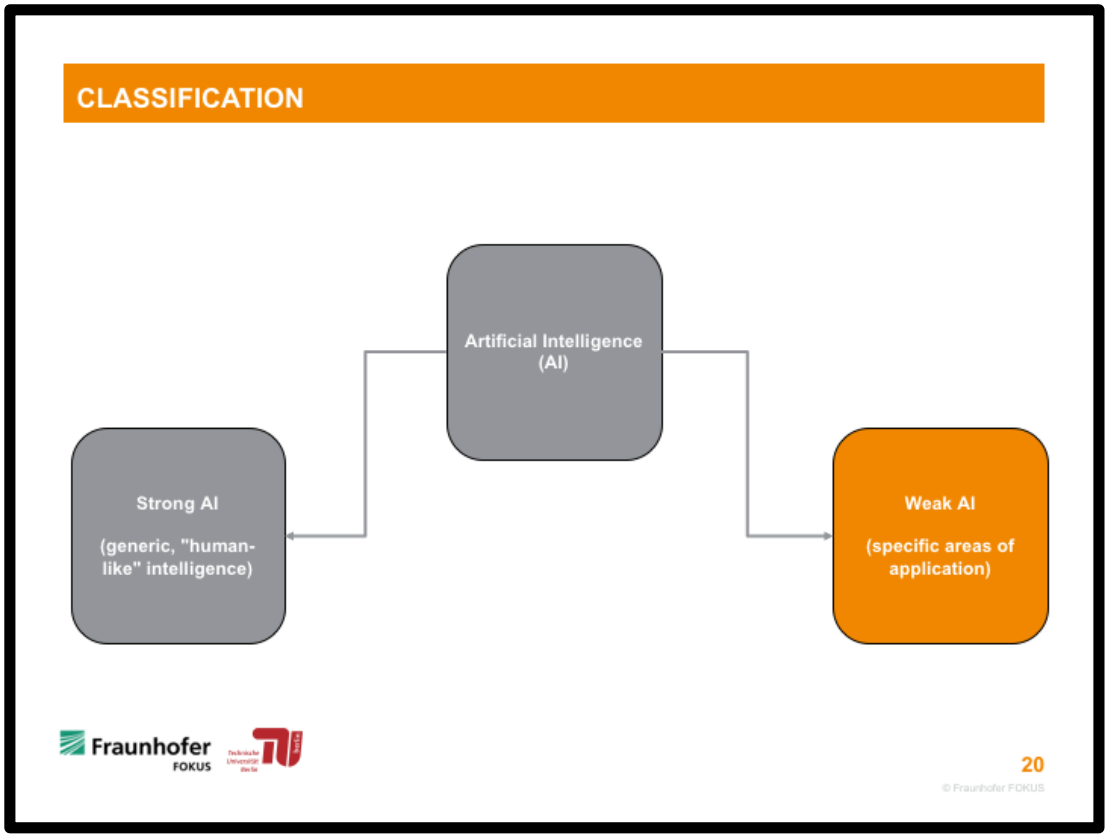
5 minutes



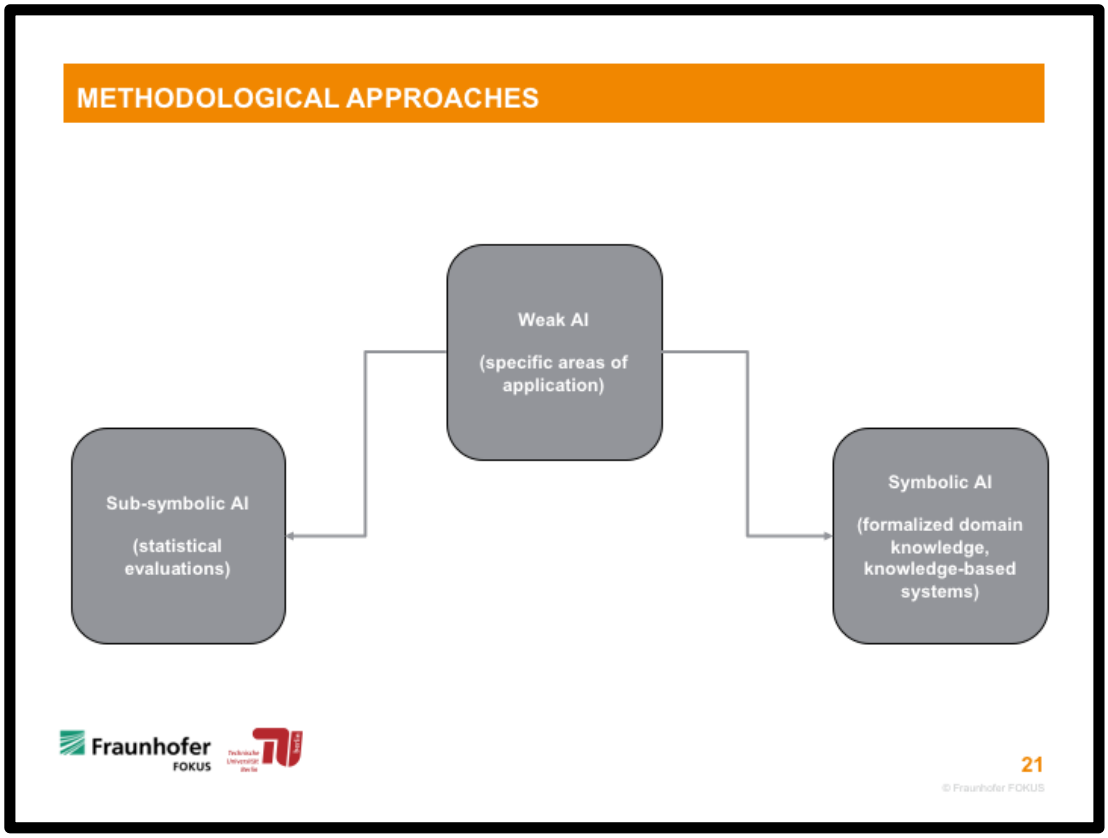
- Subfield of computer science/mathematics since the mid/end of the 1950s (see Dartmouth Conference)



- Basically distinguish between strong and weak AI

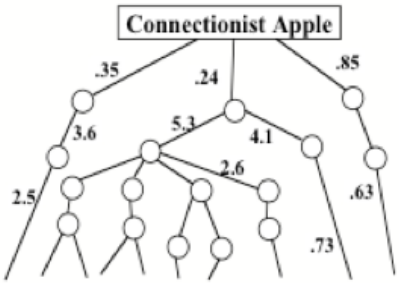


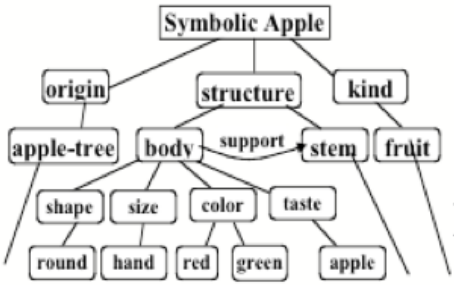
- Even if intelligent digital assistants try to simulate a strong AI (natural speech synthesis), the focus is on the combination of different weak, application-related AI technologies.
- Many know it from interactions with the digital assistants on the mobile phones.



- Basically, two methodological approaches are to be distinguished, which also complement each other
- Sub-symbolic AI: new attraction due to artificial neuronal networks
- Symbolic AI: we already started to talk about this

VISUAL COMPARISON



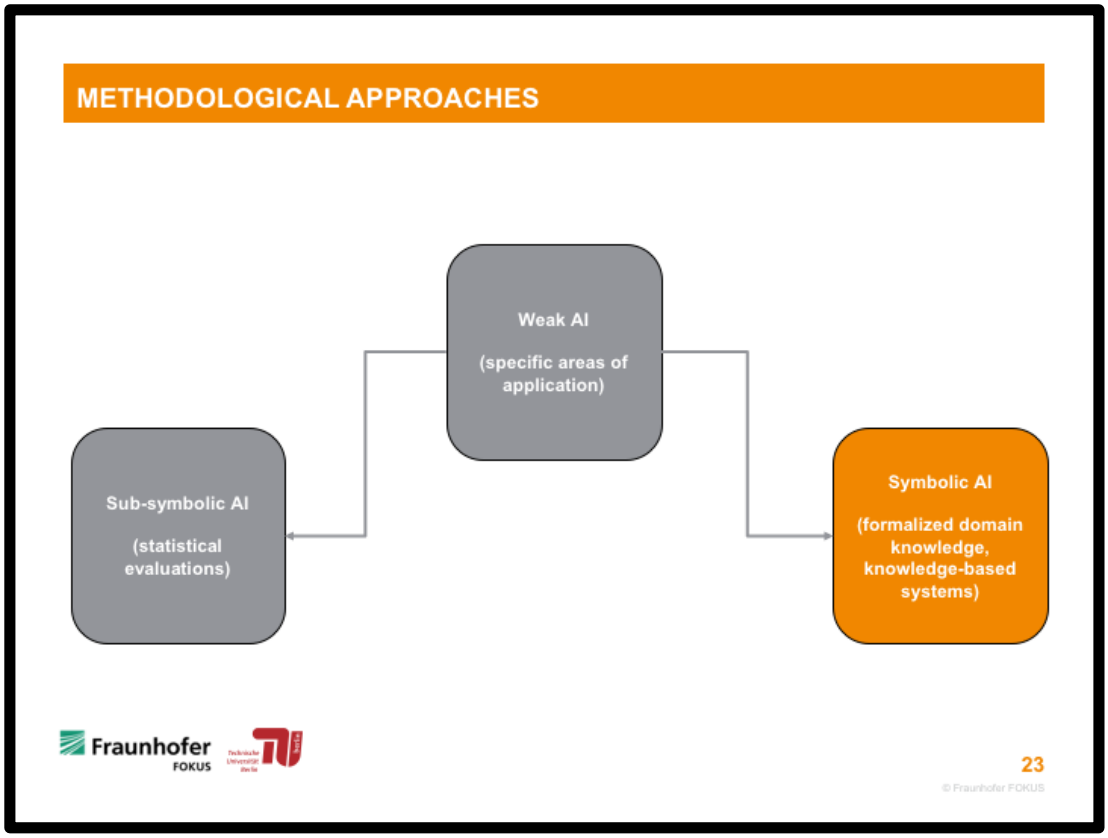


Source: Marvin Minsky: "Logical vs. Analogical or Symbolic vs. Connectionist or Neat vs. Scruffy". In Artificial Intelligence at MIT, Expanding Frontiers, Patrick H. Winston (Ed.), Vol.1, MIT Press, 1990.

Fraunhofer FOKUS

22

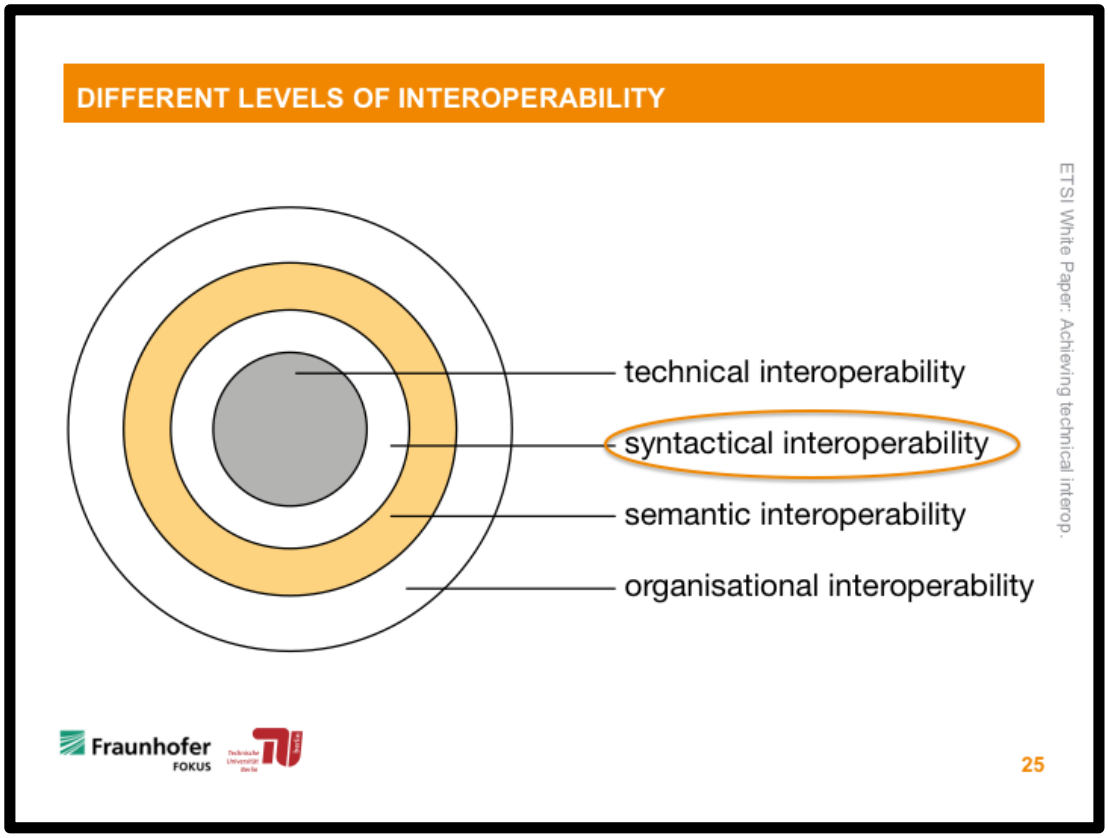
- Left: sub-symbolic AI (statistics) → implicit generation of knowledge through certain weightings / target metrics (e.g. reinforcement learning)
- Right: symbolic AI (logic) → a explicit, formal description of knowledge



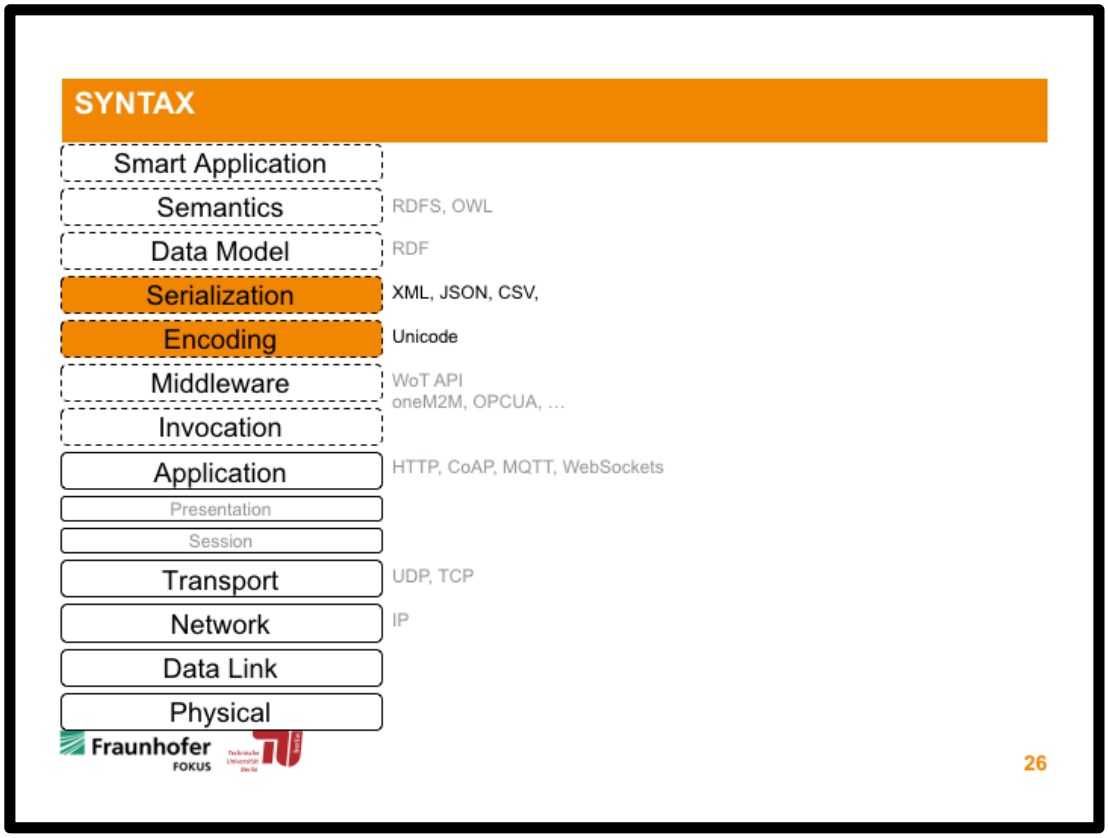
- Let us focus on symbolic KI to allow for interoperability between machines
- It's the formal description of information on which you can infer implizit knowledge based on reasoning

SYNTACTICAL INTEROPERABILITY

8 minutes



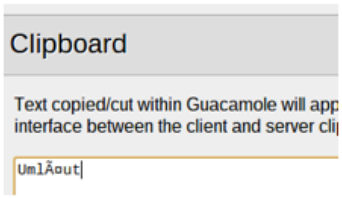
- Now moving to the next layer
- A simple and thin layer



