Resource Description Framework (RDF)

UNIVERSITY OF MANNHEIM
Data and Web Science Group

IE650 Knowledge Graphs

Overview

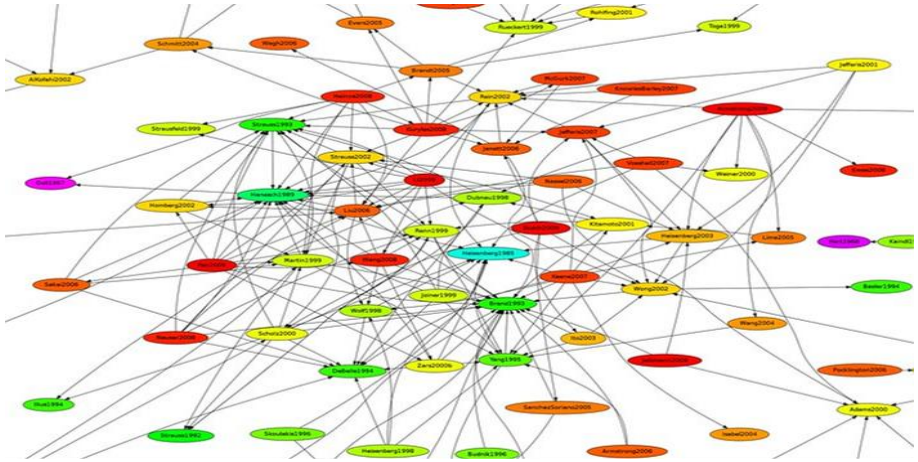
UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Graph representation in a nutshell
- A brief history of RDF
- Encodings of RDF
- Semantics and principles of RDF
- Embedding RDF in HTML – RDFa, Microdata, Microformats
- RDF Tools

Knowledge Graphs are Graphs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

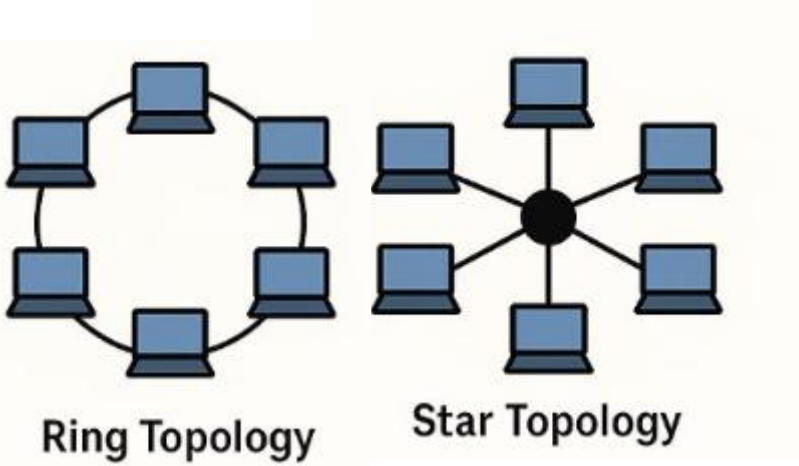
- Graphs may have different flavours
- Directed vs. undirected, e.g.:
  - Social graph: A and B know each other
  - Citation graph: paper X cites paper Y



Knowledge Graphs are Graphs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

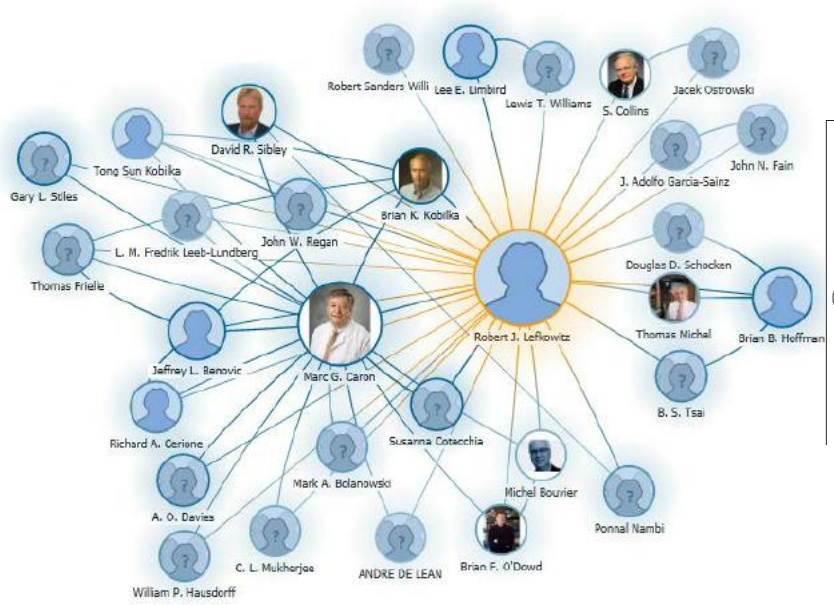
- Graphs may have different flavours
- Labeled vs. unlabeled
  - A network of highways: a particular highway links to cities
  - A network topology

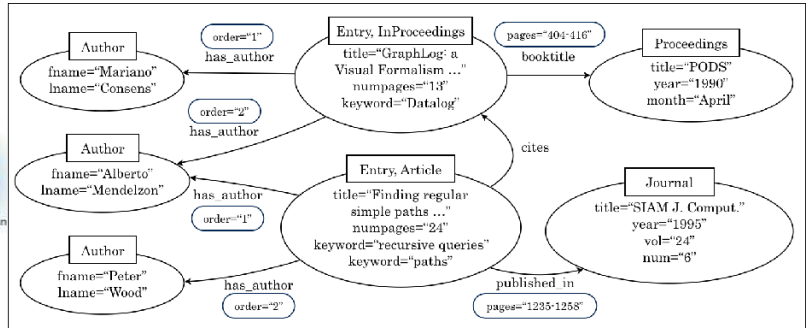


Knowledge Graphs are Graphs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Graphs may have different flavours
- Homogenous vs. heterogeneous nodes, e.g.,
  - A coauthorship graph
  - A graph of authors and publications

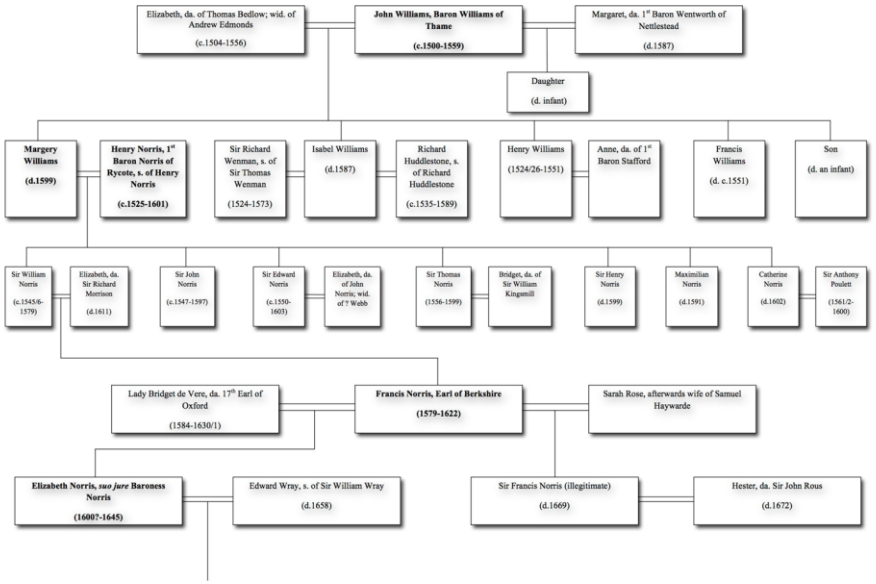


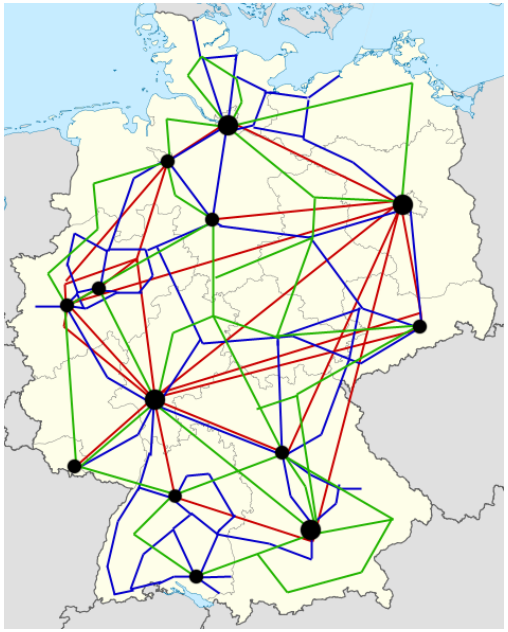


Knowledge Graphs are Graphs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Graphs may have different flavours
- Cyclic vs. acyclic
  - A family tree (acyclic graphs are often referred to as “trees”)
  - A computer network