- Two main aspects

UNICODE

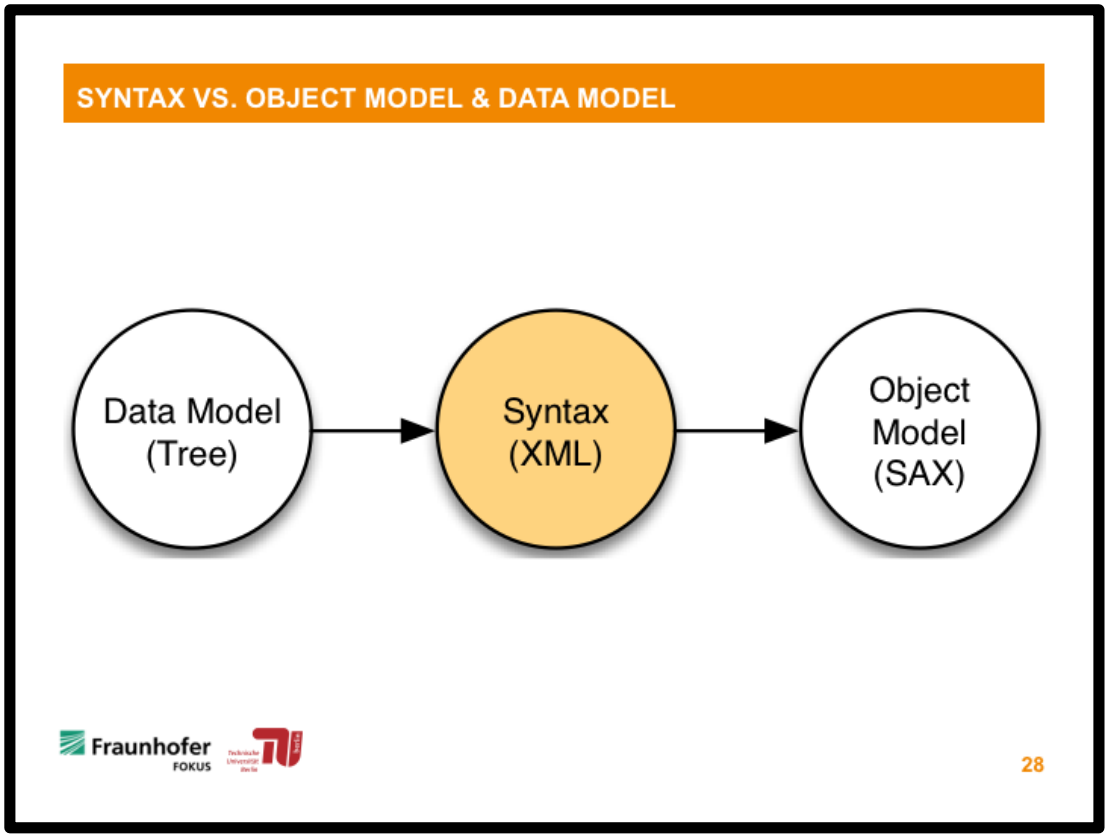
- **Description**:
  - Unicode is a character set used to translate characters into numbers with currently over 120,000 characters and symbols, maintained by the Unicode Consortium.
  - UTF is an encoding used to translate numbers into binary data.
  - UTF-8 is commonly used on the web. First byte = 128 code points of ASCII characters, up to 4 bytes for others.
- **Examples**
  - German: Walter Schultheiß
  - Hindi: माधुरी दीक्षित
  - Braille: ⠍⠁⠓⠥⠗⠊ ⠙⠊⠅⠎⠓⠊⠞
  - Korean / Chosun: 셍종 / 세종대왕
  - Emoticons: 😀 😉 🤠 🥔 🚚 📸
- **Often**: Encoding issues (assumed local encoding)



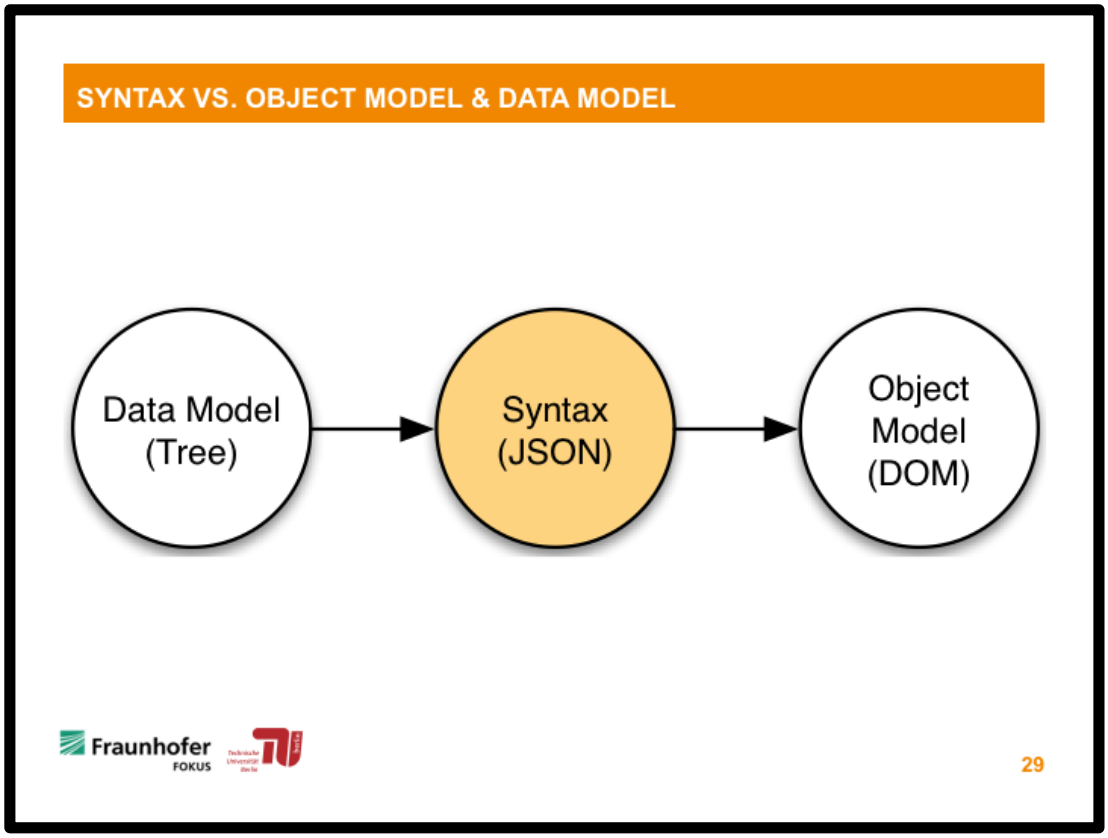
Fraunhofer FOKUS

27

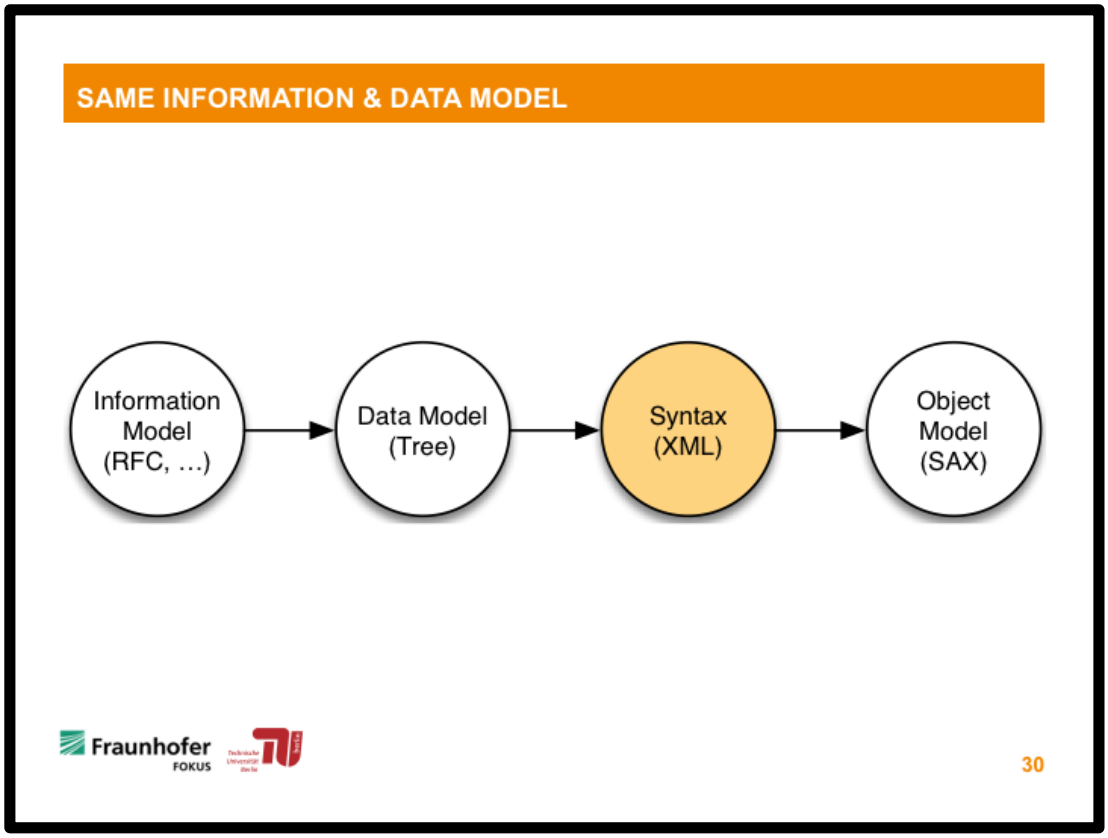
- Basic concept for any serialization
- Unicode is NOT the same as UTF-8
- Encoding issues still nowadays. Yester downloaded a vcf file with German umlauts…



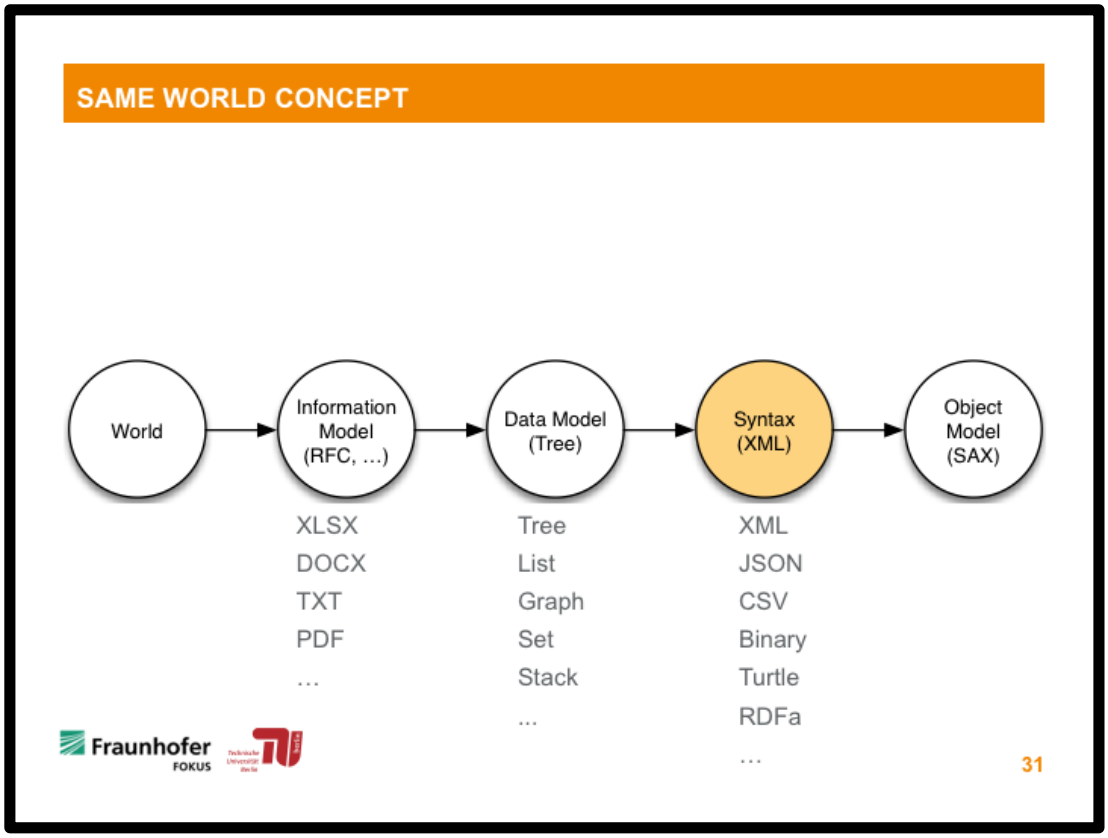
- Typical view as developer, e.g. XML Schema, XML file, SAX object
- SAX=Java API for XML



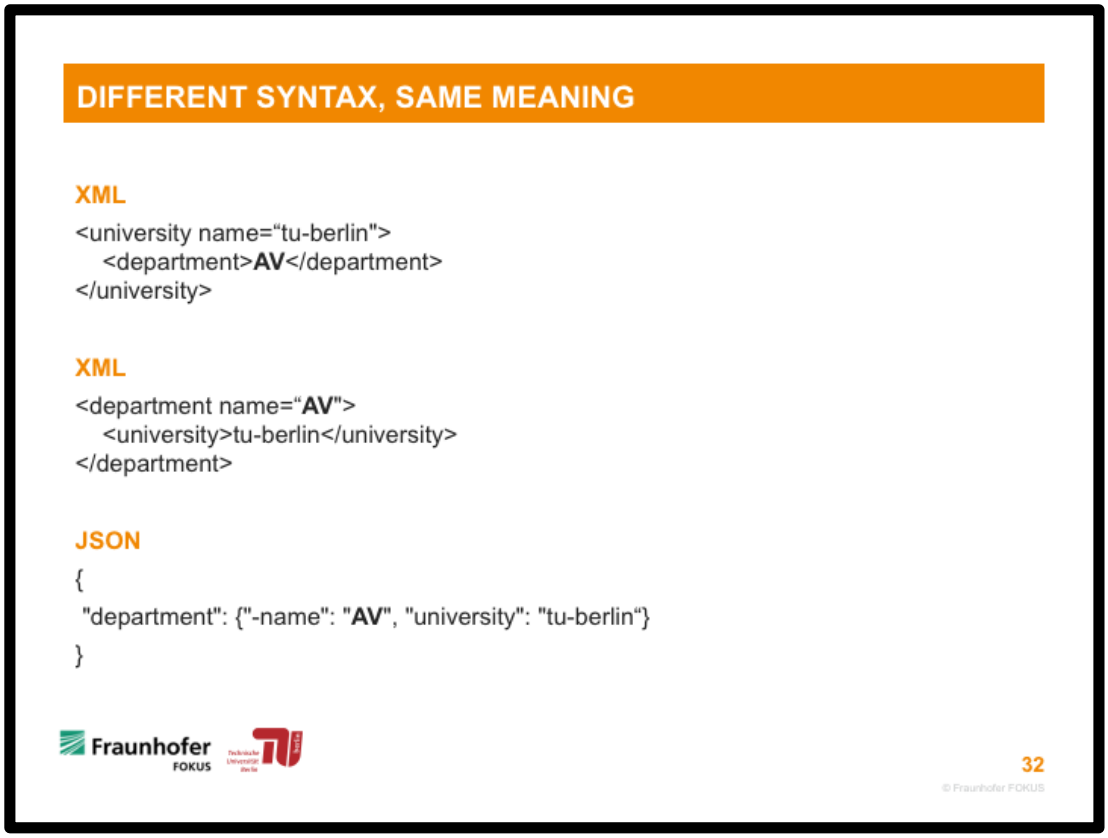
- JSON is just another serialization of a tree data model
- DOM = Document Object Model



- Both might describe the same underlying information



- This information model is an abstraction of a real word concept
- Usually human readable (from humans for humans)
- Examples



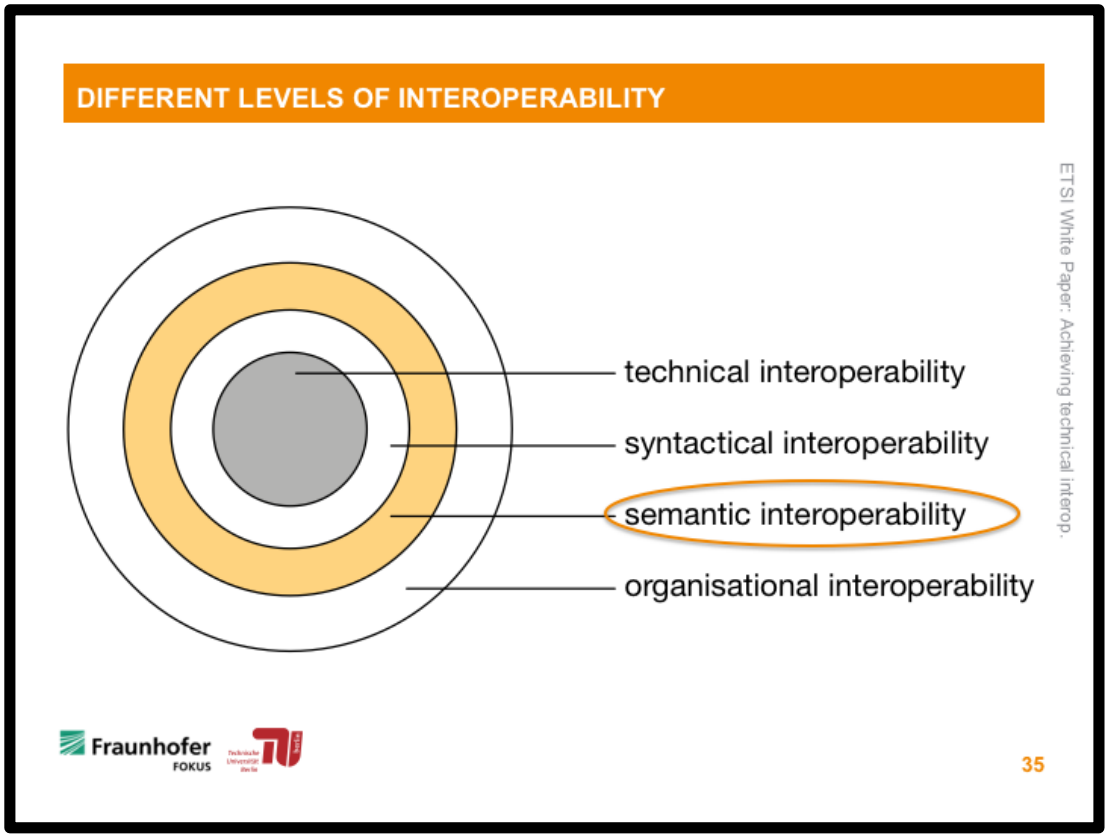
- Same information encoded differently
- Structure implied “semantics”
- At the end, the serialization is not important
- However, for interoperability we either have to use the same serialization or have translators

SYNTACTICAL INTEROPERABILITY

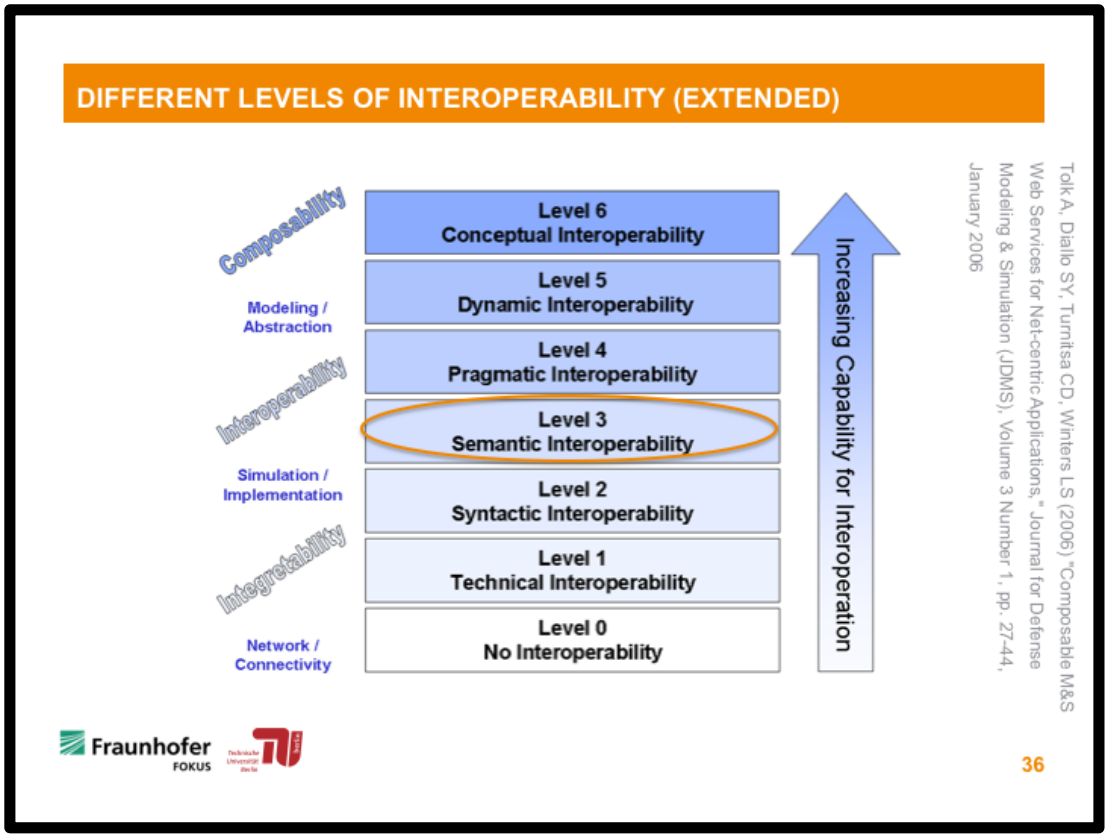
Questions?