Knowledge Graphs are Graphs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

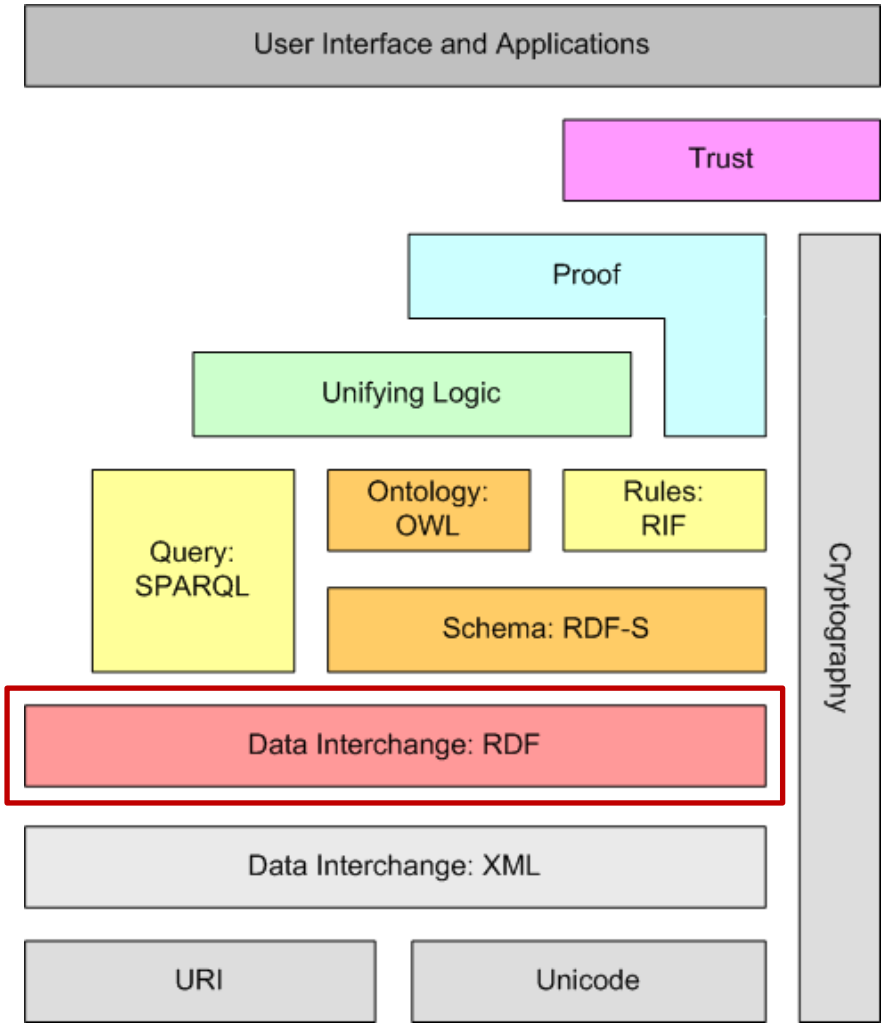
- So when we talk about a knowledge graph, what kind of graph do we mean?
- No formal definition, but de facto consensus:
  - Directed, labeled graph
  - Heterogeneous node types (and edges)
  - Need not be cycle free
- Node types (“classes”) and edge types (“properties”)
  - Are also referred to the “schema” of the graph (aka “ontology”)
  - Can be defined with further restrictions
    - e.g., an edge of type “author” links a publication to a person

Semantic Web – Architecture

here be dragons...

Knowledge Graph Technologies
(This lecture)

Technical
Foundations



Berners-Lee (2009): *Semantic Web and Linked Data* http://www.w3.org/2009/Talks/0120-campus-party-tbl/

History: Metadata on the Web

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Goal: more effective rating and ranking of web contents, e.g., by search engines

- Who has created this page?
- When has it been changed the last time?
- What is its topic?
- Which is the content's license?
- How does it relate to other pages?

Metadata on the Web: Dublin Core

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Developed in 1995 at a workshop in Dublin, Ohio
- 15 predefined tags
- A widely accepted standard (ISO 15836:2009)

- May be embedded into HTML:

```
<html>
  <head profile="http://dublincore.org/documents/2008/08/04/dc-html/">
    <title>Knowledge Graphs</title>
    <link rel="schema.DC"      href="http://purl.org/dc/elements/1.1/" >
    <meta name="DC.publisher"  content="University of Mannheim" />
    <meta name="DC.subject"    content="Semantic Web Technologies" />
    <meta name="DC.creator"    content="Sven Hertling" />
    <meta name="DC.relation"   content="http://www.w3.org/2001/sw/" />
    ...
  </head>
  <body>
  ...
  </body>
</html>
```

Metadata on the Web: Dublin Core

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Identifier
- Format
- Type
- Language
- Date

- Title
- Subject
- Coverage
- Description

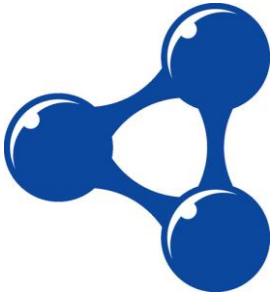
- Creator
- Publisher
- Contributor
- Rights

- Source
- Relation

What is RDF?

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- „Resource Description Framework“
- A W3C standard since 2004
- Description of arbitrary things



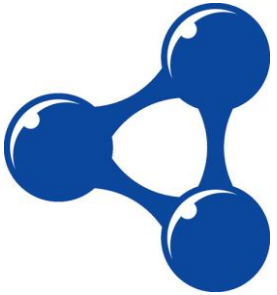
- View 1: Sentences in the form <subject, predicate, object>
- „Alice studies IE650 Knowledge Graphs.“

Some literature:
<head, relation, tail>

- View 2: Directed graphs with labeled edges

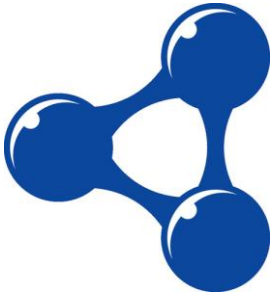
What is RDF?

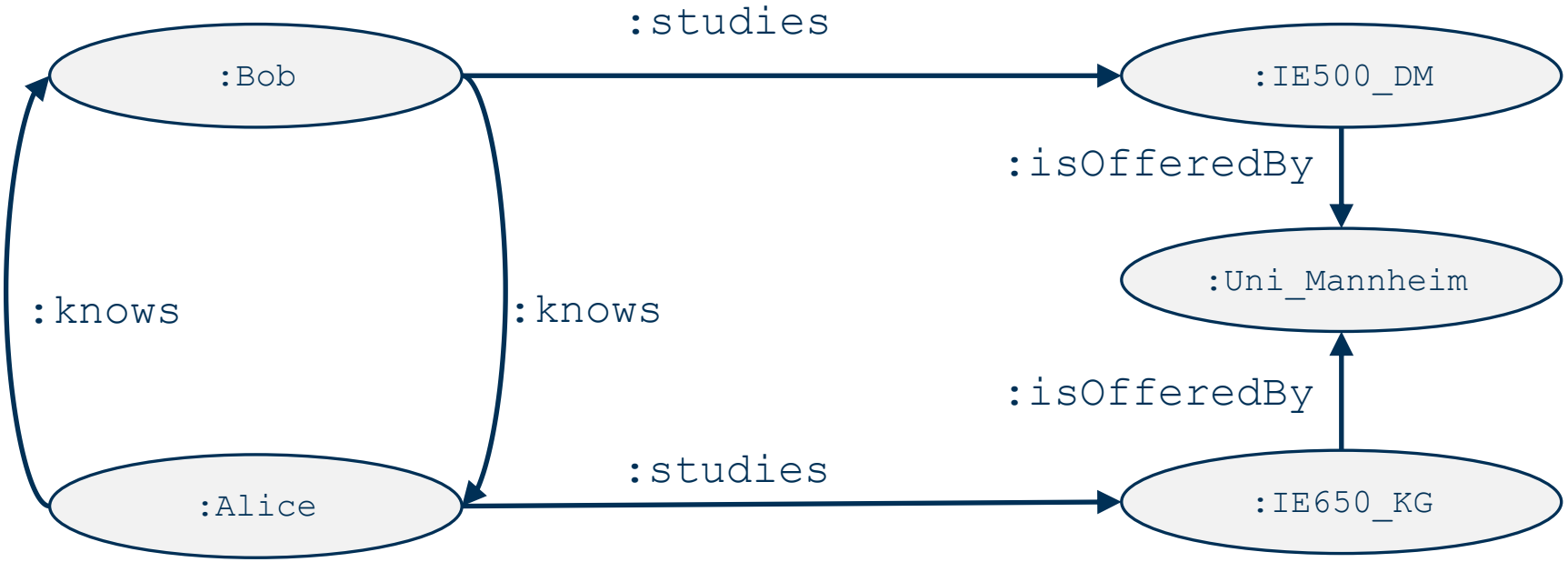
- A knowledge graph consists of *multiple* sentences



What is RDF?

- We usually think of knowledge graphs as densely connected graphs
  - Objects of one statement become subjects of another





Basic Building Blocks of RDF

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Resources
  - denote things
  - are identified by a URI
  - can have one or multiple types
- Literals
  - are values like strings or integers
  - can only be objects, not subjects or predicates (graph view: they can only have ingoing edges)
  - can have a datatype or a language tag (but not both)
- Properties (Predicates)
  - Link resources to other resources and to literals

Basic Building Blocks of RDF

- Resources
  - Alice, IE650_KG
- Properties
  - studies, code
- Literals
  - “650”

Types

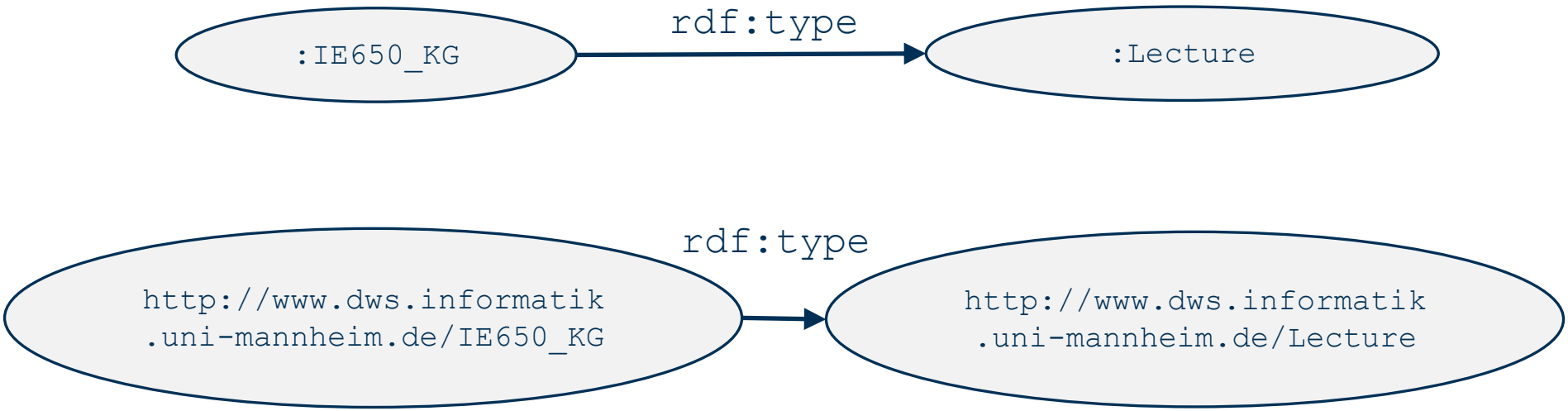
- All resources (not literals) can have a type
- Types can be arbitrarily defined
- The predefined predicate `rdf:type` defines the type of a resource

- Knowledge Graphs is a lecture

<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>

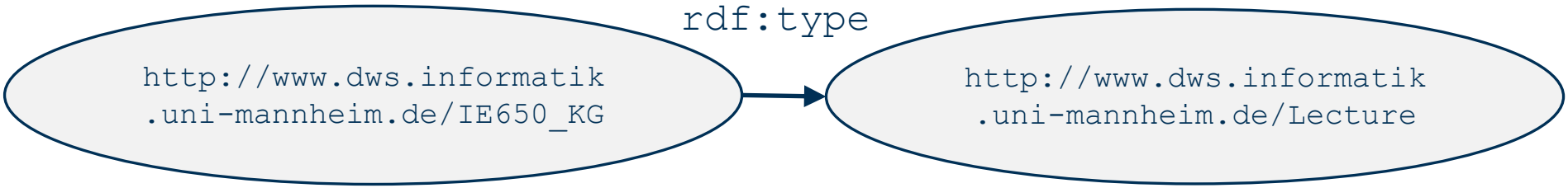
URIs

- Each ":" represents a URI prefix
  - E.g. `http://www.dws.informatik.uni-mannheim.de/`



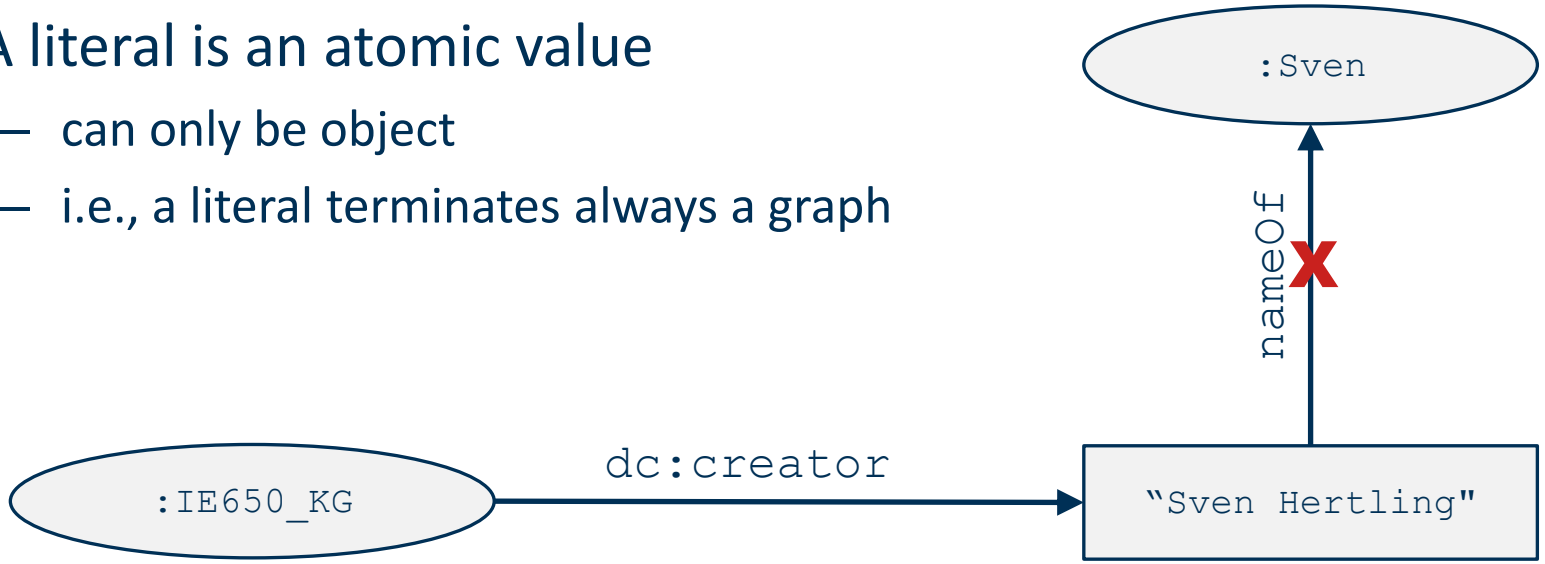
Vocabulary

- There are already standardized vocabularies
  - One is `rdf:`
    - This prefix resolves to http://www.w3.org/1999/02/22-rdf-syntax-ns#
  - Thus `rdf:type` is
    - http://www.w3.org/1999/02/22-rdf-syntax-ns#type