SEMANTIC INTEROPERABILITY

12 minutes

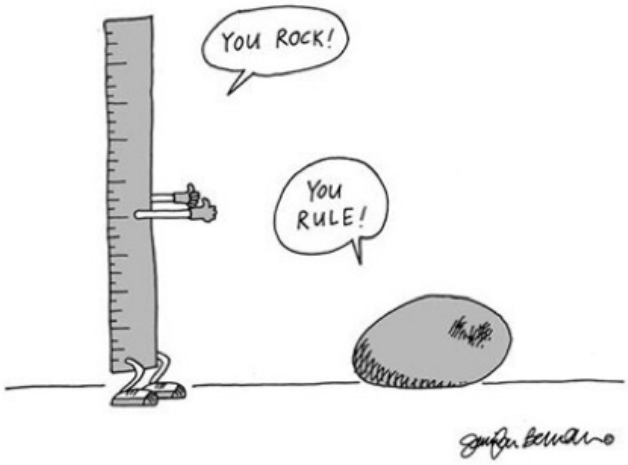


- Finally, the semantic layer
- The last layer we'll cover



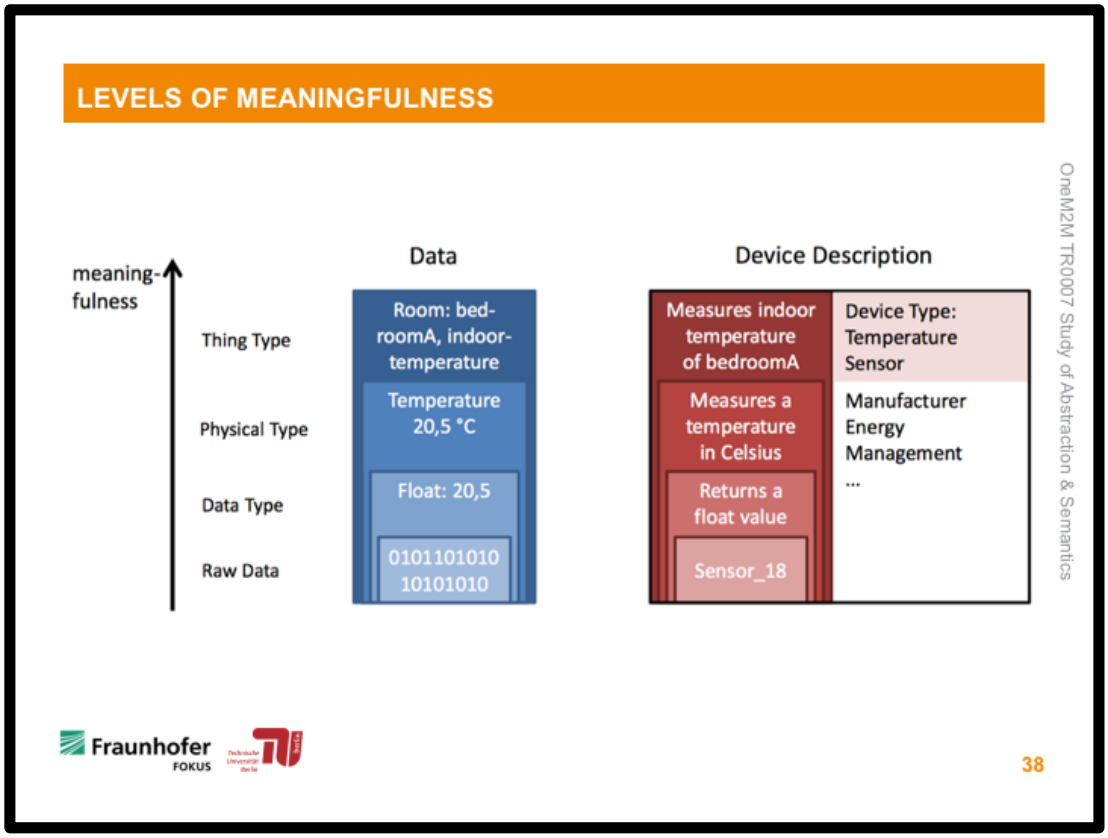
- There are more interesting levels
- Pragmatic interoperability: methods and procedures → Remember: Asset Administration Shell!

IT'S ALL ABOUT SEMANTICS (=MEANING)



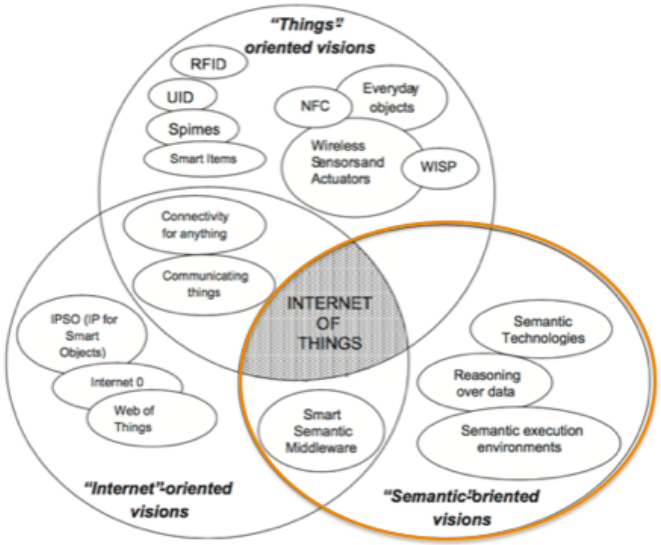
Fraunhofer FOKUS

37



- Raw data: a list of "0" and "1
- Data type: Interpretation of the raw data. Probably content of the basic lectures.
- Physical type: Metadata on the meaning of the data
- Building on this, an ever-growing knowledge graph of metadata can be generated → room temperature in Building X in City Y with Mayor Z born in Year A, as well as writer B, the...

GOT | CONVERGENCE OF DIFFERENT VISIONS

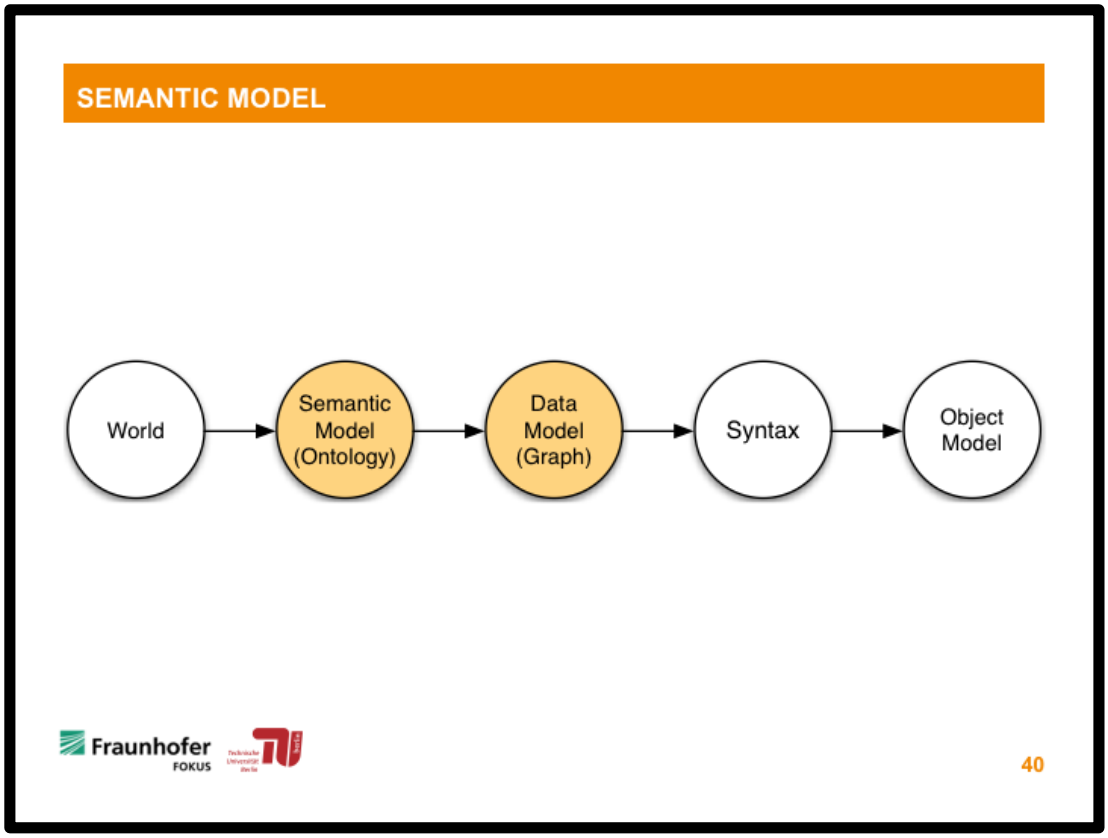


Atzori, Luigi: The Internet of Things: A survey

Fraunhofer FOKUS

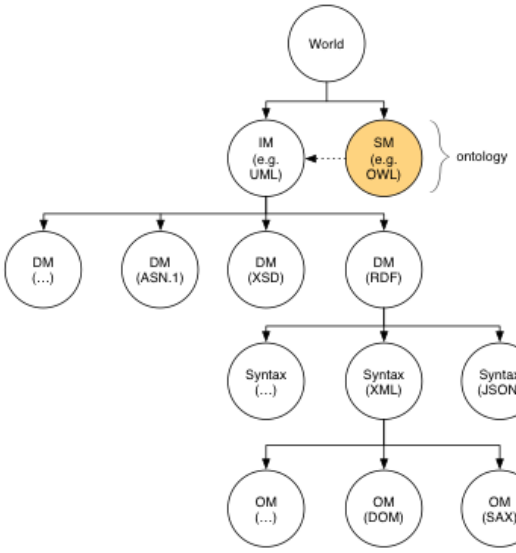
39

- Another part of the overall IoT visions: IP + Things + Semantics
- Note the different terminology used in this example!



- Moving from human readable information models to formal information models
- Moving from lists or trees to graph data models

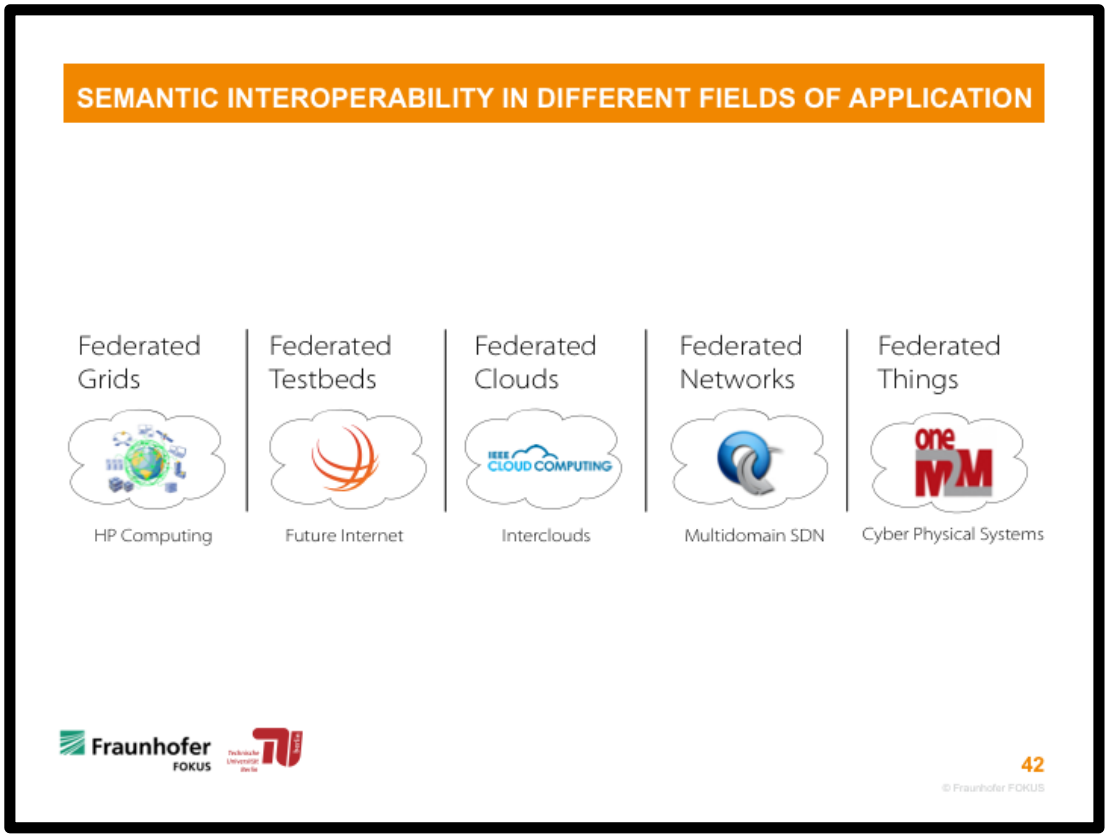
SEMANTIC MODELS



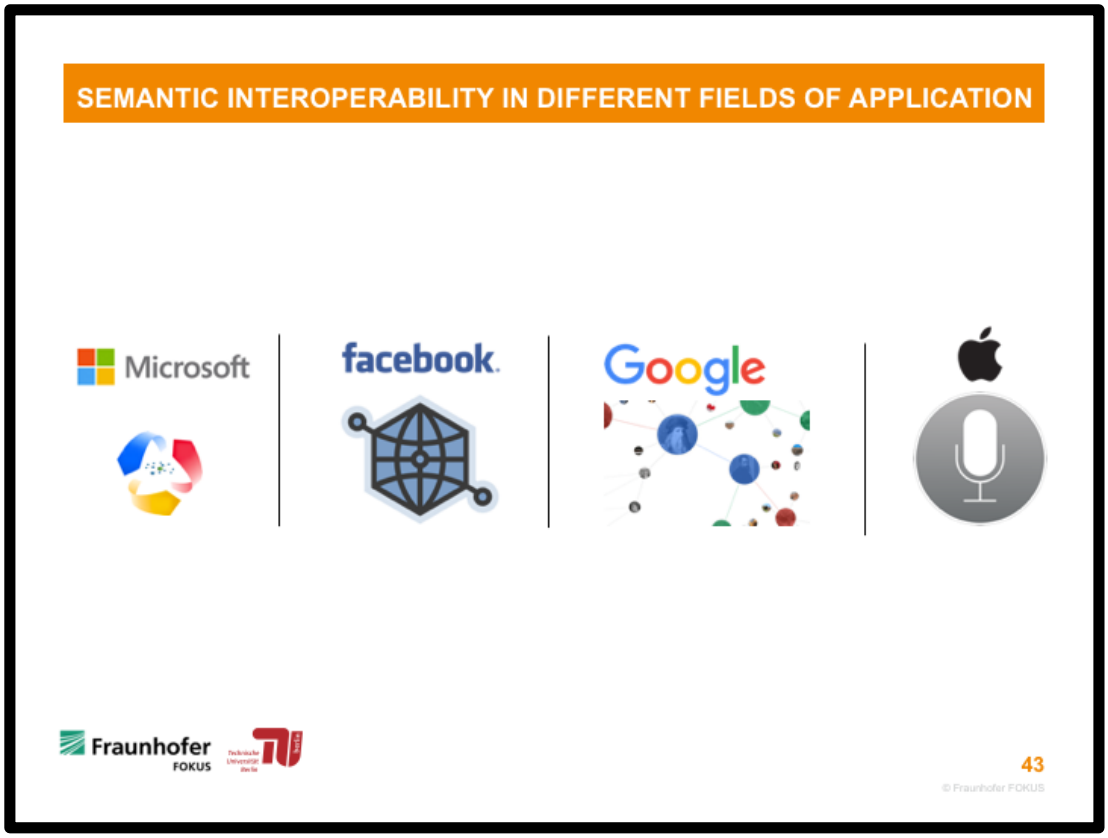
Fraunhofer FOKUS

41

- Another view
- UML is already a formal information model
- So is the OPC UA information model
- However, with semantic information models we move towards symbolic KI capabilities

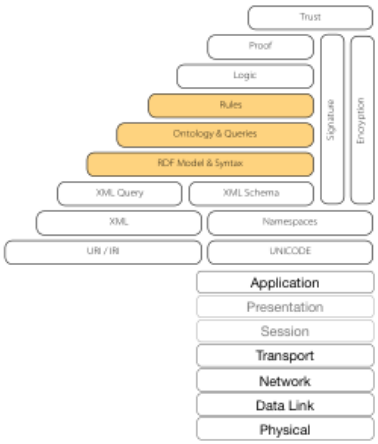


- Use of semantic information models in different fields of application



- Use of semantic information models in different products

THE SEMANTIC WEB IS ONE POSSIBLE IMPLEMENTATION



Fraunhofer FOKUS

44

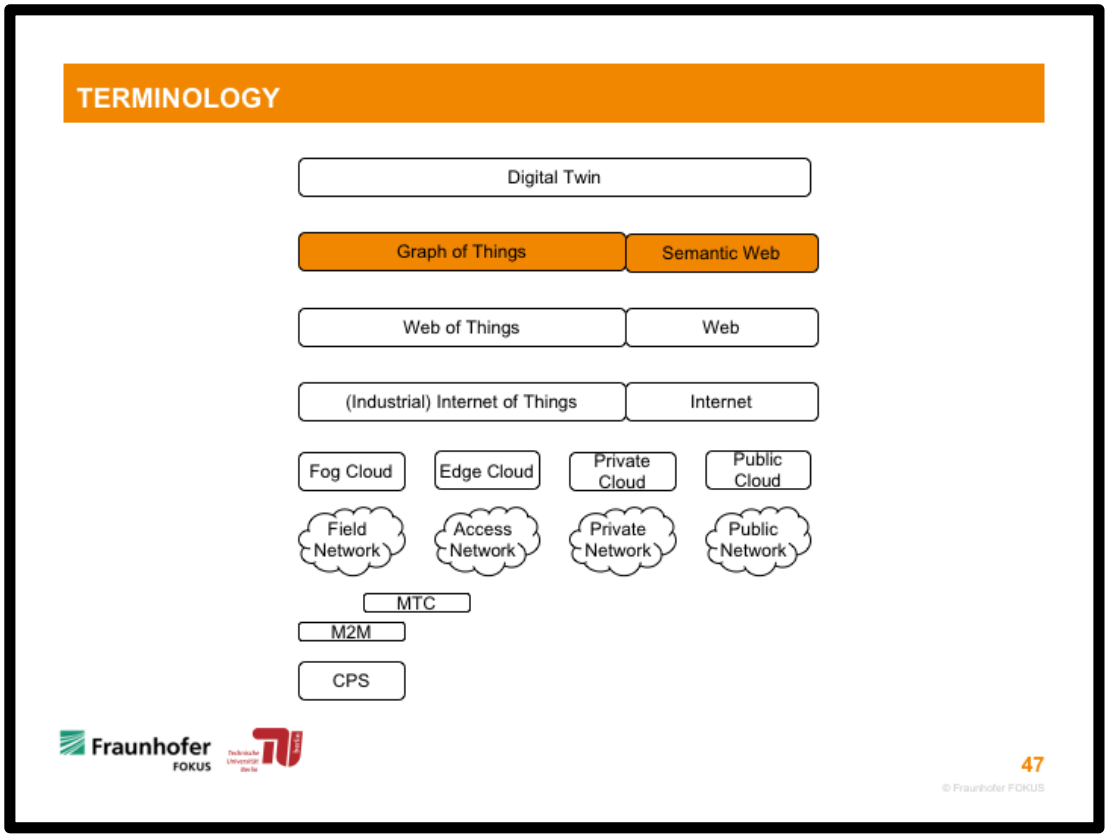
- The Semantic Web is ONE possible implementation of the related subsymbolic KI mechanisms

SEMANTIC INTEROPERABILITY

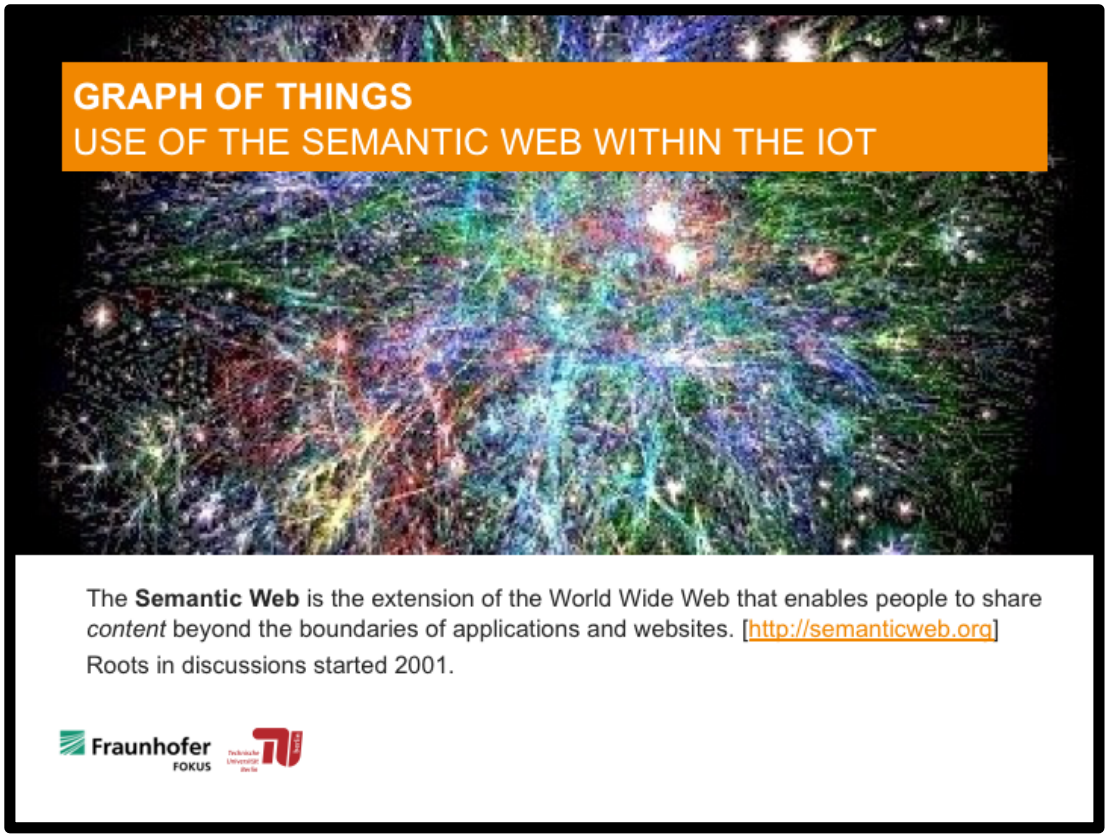
Questions?

THE SEMANTIC WEB

8 minutes



- The application of semantic web technologies for interoperable communication within the Internet of Things



- It's origin from the web community
- Formal languages as such have a way longer history

WHERE IS THE SEMANTIC WEB? (2010)

Fraunhofer FOKUS

- CAN by applied to web pages, but it's ONLY ONE example

WHERE IS THE SEMANTIC WEB?

Fraunhofer FOKUS

- Annotated within HTML → RDFa serialization (RDF in attributes)

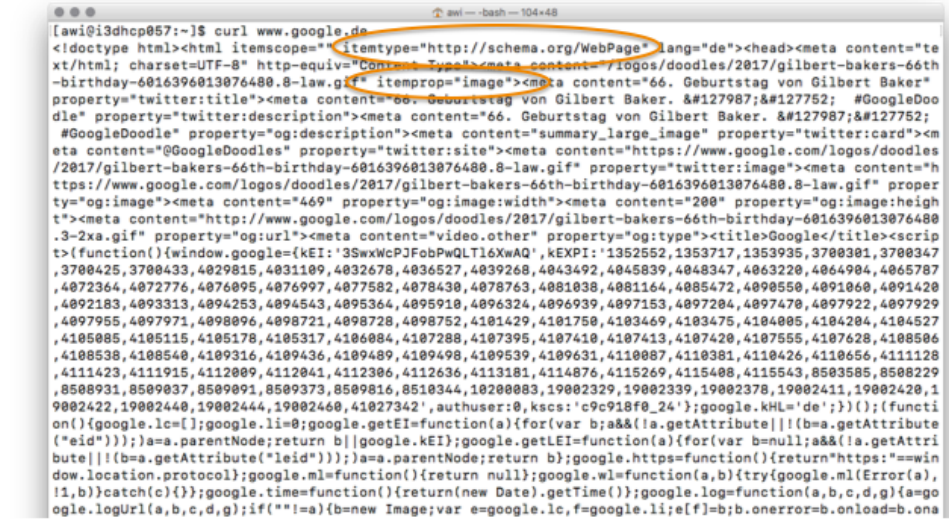
WHERE IS THE SEMANTIC WEB?

Fraunhofer FOKUS

51

- Also modern web pages are using RDFa

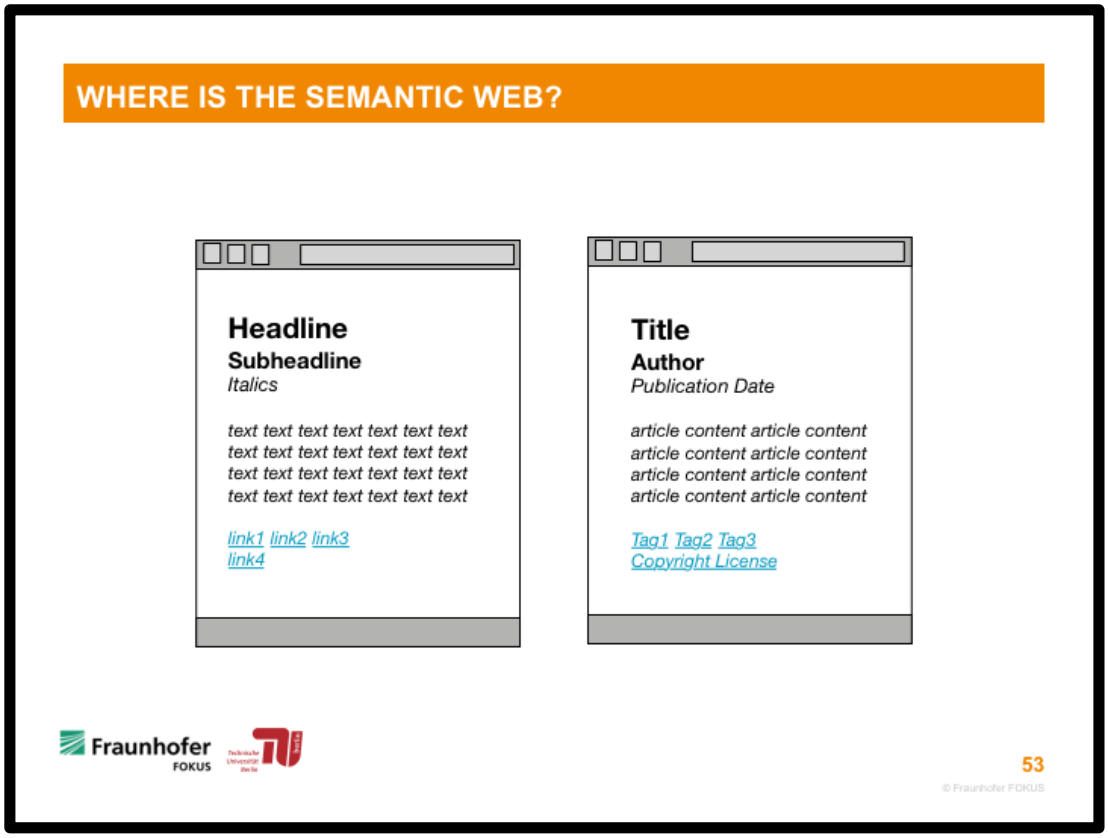
WHERE IS THE SEMANTIC WEB?



Fraunhofer FOKUS

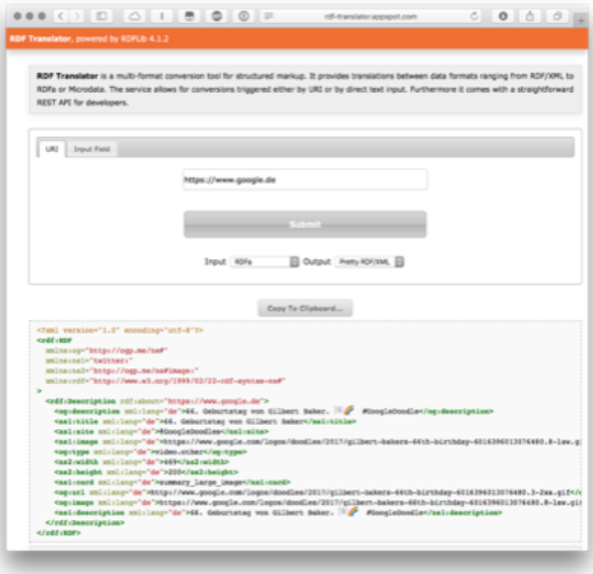
52

- Again: ONE possible application of Semantic Web Technologies are web pages (=largely distributed, heterogeneous, federated and linked resource database)
- NOT limited to web pages



- Needed as soon knowledge representation is directed towards machines, not humans -> IoT

CONVERTING FROM ONE SERIALIZATION TO ANOTHER



http://rdf-translator.appspot.com

Fraunhofer FOKUS

54

- Agnostic towards its serialization: RDFa, XML, N3, JSON, …
- Remember the last slide about syntactic interoperability

THE SEMANTIC WEB

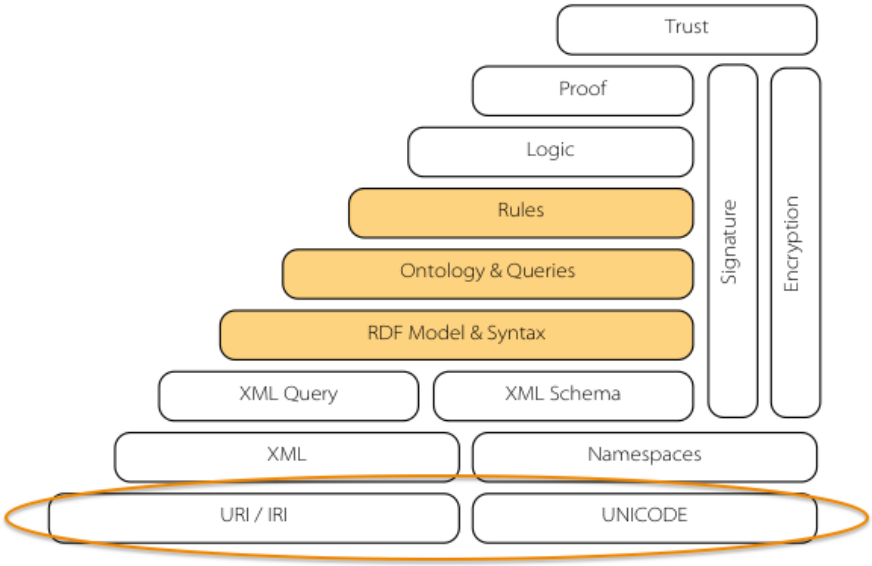
Questions?

SEMANTIC WEB LAYER CAKE

URI

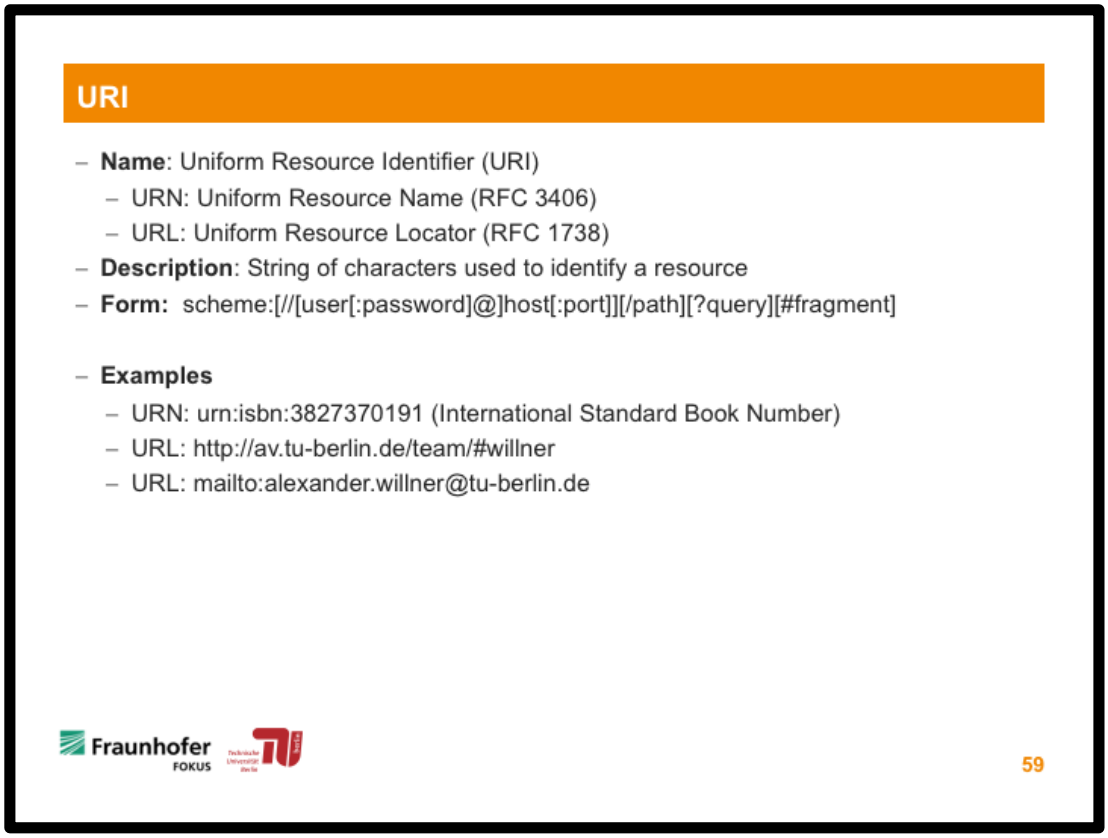
3 minutes

SEMANTIC WEB LAYER CAKE



Fraunhofer FOKUS

58



- Core concept is the unique identification of concepts in a largely distributed manner

IRI

- **Name**: Internationalized Resource Identifier (RFC 3987)
- **Description**: String of characters used to identify a resource using Unicode.