Resource vs. Literal

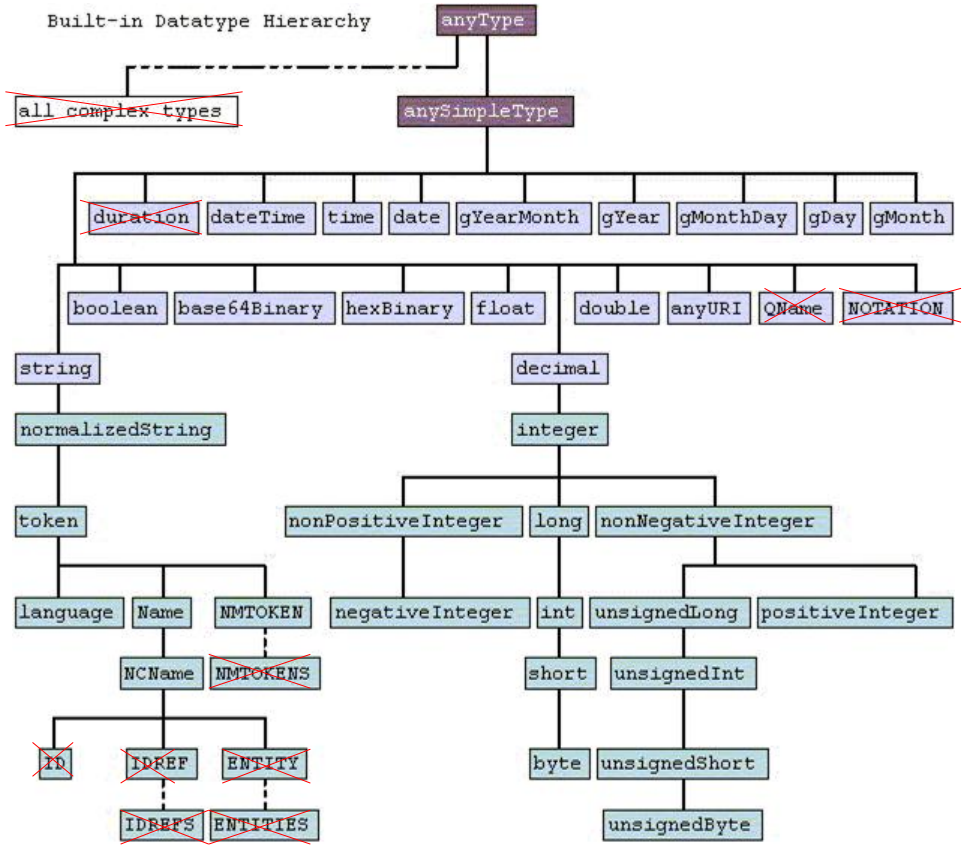
- A literal is an atomic value
  - can only be object
  - i.e., a literal terminates always a graph



Datatypes for Literals

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- (Almost) all XML Schema datatypes may be used
- Exception:
  - XML specific types
  - The underspecified type "duration"
  - sequence types



XML Schema Part 2: Datatypes Second Edition http://www.w3.org/TR/xmlschema-2/

Language Tags for Literals

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Literals may be defined in different natural languages
  - `"München"@de`
  - `"Munich"@en`
- Those can be marked
- Note: Knowledge Graphs can be multilingual!

- Language codes according to ISO 963
  - ISO 963-1 (1963): two-digit codes, 136 languages
  - ISO 963-2 (1998): three-digit Codes, 464 languages
  - if both are defined, ISO 963-1 has to be used!

ISO International Organization for Standardization

- http://www.loc.gov/standards/iso639-2/php/English_list.php

Language Tags for Literals

UNIVERSITY OF MANNHEIM
Data and Web Science Group

Library of Congress >> Standards

ISO 639.2 Language Coding Agency

ISO 639-1 Language Coding Agency Home - ISO 639-3 Language Coding Agency Home
ISO 639-5 Language Coding Agency Home
Other Standards Maintained by the Library - Library of Congress Home

Codes for the Representation of Names of Languages

Codes arranged alphabetically by alpha-3/ISO 639-2 Code

Note: ISO 639-2 is the alpha-3 code in *Codes for the representation of names of languages-- Part 2*. There are 20 languages that have alternative codes for bibliographic or terminology purposes. In those cases, each is listed separately and they are designated as "B" (bibliographic) or "T" (terminology). In all other cases there is only one ISO 639-2 code. Multiple codes assigned to the same language are to be considered synonyms. ISO 639-1 is the alpha-2 code.

ISO 639-2 Code	ISO 639-1 Code	English name of Language	French name of Language	German name of Language
aar	aa	Afar	afar	Danakil-Sprache
abk	ab	Abkhazian	abkhaze	Abchasisch
ace		Achinese	aceh	Aceh-Sprache
ach		Acoli	acoli	Acholi-Sprache
ada		Adangme	adangme	Adangme-Sprache
ady		Adyghe; Adygei	adyghé	Adygisch
afa		Afro-Asiatic languages	afro-asiatiques, langues	Hamitosemitische Sprachen (Andere)
afh		Afrihili	afrihili	Afrihili
afr	af	Afrikaans	afrikaans	Afrikaans
ain		Ainu	aïnou	Ainu-Sprache
aka	ak	Akan	akan	Akan-Sprache
akk		Akkadian	akkadien	Akkadisch
alb (B) sqi (T)	sq	Albanian	albanais	Albanisch
ale		Aleut	aléoute	Aleutisch
alg		Algonquian languages	algonquines, langues	Algonkin-Sprachen (Andere)
alt		Southern Altai	altai du Sud	Altaisch
amh	am	Amharic	amharique	Amharisch

https://www.loc.gov/standards/iso639-2/php/code_list.php

Datatypes in RDF

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Examples:

```
:Munich :hasName "München"@de .
:Munich :hasName "Munich"@en .
:Munich :hasPopulation "1356594"^^xsd:integer .
:Munich :hasFoundingYear "1158-01-01"^^xsd:date .
```

Literals can have either a datatype or a language tag – not both

- Note: there are no default datatypes (not even “string”!)

- These are three different literals:
  - `"München"`
  - `"München"@de`
  - `"München"^^xsd:string`

Triple Notation

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- A W3C standard (2004)
- Triples consist of a subject, predicate, and object
- An RDF document is an *unordered* set of triples

- Simple triple:

```
<http://www.dws.informatik.uni-mannheim.de/teaching/IE650_KG>
<http://purl.org/dc/elements/1.1/relation>
<http://www.w3.org/2001/sw/> .
```

- Literal with language tag:

```
<http://www.dws.informatik.uni-mannheim.de/teaching/IE650_KG>
<http://purl.org/dc/elements/1.1/subject>
“Knowledge Graphs"@en .
```

- Type literal:

```
<http://www.dws.informatik.uni-mannheim.de/teaching/IE650_KG>
<http://www.uni-mannheim.de/mhb/creditpoints>
"6"^^<http://www.w3.org/2001/XMLSchema#integer> .
```

Turtle Notation

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- A simplified triple notation
- Central definition of namespaces (and a default namespace):

```
@prefix dc: <http://purl.org/dc/elements/1.1/>
@prefix : <http://www.dws.informatik.uni-mannheim.de/teaching/>
:IE650_KG dc:subject "Knowledge Graphs"@en .
```

- Triples sharing the same subject or subject+predicate:

```
:IE650_KG dc:subject "Wissensgraphen"@de ,
                     "Knowledge Graphs"@en ;
          dc:creator "Sven Hertling".
```

- Shorthand notation for `rdf:type`:

```
:IE650_KG a :lecture .
```

Notation RDF/XML

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- A W3C standard since 2004
- Encodes RDF in XML
- Suitable for machine processing (plenty of XML tools!)

- Defining resources

```
<rdf:Description rdf:about="http://dws.informatik.uni-mannheim.de/teaching/IE650_KG">
  <dc:creator>Sven Hertling</dc:creator>
</rdf:Description>
```

- Defining typed resources

```
<rdf:Description rdf:about="http://dws.informatik.uni-mannheim.de/teaching/IE650_KG">
  <rdf:type rdf:resource="http://www.uni-mannheim.de/mhb/Lecture"/>
</rdf:Description>
```

- Alternative representation

```
<mhb:Lecture rdf:about="http://dws.informatik.uni-mannheim.de/teaching/IE650_KG"
 xmlns:mhb="http://www.uni-mannheim.de/mhb/" />
```

Notation RDF/XML

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Relations between resources by nesting

```
<mhb:Lecture rdf:about="http://dws.informatik.uni-mannheim.de/teaching/IE650_KG">
  <mhb:givenBy>
    <mhb:Lecturer rdf:about="http://dws.informatik.uni-mannheim.de/sven"/>
  </mhb:givenBy>
</mhb:Lecture>
```

- Relations between resources by explicit links

```
<mhb:Lecturer rdf:about="http://dws.informatik.uni-mannheim.de/sven" />

<mhb:Lecture rdf:about="http://dws.informatik.uni-mannheim.de/teaching/IE650_KG">
  <mhb:givenBy rdf:resource="http://dws.informatik.uni-mannheim.de/sven"/>
</mhb:Lecture>
```

JSON-LD Notation

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- JSON is popular in script programming
- JSON-LD: Standard for serializing RDF in JSON
  - JSON-LD 1.1 W3C standard (2020)

```
{
    "@id": "http://www.svenhertling.com/",
    "http://dws.informatik.uni-mannheim.de/name": "Sven Hertling",
    "http://dws.informatik.uni-mannheim.de/teaches":
    {
        "http://http://www.w3.org/1999/02/22-rdf-syntax-ns#type":
            "http://dws.informatik.uni-mannheim.de/Lecture",
        "http://dws.informatik.uni-mannheim.de/title":
            "Knowledge Graphs"
    }
}
```

RDF Lists

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- In case you want to model a list of items
  - E.g. recipe contains milk, eggs, flour