- **Example** (percent-encoding)
  - https://en.wiktionary.org/wiki/Πόδος → https://en.wiktionary.org/wiki/%E1%BF%AC%CF%8C%CE%B4%CE%BF%CF%82

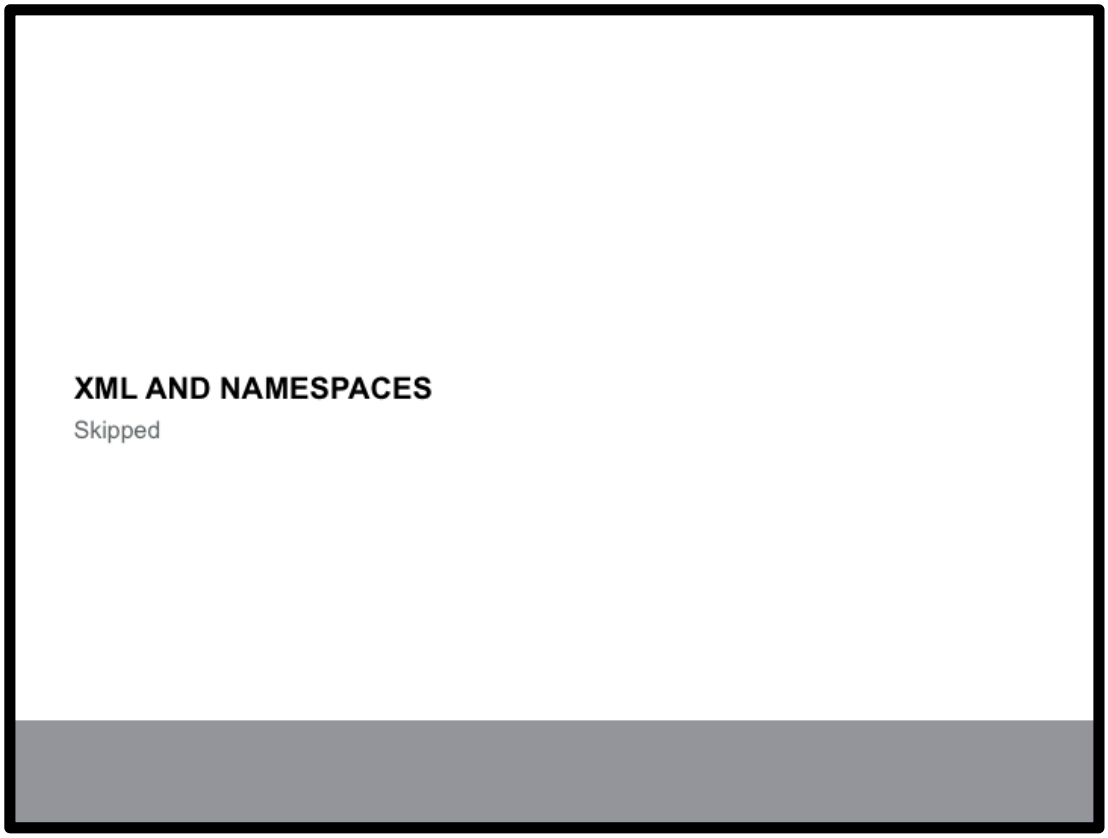
- **Examples** (IDN - Internationalized Domain Name, punycode translation)
  - http://例子.卷筒纸 → http://xn--fsqu00a.xn--3lr804guic/
  - http://www.paypal.com → http://www.xn--80aa0cbo0j.com (not www.paypal.com)
  - http://example.com⁄a-top-level-domain.com/ (not http://example.com/a-top-level-domain.com/

Fraunhofer FOKUS

60

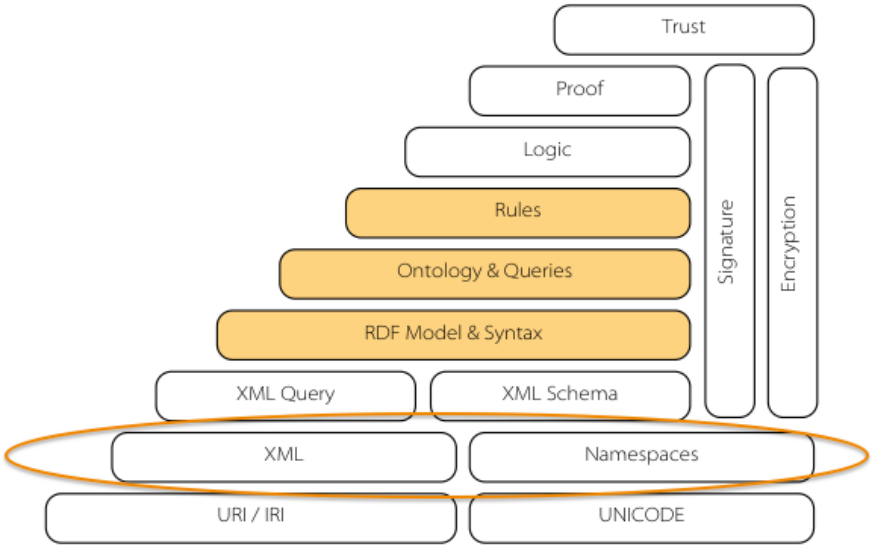
URI

Questions?



- As the serialization doesn’t matter

SEMANTIC WEB LAYER CAKE



Fraunhofer FOKUS

63

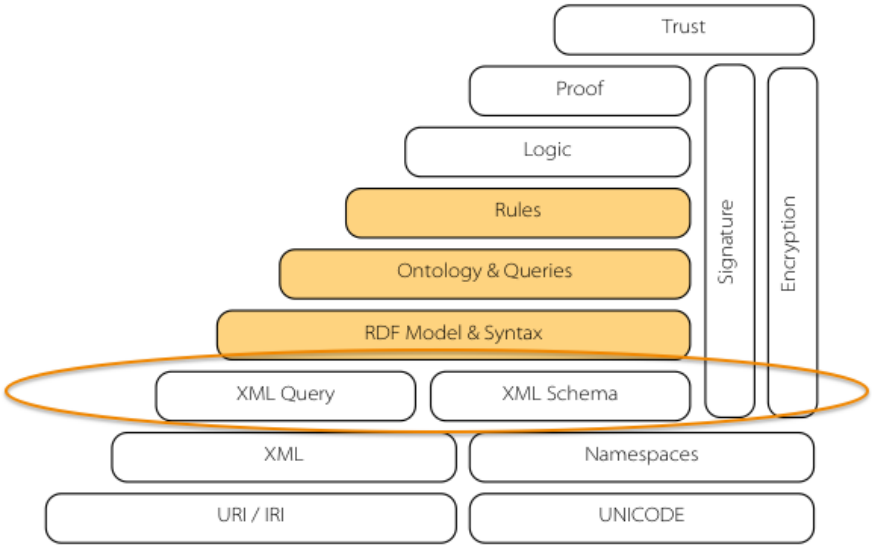
XML AND NAMESPACES

Skipped

XML QUERY AND SCHEMA

Almost skipped

SEMANTIC WEB LAYER CAKE



Fraunhofer FOKUS

67

JSON SCHEMA

- **JSON Schema**: Defines the syntax of a JSON document (it's a tree structured data model)
- **Example JSON**:
  - { "name": "George Washington",
    "birthday": "February 22, 1732"}
  - { "first_name": "George",
    "last_name": "Washington",
    "birthday": "1732-02-22"}
- **Example JSON Schema:**
  - ```
    { "type": "object",
      "properties": {
        "first_name": { "type": "string" },
        "last_name": { "type": "string" },
        "birthday": { "type": "string", "format": "date-time" }
      }
    }
    ```
- **Basic types:** string, integer, number, object, array, boolean, null

Fraunhofer FOKUS

69

- As seen in the WoT TD example
- To be used for simple verification processes
    ```

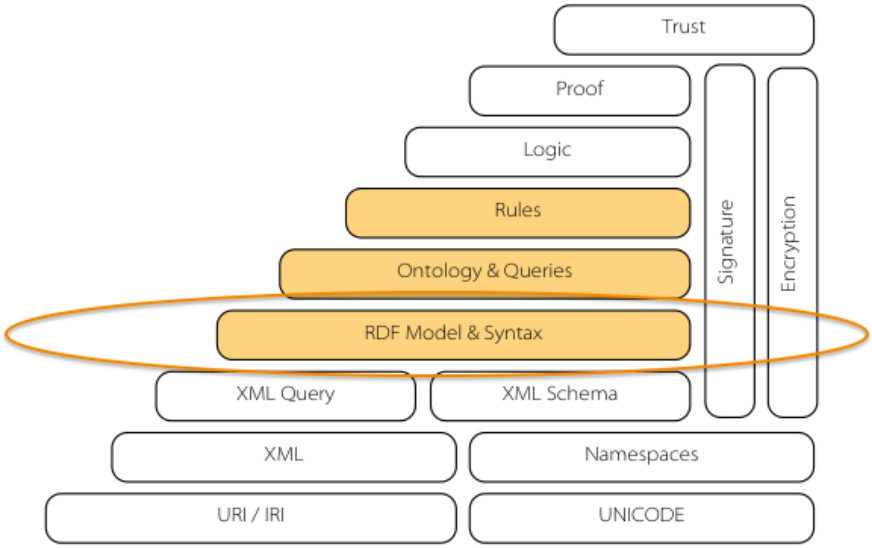
XML QUERY AND SCHEMA

Questions?

RDF

10 minutes

SEMANTIC WEB LAYER CAKE



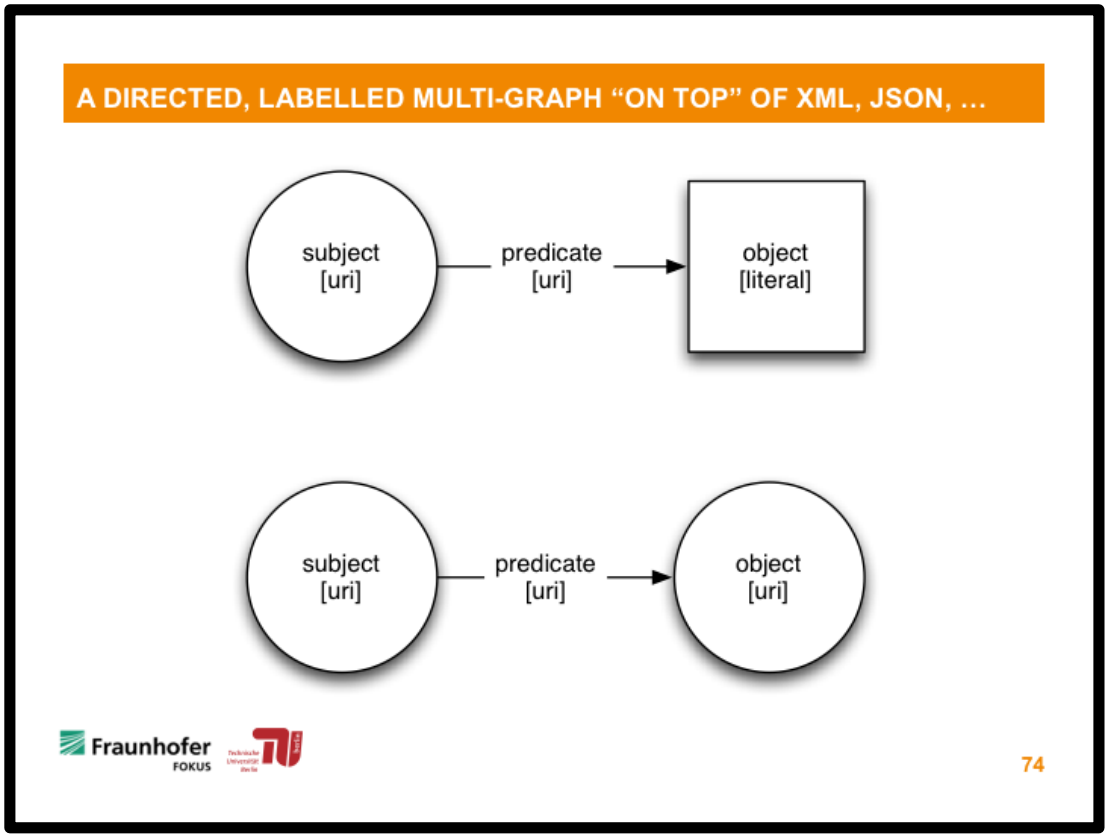
Fraunhofer FOKUS

72

RDF

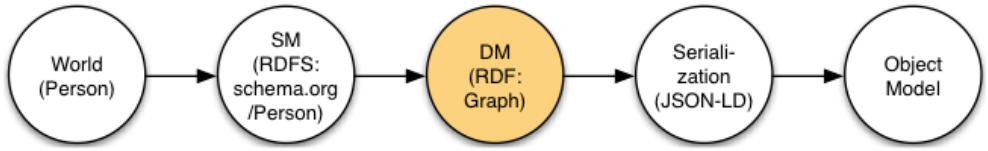
- **Name**: Resource Description Framework
- **Model**: It's just a labeled, directed multi-graph (not an XML tree)
- **Idea**: Triplets using URIs (Subject - Predicate - Object)
- **Advantages**: RDF is actually both, a graph data model and an initial vocabulary to define types and relationships (first simple form an ontology)
- **Examples**:
  - <urn:nasa:sensor:hubble> <rdf:type> <http://nasa.gov/resource/type/telescope>
  - <mail:alexander.willner@tu-berlin.de> <foaf:name> "Alex"

Fraunhofer FOKUS



- The object can either be a literal (concrete data, no extension from here) or the object can be a URI (possible extensions from here)
- **Important**: this combination (subject-predicate-object) is also called a triplet

RDF IS A GRAPH DATA MODEL



Fraunhofer FOKUS

75

- RDFS = RDF Schema
- http://schema.org/Person formally describes what a person is

RDF

- **Serialization**:
  - XML (application/rdf+xml)
  - N3 (Notation 3)
  - Turtle
  - JSON-LD
  - RDFa
  - ...
- **Disadvantage**: RDF defines mainly the way how to write information and only partly what to write.
- **Example** (in turtle):

```
:alex :talksWith :duck .
```

Fraunhofer FOKUS

76

- RDF is a model, not a syntax

RDF

- **Example** (JSON-LD)

```
{  "@context": "http://schema.org/",
   "@type": "Person",
   "name": "Alexander Willner",
   "url": "http://av.tu-berlin.de/willner"
}
```

- **Example** (N3)

```
[] a schema:Person ;
   schema:name "Alexander Willner" ;
   schema:url <http://av.tu-berlin.de/willner> .
```

- **Example** (RDF/XML)

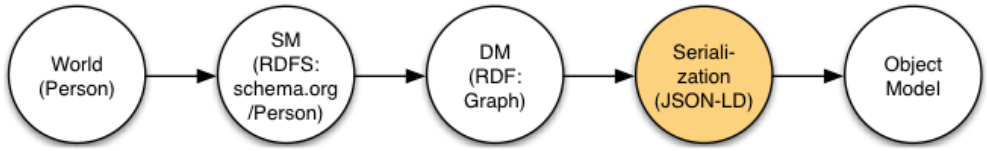
```
<rdf:RDF>
 <rdf:Description rdf:nodeID="xxxx">
  <rdf:type rdf:resource="http://schema.org/Person"/>
  <schema:name>Alexander Willner</schema:name>
  <schema:url rdf:resource="http://av.tu-berlin.de/willner"/>
 </rdf:Description>
</rdf:RDF>
```

Fraunhofer FOKUS

77

- Three different serializations of the same information
- JSON-LD example -> see W3C WoT

JSON-LD IS ONE POSSIBLE SERIALIZATION OF AN RDF GRAPH



Fraunhofer FOKUS

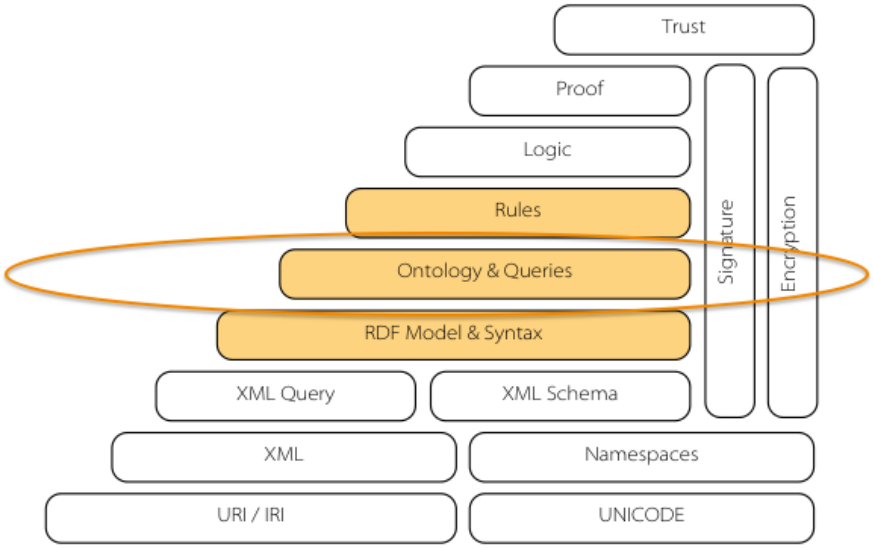
RDF

Questions?

ONTOLOGIES AND QUERIES

15 minutes

SEMANTIC WEB LAYER CAKE



Fraunhofer FOKUS

81

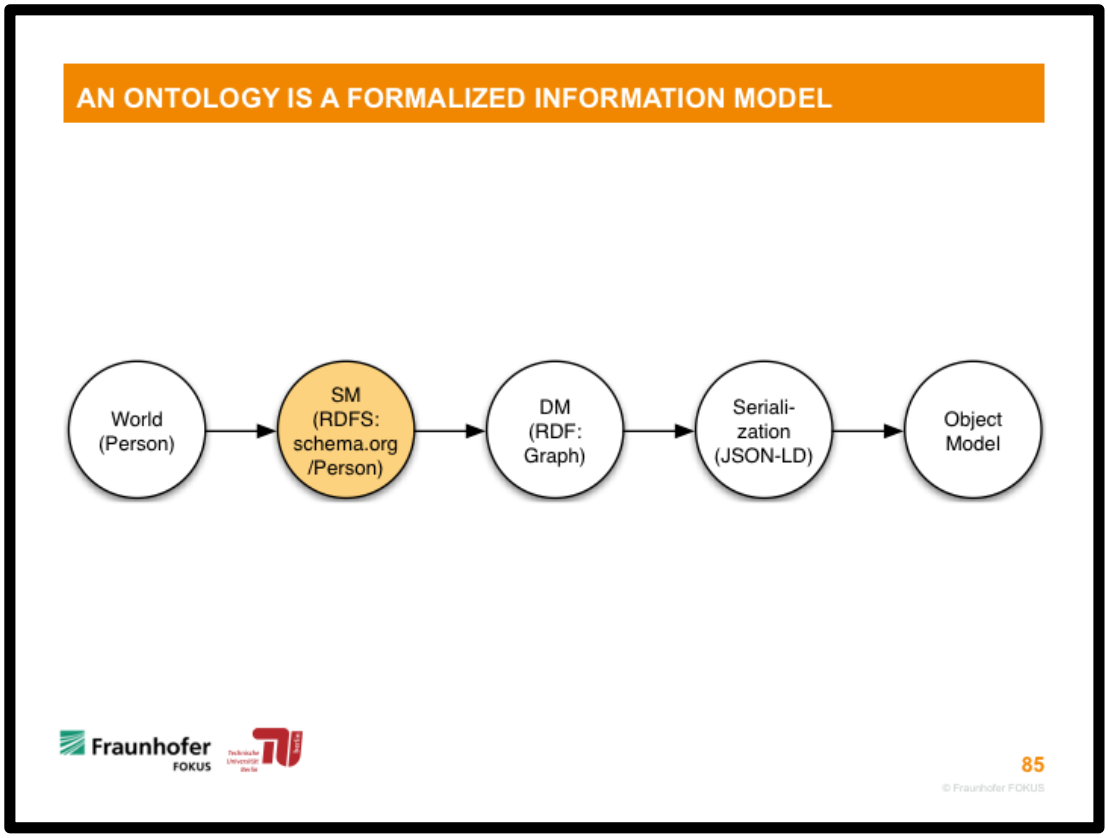
ONTOLOGIES | DEFINITION

- **In General**: The terminology and concepts are often not well used (fuzzy)
- **Vocabulary**: list of unambiguous, non-redundant explicit terms (maybe w/o meaning)
- **Taxonomy**: organized vocabulary into a (poly-)hierarchical structure
- **Thesaurus**: networked vocabulary using associative relationships
- **Meta-Model**: rules on how to construct models within a domain
- **Definition**: An ontology
  - formally represents knowledge
  - as a hierarchy of concepts
  - within a domain,
  - using a shared vocabulary
  - to denote the types, properties and interrelationships
  - of those concepts.

Fraunhofer FOKUS

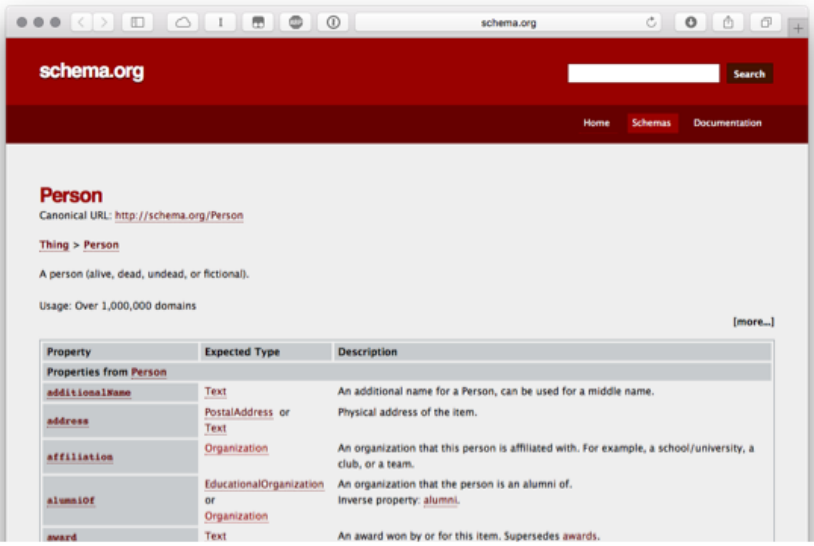
ONTOLOGIES | RDFS

- **Ontology**: Describes the semantic – the meaning of a resource.
- **Resource Description Framework (RDF)**:
  Has already limited ontology support by providing **rdf:type**.
- **Resource Description Framework Schema (RDFS)**
  - Officially the "RDF Vocabulary Description Language"
  - The term "Schema" is retained for historical reasons.
  - Defines legal uses of various classes and relationships.
  - Introduces for example rdfs:subClassOf or rdfs:range.
- **Example** – schema.org:
  "[…] a collaborative […] activity […] to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond." (schema.org)



- http://schema.org/Person formally describes what a person is

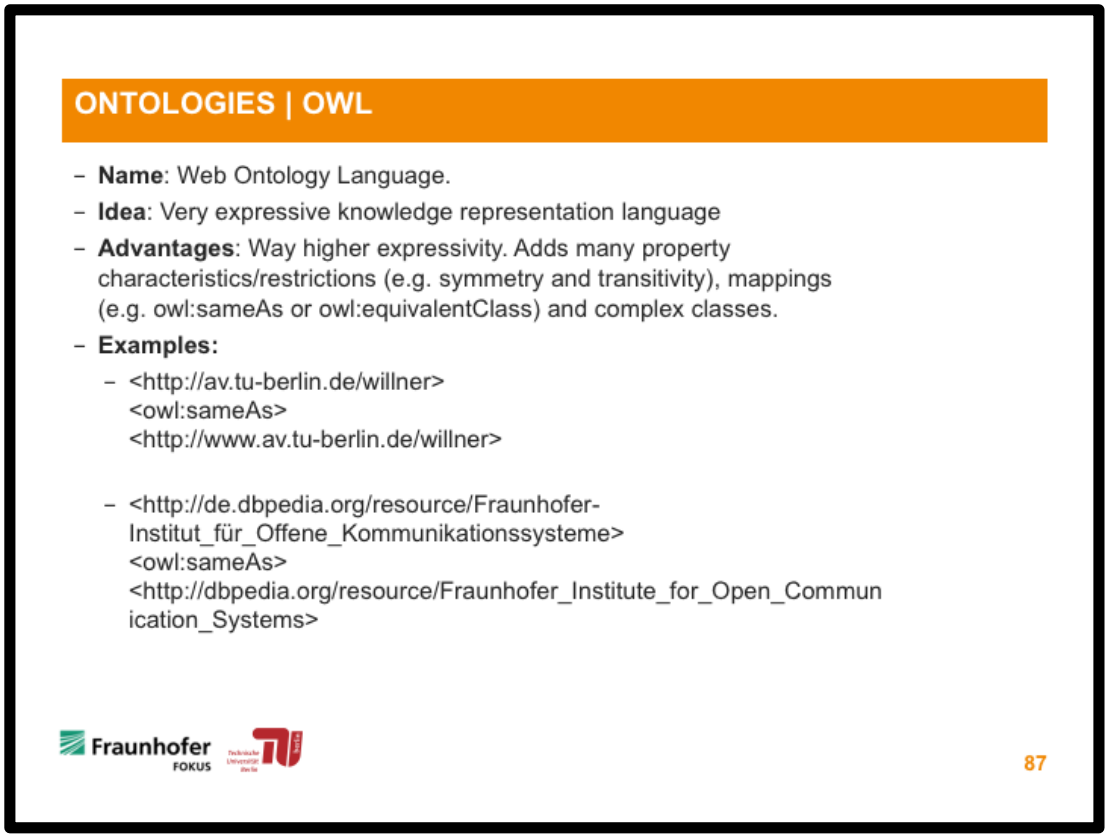
VOCABULARY: SCHEMA.ORG



http://schema.org

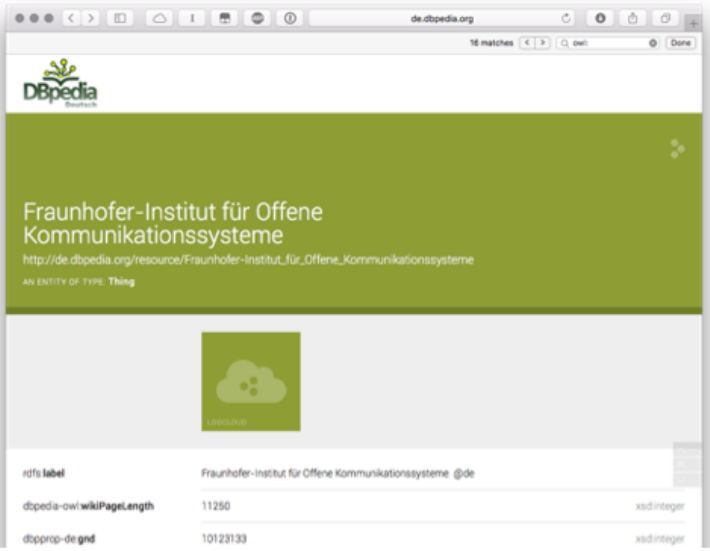
Fraunhofer FOKUS

86



- OWL is based on description logics, that are a family of logics that are decidable fragments of first-order logic

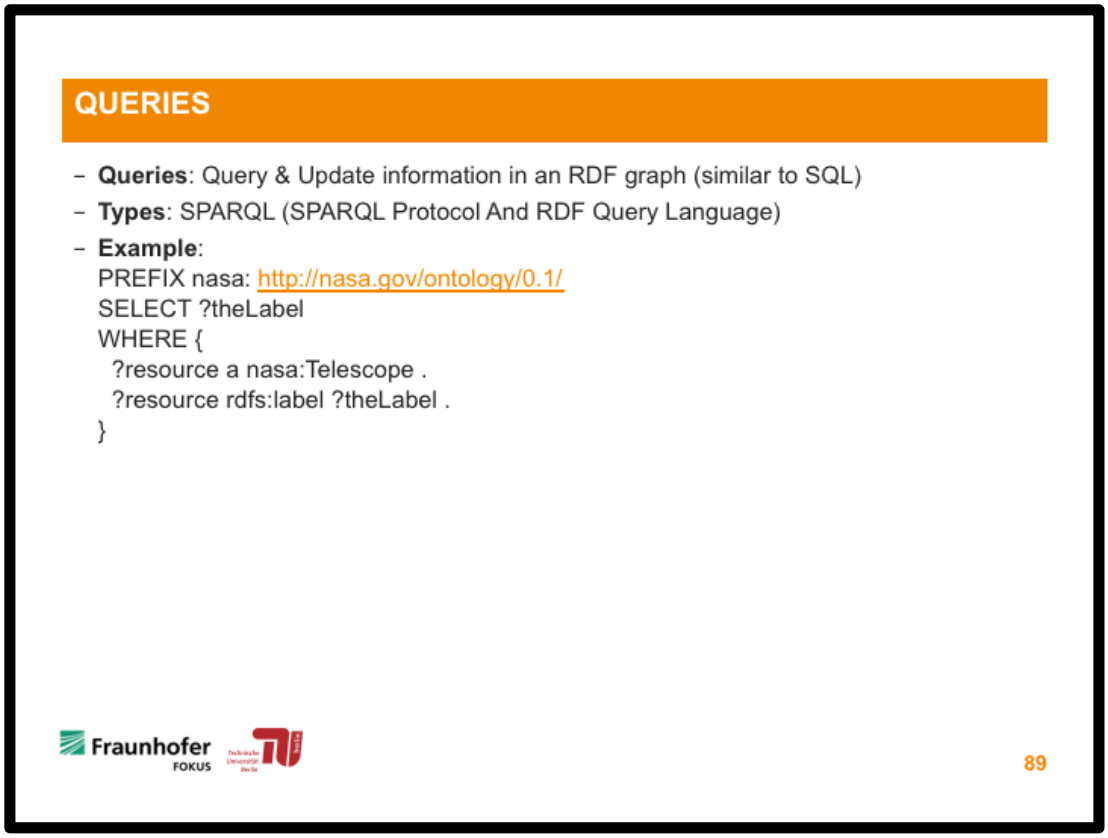
REPOSITORY: DBPEDIA.ORG



http://dbpedia.org

Fraunhofer FOKUS

88

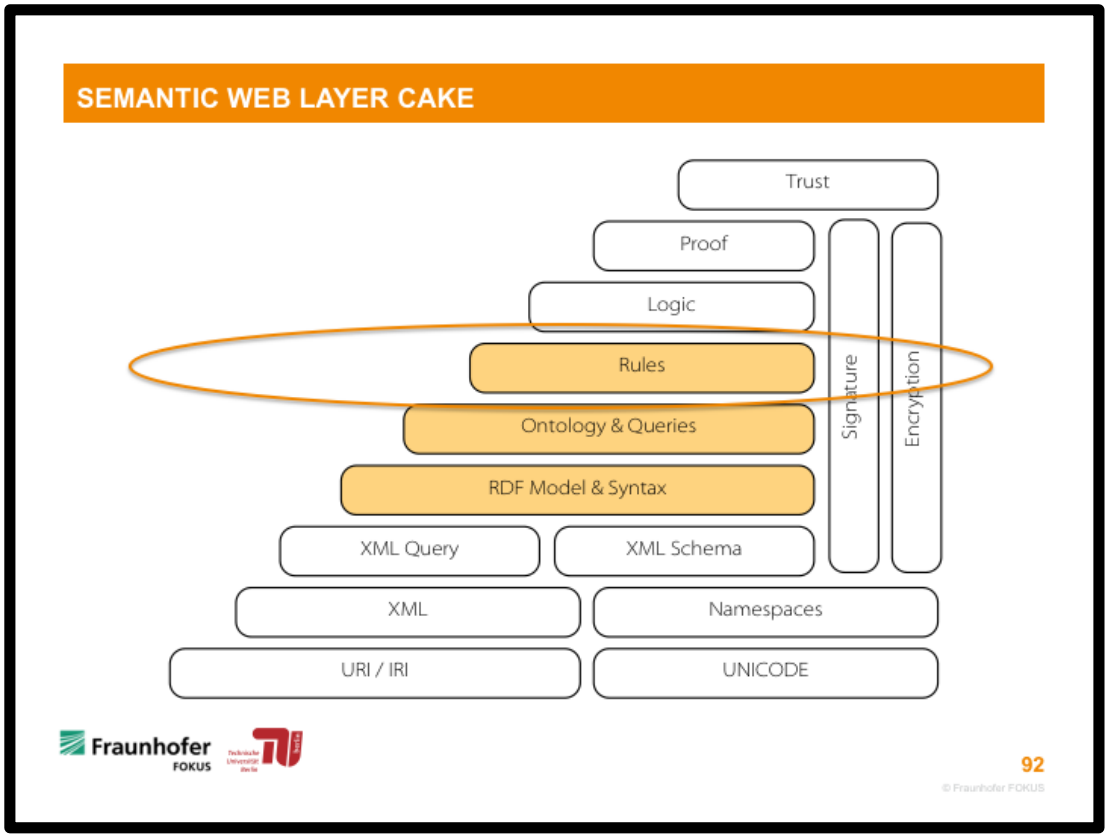