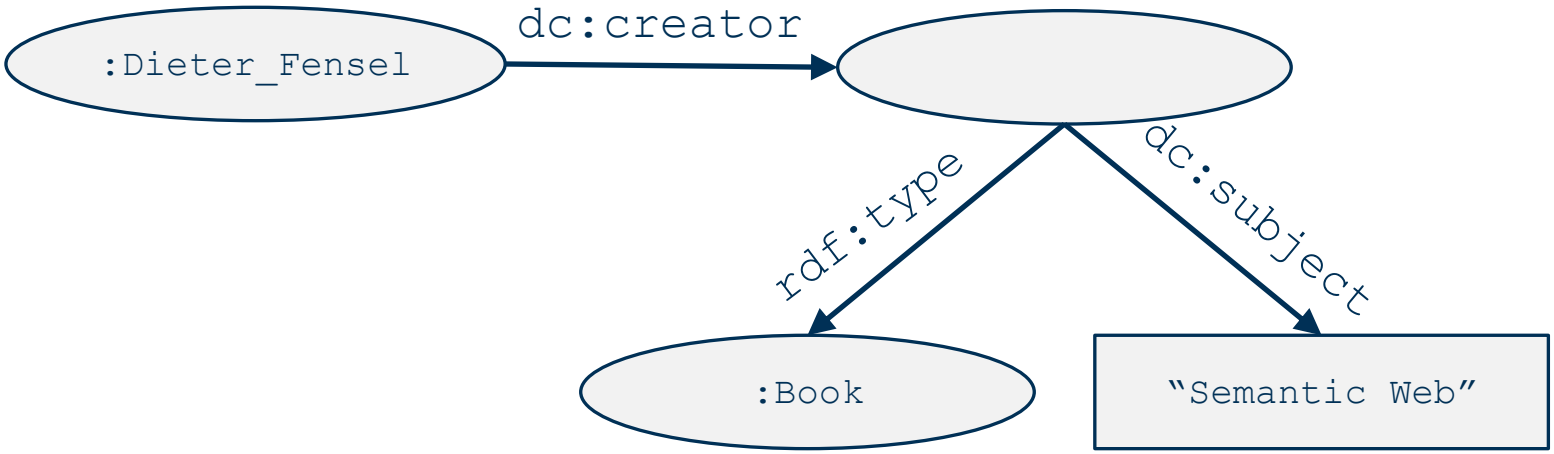
```
:recipe :ingredients :list0.
:list0 rdf:type rdf:List
:list0 rdf:first :milk.
:list0 rdf:rest :list1.

:list1 rdf:first :eggs.
:list1 rdf:rest :list2.

:list2 rdf:first :flour.
:list2 rdf:rest rdf:nil .
```

```
:recipe :ingredients ( :milk :eggs :flour).
```

Shorthand notation in turtle

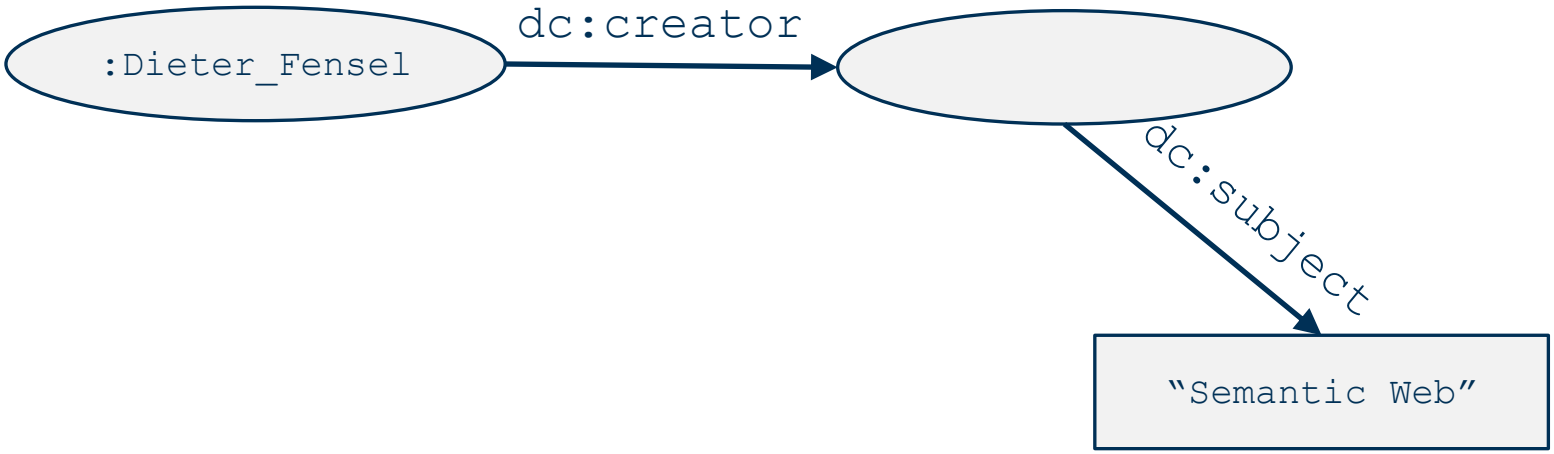
Blank Nodes

- Information that is not or cannot be specified
  - "Dieter Fensel has written a book about the Semantic Web"



Blank Nodes

- Information that is not or cannot be specified
  - "Dieter Fensel has written something about the Semantic Web."



Blank Nodes in Turtle

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Variant 1: explicitly named with an underscore

```
:Dieter_Fensel dc:creator _:x .
_:x a :Book ;
    dc:subject "Semantic Web" .
```

- Variant 2: unnamed with square brackets

```
:Dieter_Fensel dc:creator [
    a :Book;
    dc:subject "Semantic Web“
].
```

- Notes
  - both are equivalent
  - changing blank node names does not change the semantics!

Application of Blank Nodes: n-ary Predicates

UNIVERSITY OF MANNHEIM
Data and Web Science Group

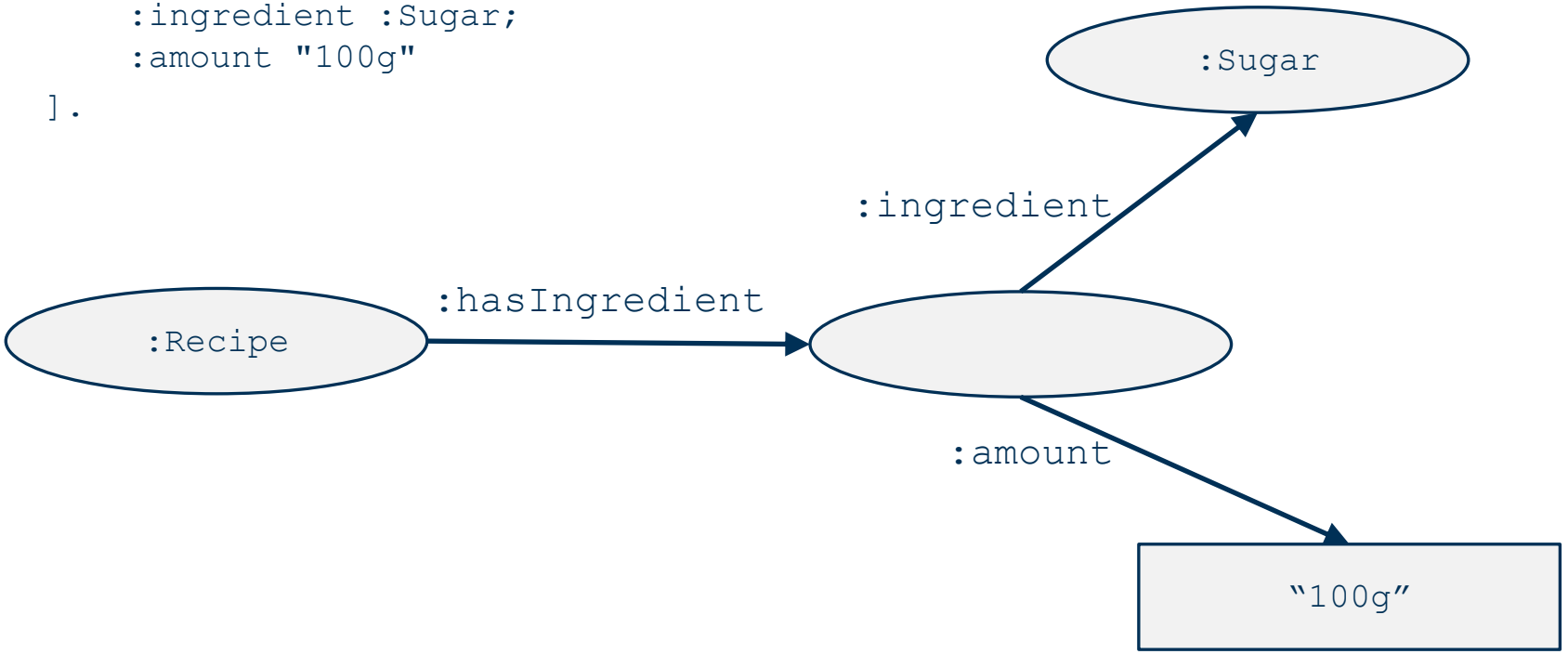
- RDF predicates always connect a subject and an object
  - i.e., in the sense of predicate logic, they are binary predicates

```
:Sven :works_for :UniMannheim .
```
⇔ works_for(Sven, UniMannheim) .

- Sometimes, n-ary predicates are needed
  - has_ingredient(Recipe, Sugar, 100g)

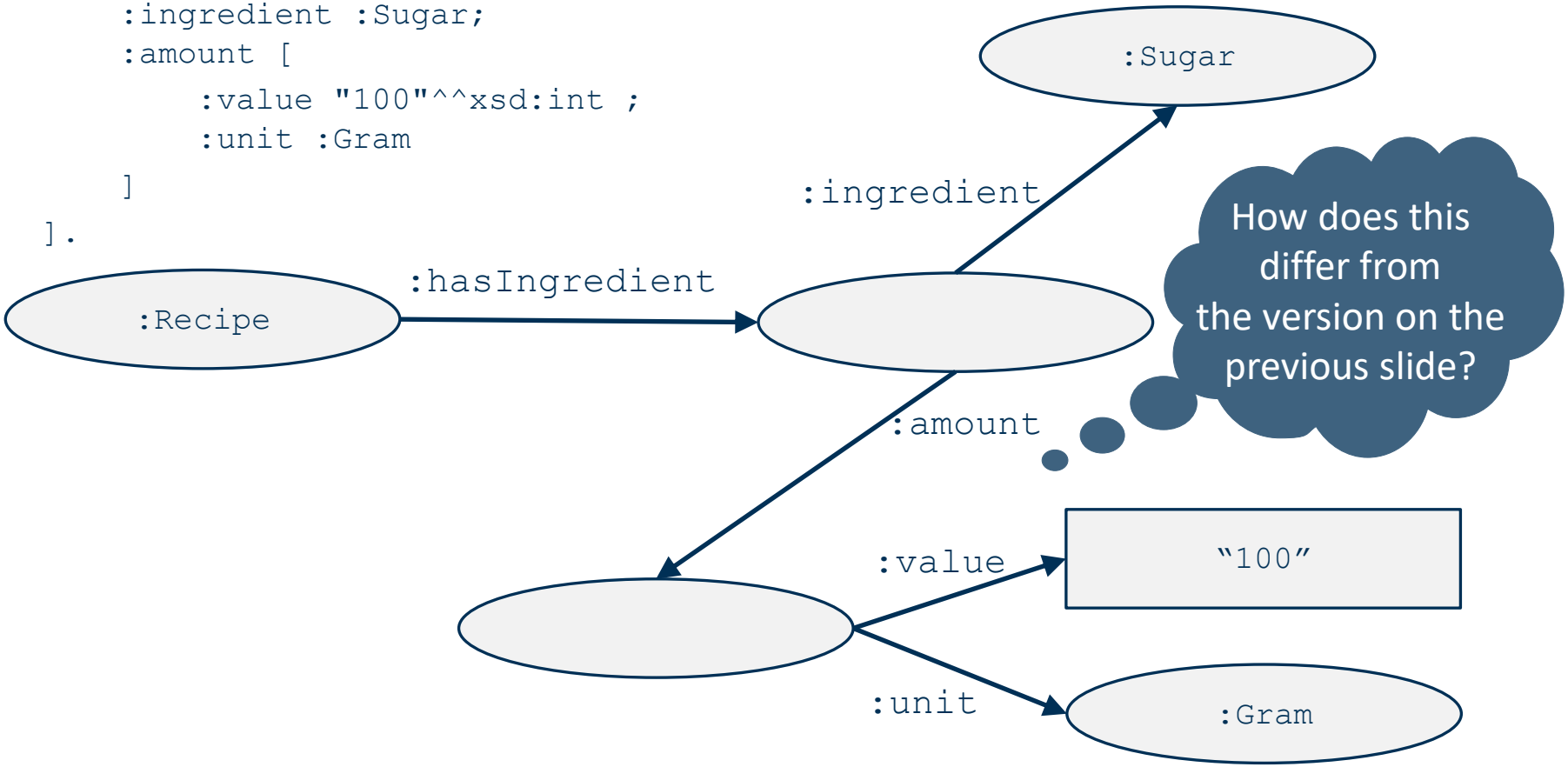
Application of Blank Nodes: n-ary Predicates

```
:Recipe :hasIngredient [
    :ingredient :Sugar;
    :amount "100g"
].
```



Application of Blank Nodes: n-ary Predicates

```
:Recipe :hasIngredient [
    :ingredient :Sugar;
    :amount [
        :value "100"^^xsd:int ;
        :unit :Gram
    ]
].
```



Break

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- How to model the following sentences
  - Mannheim is a city and has a population of 319.536
  - Alice studies at the University of Mannheim which is located in Mannheim

Semantic Principles of RDF

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- On the Web, "anybody can say anything about anything"
  - This is called the AAA principle (Allemang & Hendler)

- The AAA principle is also used for RDF knowledge graphs
  - History: information sharing on the Web

Semantic Principles of RDF

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- One thing can have multiple names

```
:Munich :capitalOf :Bavaria .
:München :capitalOf :Bayern .
```

- On the semantic web, there is not just one name for each thing
  - this is called the *Non-unique name assumption*

- This means: Just that two things have different names does not mean that they are different!

RDF: Intuition and Actual Semantics

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Let us consider the following example:

```
:Peter :fatherOf :Julia ,
                 :Mary .
```

- How many children does Peter have?

- Intuitively, we assume that Julia and Mary are two different persons
- However, this is not trivial for a machine
  - (and the assumption may even be wrong)

Semantic Principles of RDF

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Historical argumentation:
  - We (probably) do not know all the contents of the Semantic Web*
- Therefore, there may be more information on a resource than what we have

- This principle is called "Open World Assumption"

* this may not be true for enterprise knowledge graphs

RDF: Intuition and Actual Semantics

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Let us consider this example again:

```
:Peter :fatherOf :Julia ,
                 :Mary .
```

- How many children does Peter have?

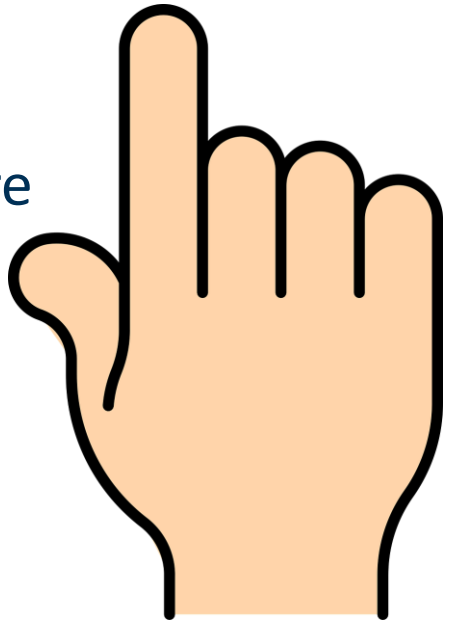
- Intuition says: two children
- However, he could also have three or more
  (or also just one, as we have learned just a minute ago)

RDF: Intuition and Actual Semantics

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Both
  - Non-unique Name Assumption and
  - Open World Assumption
- will re-occur quite a few times in this lecture

- Hint: consider those two whenever something seems weird when interpreting RDF data



RDF and HTML

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- The Semantic Web uses RDF
- The “classic” Web uses HTML
- Does that mean that each information has to be encoded twice?
  - once for humans, once for machines

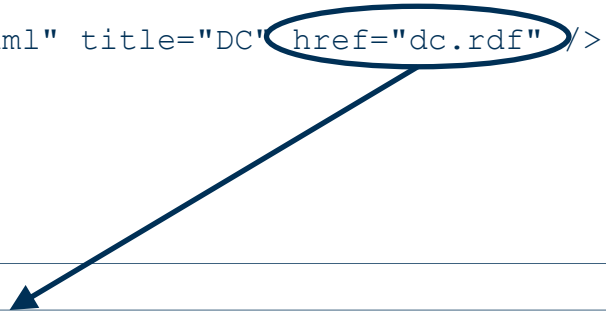
```
<html>
  …
  <b>Dr. Mark Smith</b>
  <i>Physician</i>
  Main St. 14
  Smalltown
  Mon-Fri 9-11 am
  Wed 3-6 pm
  …
</html>
```

```
:p a :Physician .
:p :hasDegree "Dr." .
:p :hasName "Mark Smith" .
:p :hasAddress :a .
:a :street "Main Street" .
:a :number "14"^^xsd:int .
:a :city "Smalltown" .
:p :hasOpeningHours [
  a rdf:Bag ;
  [ :day :Monday;
    :from "9"^^xsd:int;
    :to "11"^^xsd:int;
  ]
  ...
```