- More powerful than SQL (e.g. it support the search for transitive relations in the database)

ONTOLOGIES AND QUERIES

RULES

5 minutes



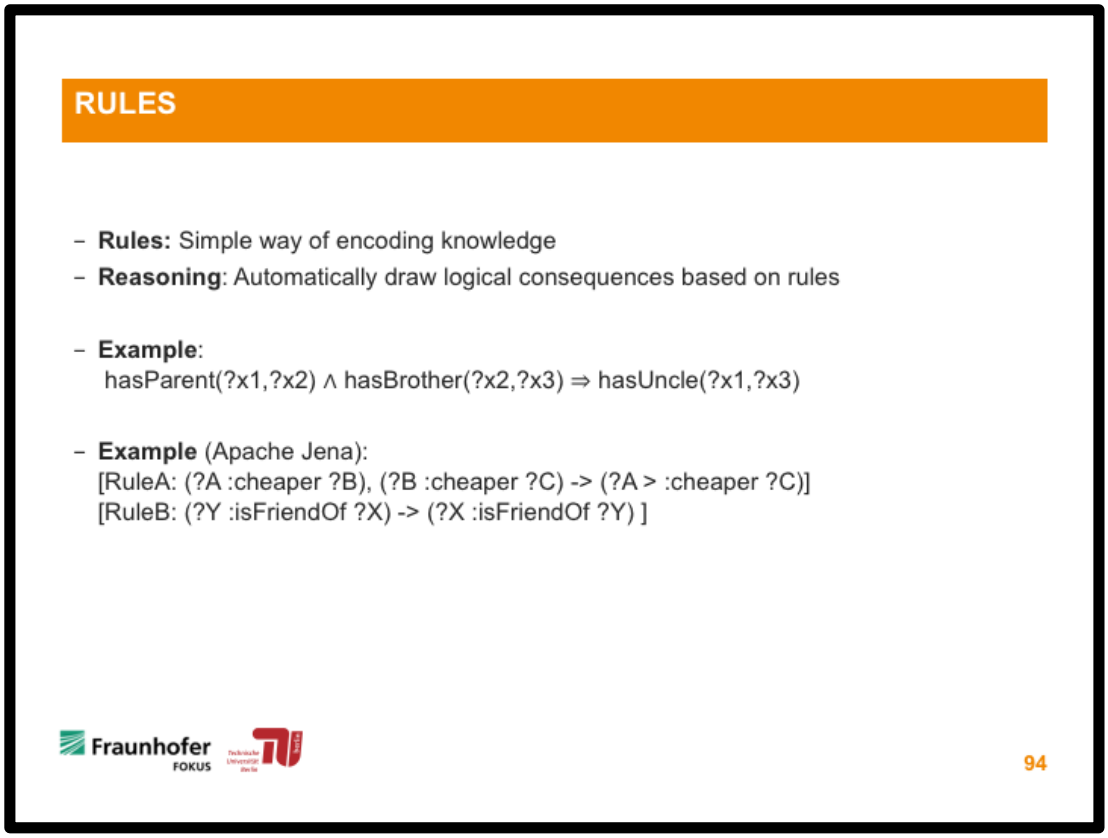
Underlying concept is formal logic: propositional logic, predicate logic, description logic, rule-based systems, …

RULES

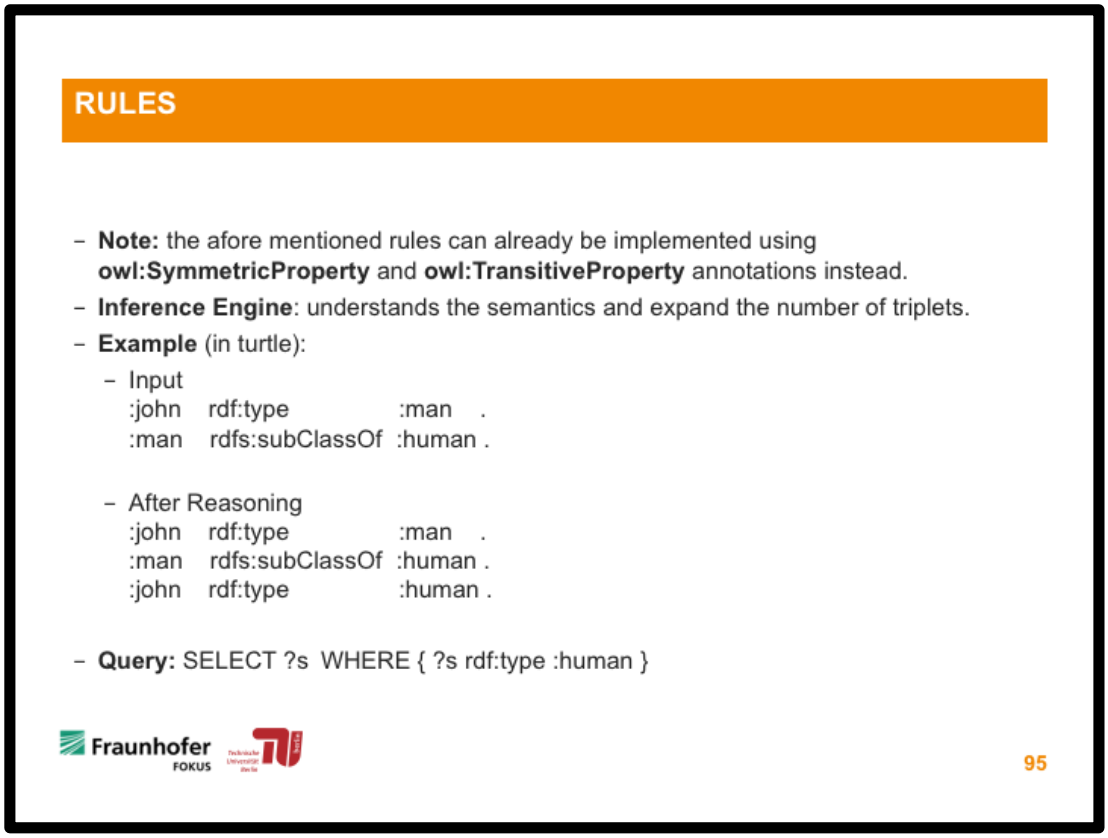
Fraunhofer FOKUS

93

- Based on this formal knowledge representation we can now infer implicit knowledge using reasoning
- In this example we also gain the knowledge that penguins are not very good in reasoning



- RuleA → transitiveRule
- RuleB → symmetricRule
- It's possible to infer knowledge that has not been described explicitly in the graph



- So after the reasoning process I can query for all object of type ":human" and will get back ":john", although this was not defined explicitly beforehand

RULES

Questions?

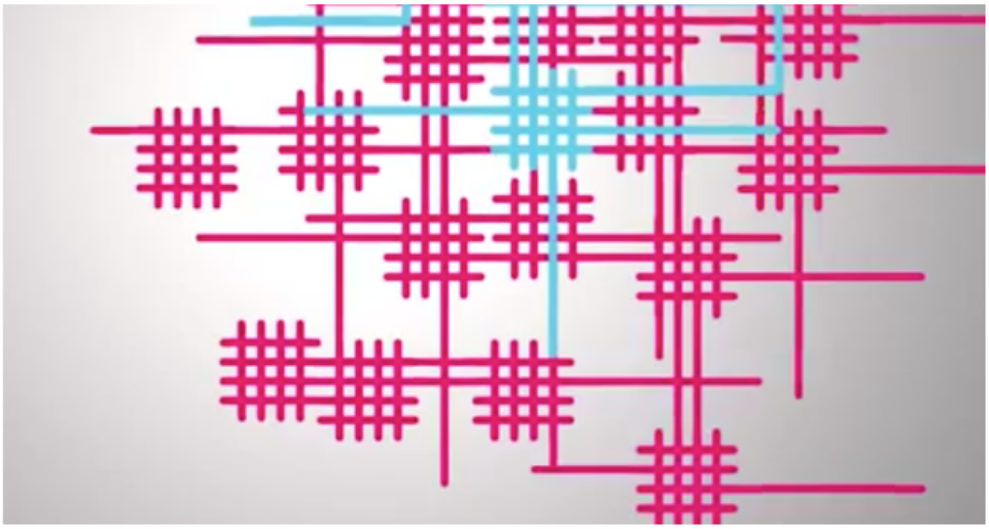
LINKED OPEN DATA

10 minutes

- To building formal knowledge graphs we need data

- we need (ideally open) linkable, formal data

LINKED OPEN DATA (4 MIN)



https://youtu.be/ijju4wT9uBIA

Fraunhofer FOKUS

5 STAR LINKED OPEN DATA

Tim Burners Lee

Fraunhofer FOKUS

5 STAR LINKED OPEN DATA


Available on the web (whatever format), but with an open license (e.g. an image scan)


Available as machine-readable structured data (e.g., excel instead of an image scan)


As before, but using a non-proprietary format (e.g., CSV instead of excel)


All the above, plus use open standards (RDF & Co.) to identify things, so that people could point at your stuff


All the above, plus link your data to other people's data to provide context

Fraunhofer FOKUS

Ivan Herman, W3C, Tutorial on Semantic Web

102

LINKED RDF

Rules

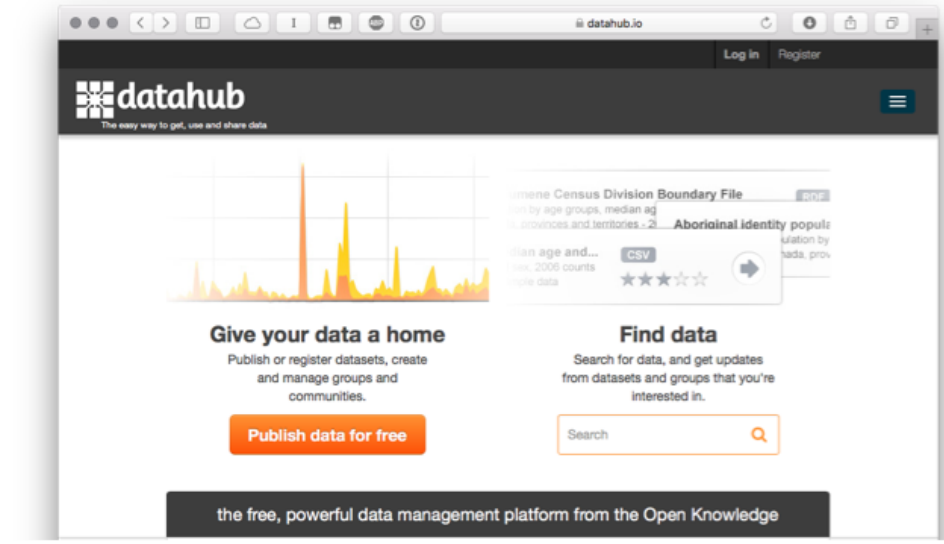
1. All kinds of conceptual things have names that start with HTTP.
2. Looking up these HTTP names I will get back some data in a standard format.
3. That information got relationships to other things that is related and starts with HTTP.

Basics

1. Use URLs as names for things
2. Use HTTP URLs so that people can look up those names.
3. Provide useful information about what a name identifies (RDF, SPARQL, etc.)
4. Include links to other URLs. So that they can discover more things.

Fraunhofer FOKUS

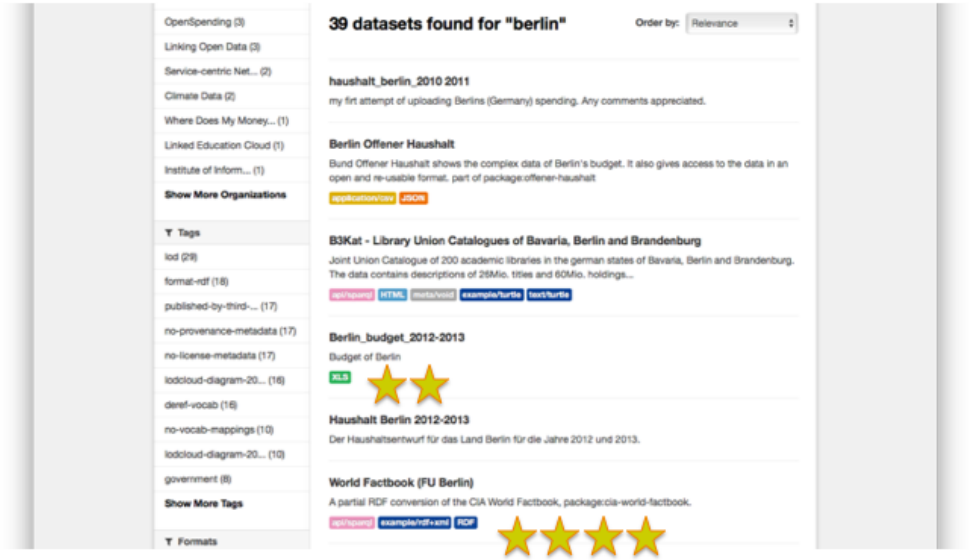
DATA HUB REPOSITORY FOR OPEN DATA



Fraunhofer FOKUS

104

DATA HUB REPOSITORY FOR OPEN DATA



Fraunhofer FOKUS

105

LINKED OPEN DATA CLOUD DIAGRAM

http://lod-cloud.net

Fraunhofer FOKUS

106

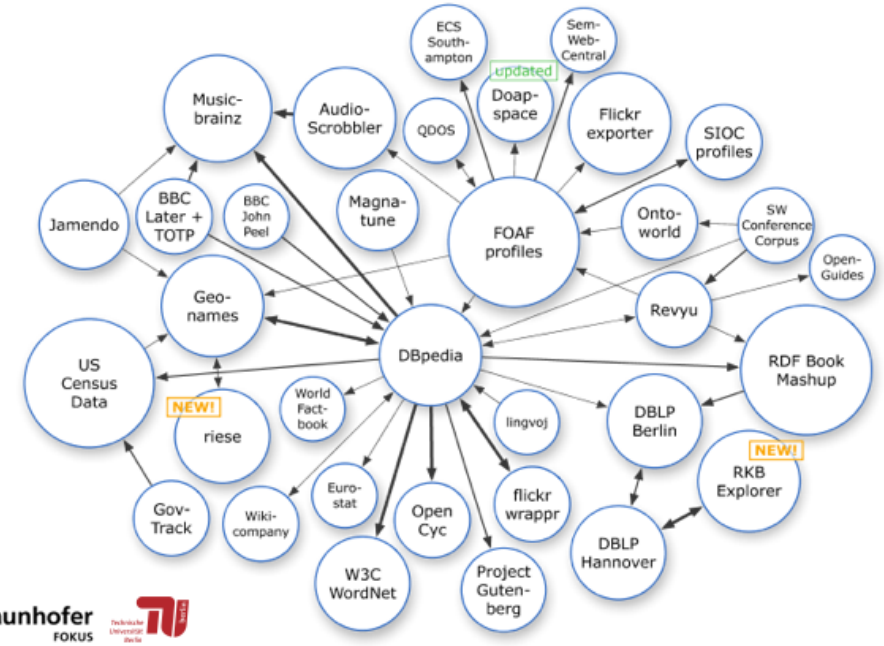