Using RDF and HTML Together – Variant 1

- Explicit reference to a RDF version
  - an agent stumbling on the HTML page can download the RDF data file

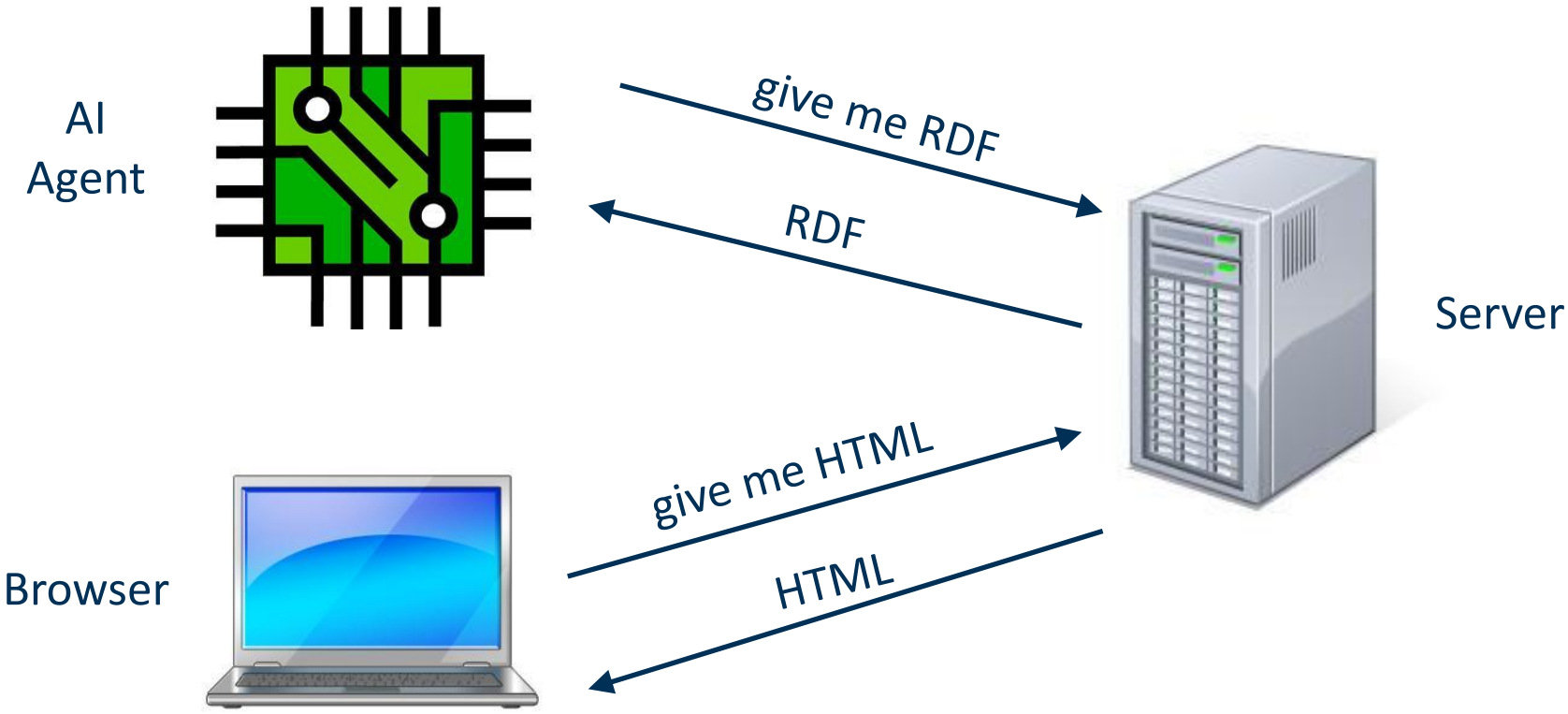
```
<html>
  <head>
    <link rel="meta" type="application/rdf+xml" title="DC" href="dc.rdf" />
  </head>
  <body>
  ...
  </body>
</html>
```



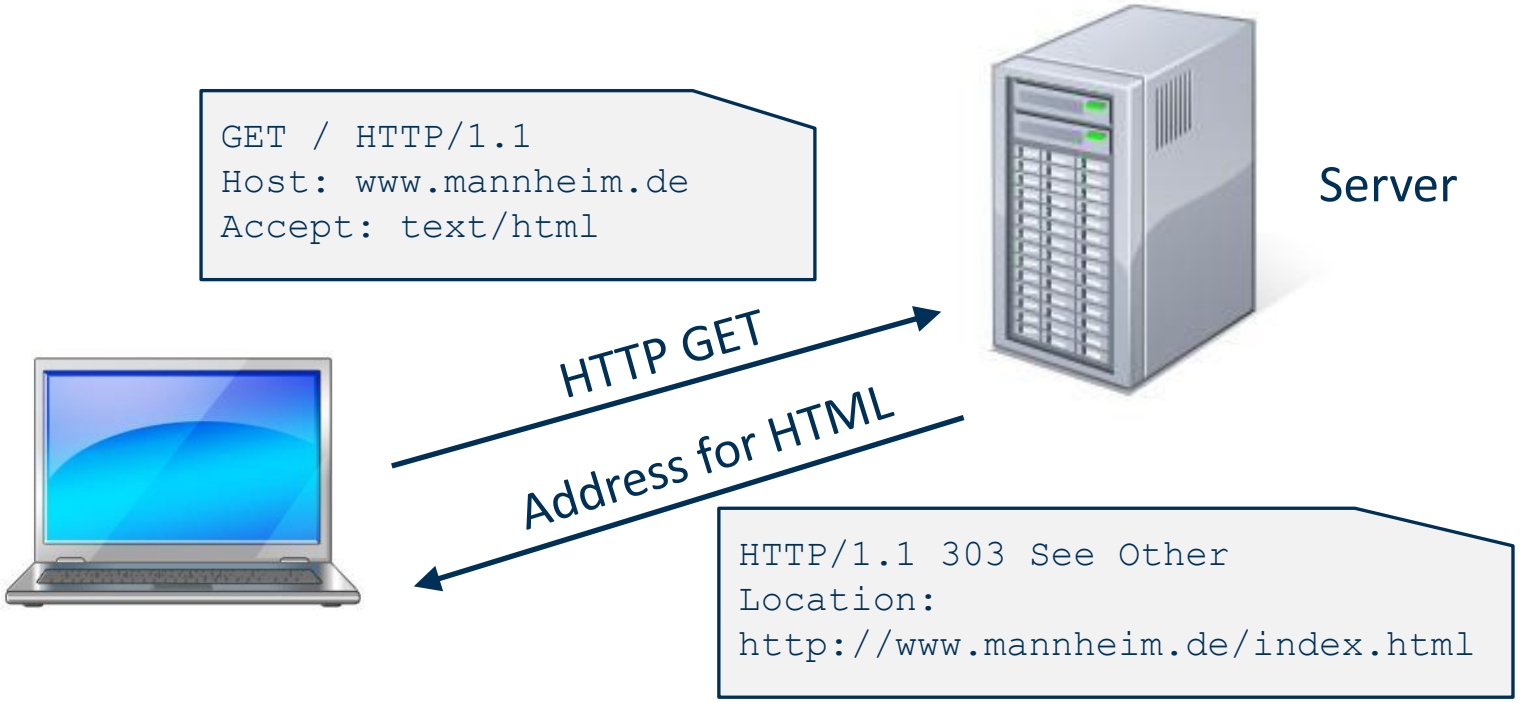
```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://www.dws.informatik.uni-mannheim.de/mhb/IE650_KG">
    <dc:publisher>University of Mannheim</dc:publisher>
    <dc:subject>Knowledge Graphs</dc:subject>
    <dc:creator>Sven Hertling</dc:creator>
    <dc:relation rdf:resource="http://www.w3.org/2001/sw/" />
  </rdf:Description>
</rdf:RDF>
```