LOD CLOUD REQUIREMENTS

"[…] The LOD cloud diagram […] shows datasets that have been published in Linked Data format […]. It is based on metadata collected and curated by contributors to the Data Hub." (http://lod-cloud.net)

1. There must be *resolvable http:// (or https://) URIs*.
2. They must resolve, with or without content negotiation, to *RDF data*.
3. The dataset must contain *at least 1000 triples*.
4. The dataset must be connected via *RDF links* to a dataset that is already in the diagram.
5. Access of the *entire* dataset must be possible via *RDF crawling*, *dump*, or *SPARQL*.

Fraunhofer FOKUS

LINKED OPEN DATA (LOD) CLOUD DIAGRAM (2008, > 30 DATASETS)

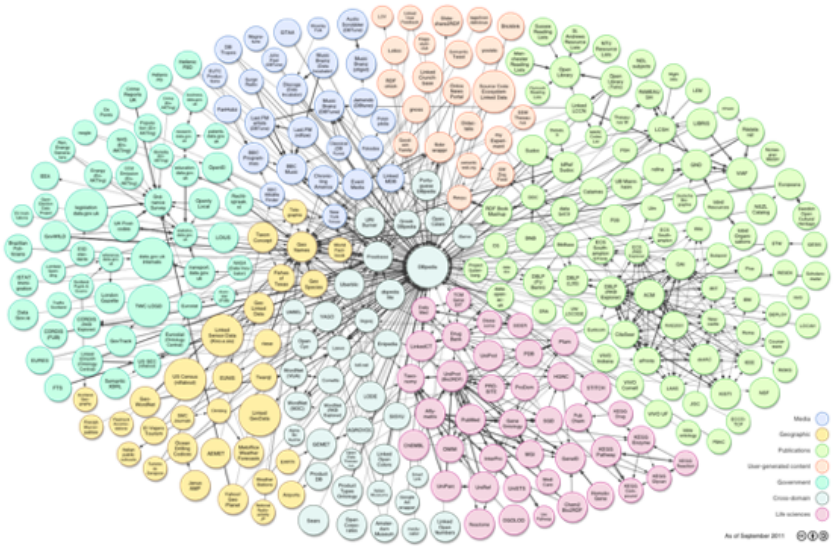


Fraunhofer FOKUS

108

Courtesy of Richard Cyganiak and Anja Jentzsch

LINKED OPEN DATA (LOD) CLOUD DIAGRAM (2014, > 500 DATASETS)



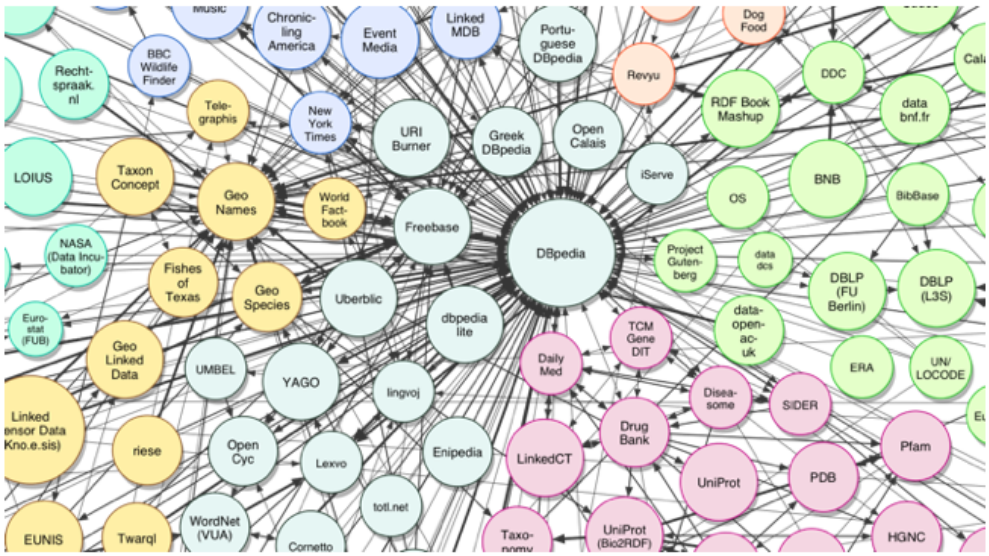
Source:http://lod-cloud.net

Fraunhofer FOKUS

109

- This version is the most used one in other presentations and publications

LINKED OPEN DATA (LOD) CLOUD DIAGRAM (2014, > 500 DATASETS)

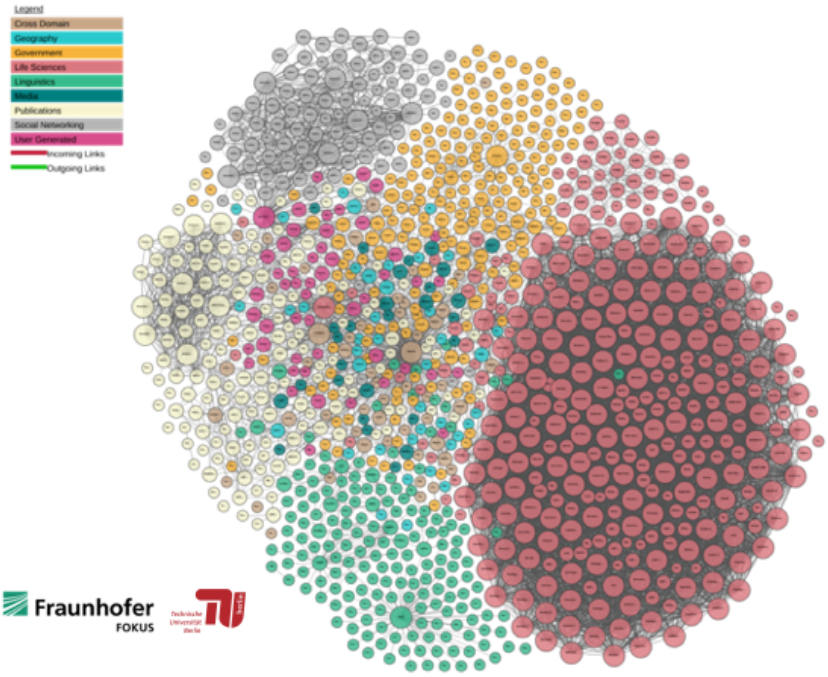


Fraunhofer FOKUS

110

- Dbpedia as a central dataset

SEMANTIC WEB | LINKED OPEN DATASETS (2017, > 1000 DATASETS)



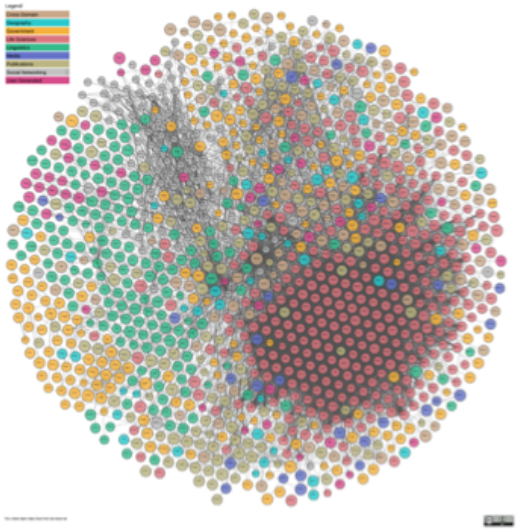
http://lod-cloud.net

Fraunhofer FOKUS

Technische Universität Berlin

111

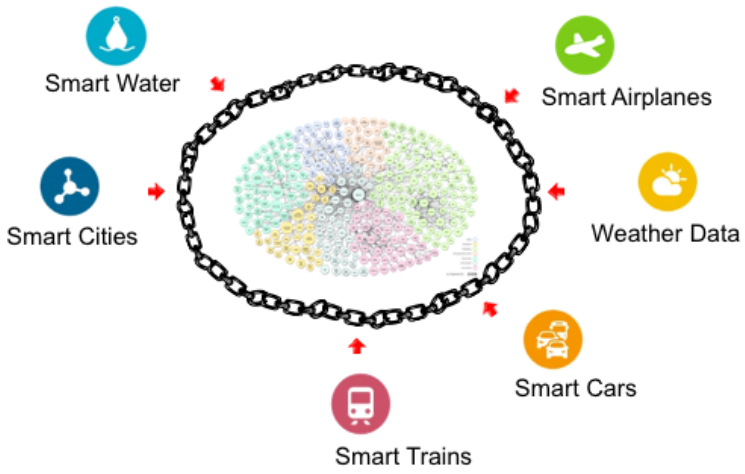
SEMANTIC WEB | LINKED OPEN DATASETS (2018, > 1000 DATASETS)



http://lod-cloud.net

Fraunhofer FOKUS

VISION: CONNECT ALL DATA FROM ALL (IOT) DATA SOURCES



Fraunhofer FOKUS

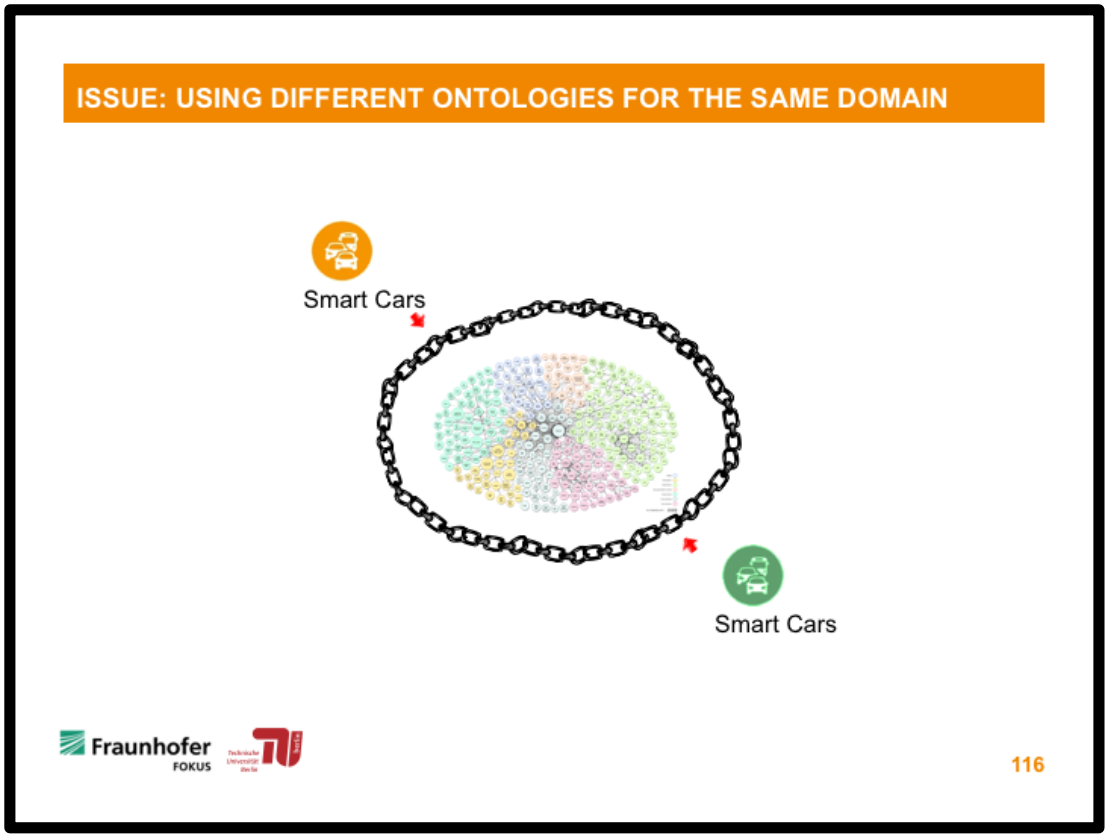
113

LINKED DATA

Questions?

ONTOLOGY MAPPING PROBLEM

Skipped



- Assuming the same concept (e.g. car) is described with two different ontologies

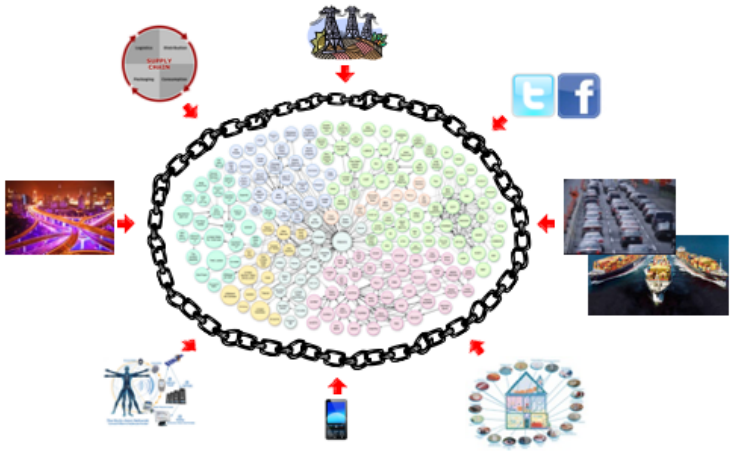
ONTOLOGY MAPPING PROBLEM

Skipped

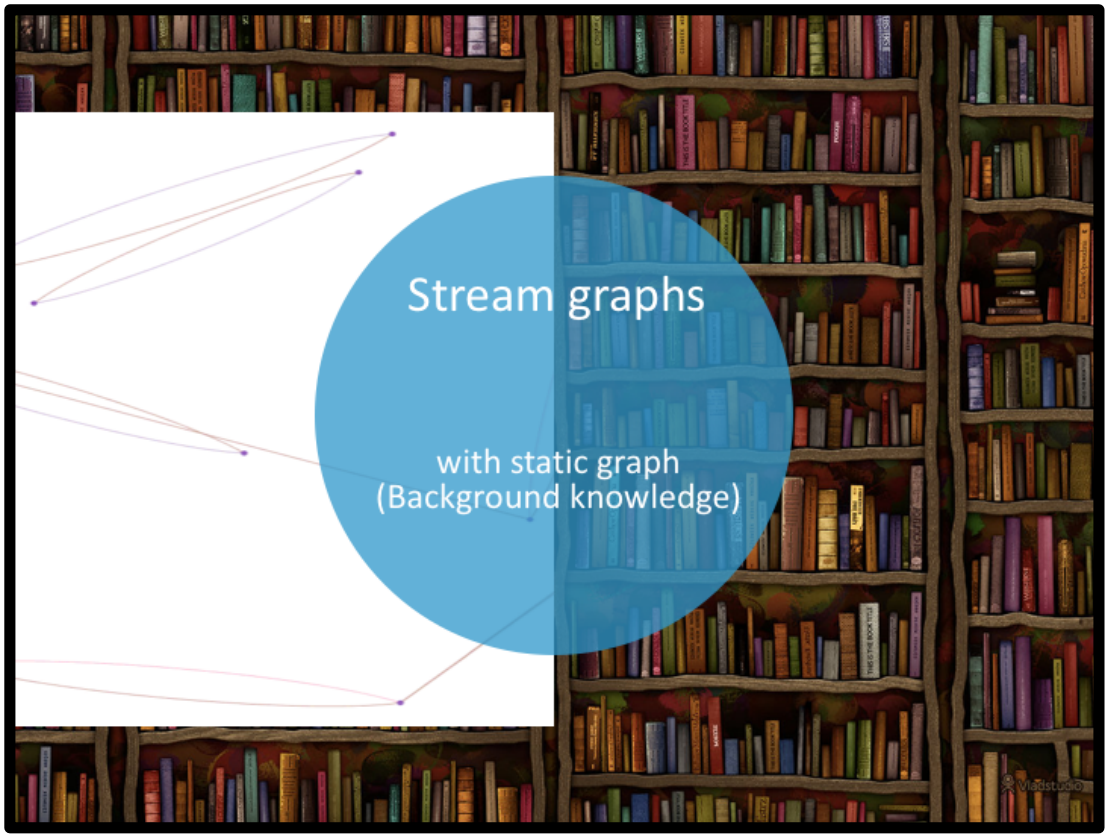
STREAM PROCESSING EXAMPLE

3 minutes

ADDING DYNAMIC DATA TO THE LOD CLOUD

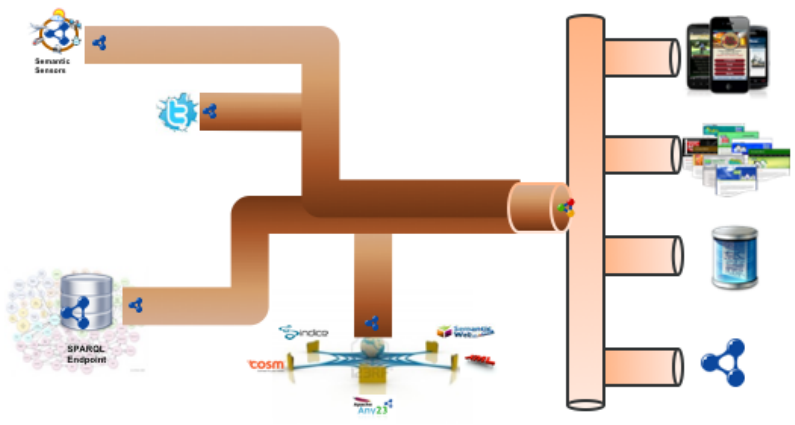


Fraunhofer FOKUS



- Dynamic part
- Static part

LINKED STREAM DATA



Fraunhofer FOKUS

125

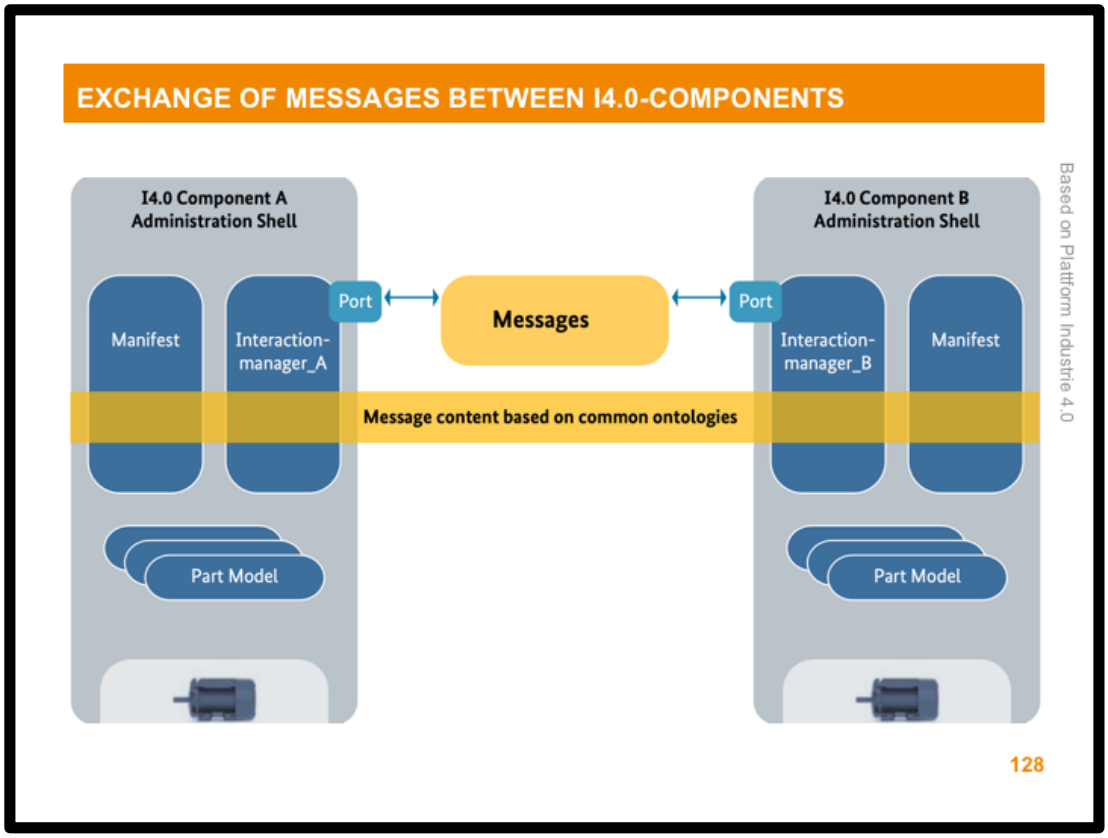
- Dynamic RDF triplets are combined with queries from static triple stores to answer application specific questions

STREAM PROCESSING EXAMPLE

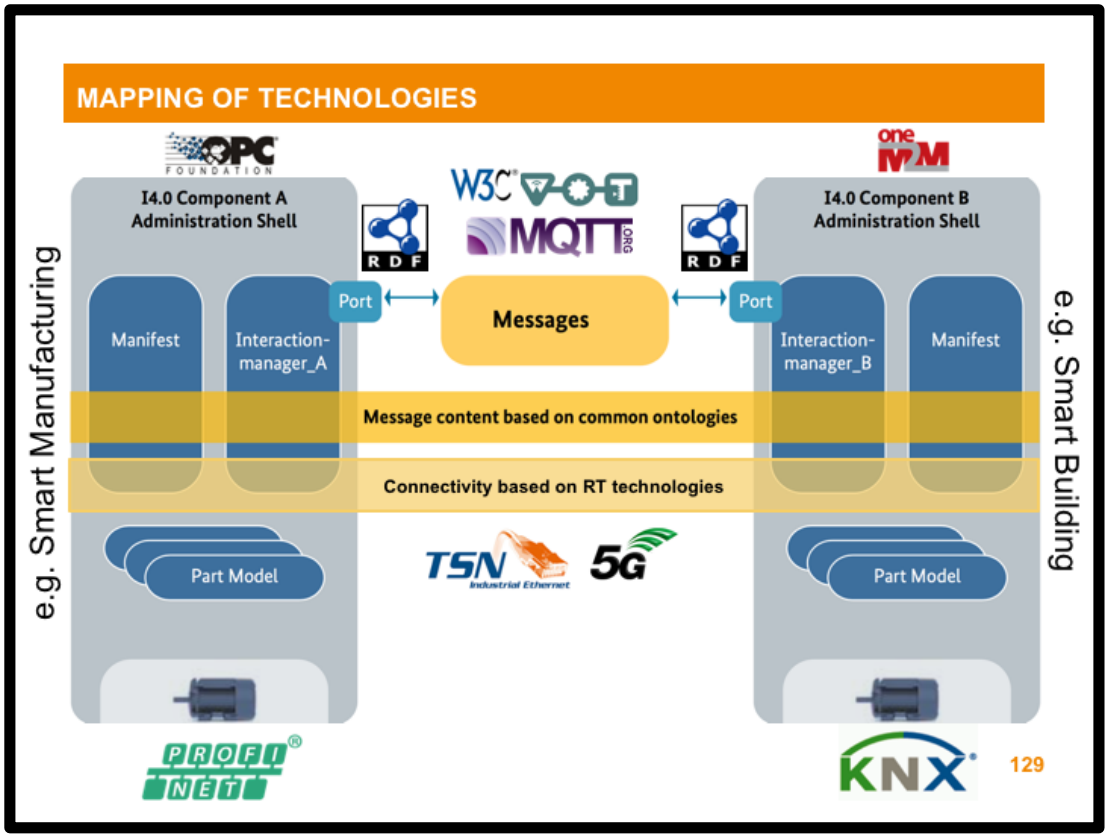
Questions?

OUTLOOK

2 minutes



- Representation by the Plattform Industrie 4.0
- We already talked about I4.0 components and what an administration shell is



- How the technologies we already learned about relate to each other

WHERE IS THE CLOUD IN THIS CONTEXT?