Using RDF and HTML Together – Variant 2

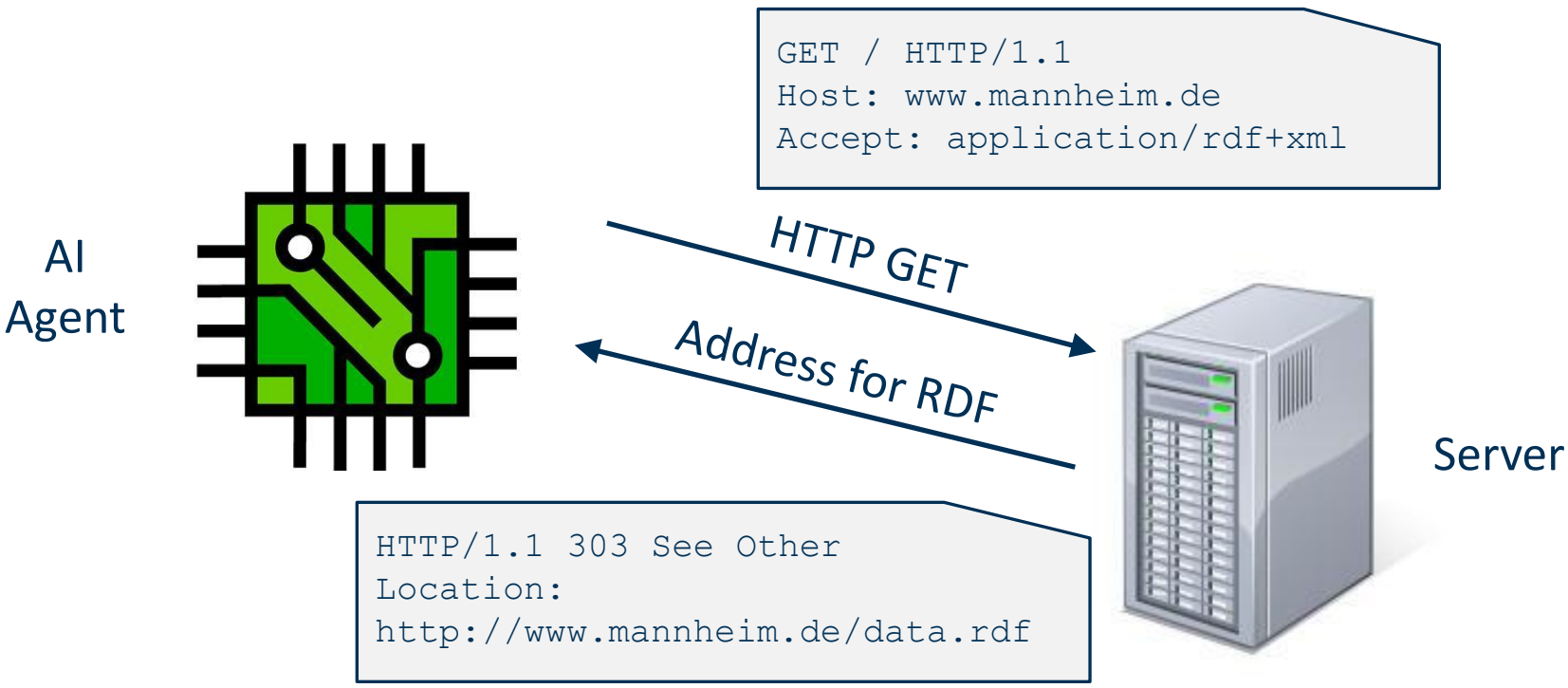
- Content Negotiation



Content Negotiation in Detail



Content Negotiation in Detail



LLMs are starving for KGs

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- **List of crawlers visiting KGs from eurecom (1 month):**
  - OpenAI(OpenAI-Search, ChatGPT-User, **GPTBot**)
  - Bytedance(Bytespider)
  - Apple (Applebot)
  - Meta AI(MetaExternal)
  - Anthropic (**ClaudeBot**)
  - Microsoft (BingBot, LinkedInBot)
  - DuckDuckGo (DuckAssist)
  - CommonCrawl(CCBot)
  - Amazon (**Amazonbot**)
  - Perplexity (Perplexity-User, Perplexitybot)

data.doremus.org: 3,430,585 requests
memad.eurecom.fr: 1,831,538 requests
data.cimple.eu: 1,280,429 requests
explorer.cimple.eu: 676,140 requests
vocab.odeuropa.eu: 303,161 requests
asrael.eurecom.fr: 241,378 requests
data.silknow.org: 103,334 requests
skosmos.silknow.org: 68,660 requests
ada.silknow.org: 12,402 requests
data.odeuropa.eu: 9,688 requests
std.eurecom.fr: 217 requests
data.edf.eurecom.fr: 117 requests
yang.eurecom.fr: 48 requests
sirene.eurecom.fr: 17 requests

https://www.eurecom.fr/en/publication/8246

Content Negotiation: MIME Types

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- MIME: Multipurpose Internet Mail Extensions
- Original purpose: classifying e-mail attachements
  - Text, PDF, ..
- First version: 1996
- Administrated by IANA
- Important MIME types for Semantic Web programming
  - application/rdf+xml
  - text/turtle
  - text/n3
  - application/json
  - application/sparql-query
  - application/sparql-results+xml

Using RDF and HTML Together

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Link to RDF Document
  - Can be done with a simple HTML editor
  - No special server configuration needed
- Content Negotiation
  - Requires particular server setup
  - *One* URI can be used for different representations

- Both cases require
  - two different representations
  - “double bookkeeping”

  → Potential source of inconsistencies!

RDF in Attributes (RDFa)

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Idea of RDFa
  - Why not encode HTML and RDF in *one* document
  - The essential information only has to be encoded once

- RDFa combines XHTML with RDF
  - W3C Standard since 2008

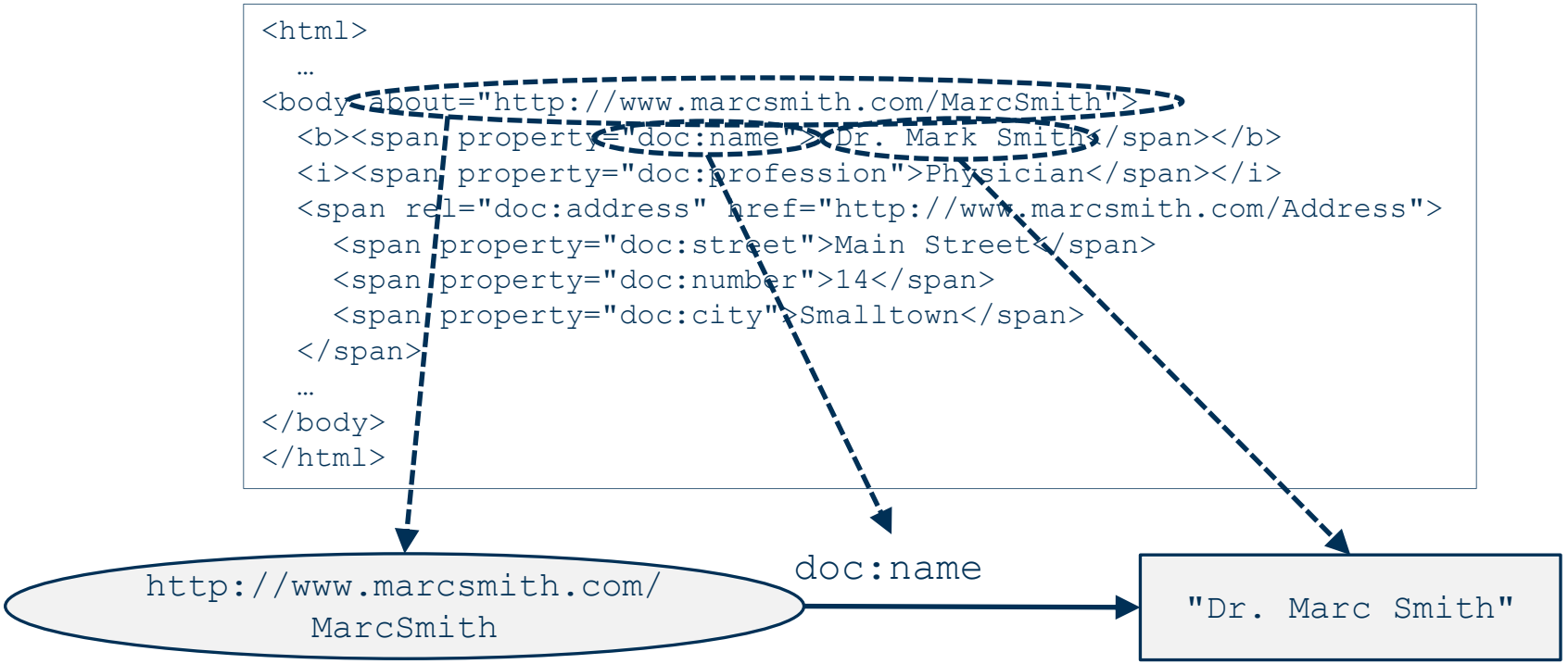
```
<html>
  …
<body about="http://www.marcsmith.com/MarcSmith">
  <b><span property="doc:name">Dr. Mark Smith</span></b>
  <i><span property="doc:profession">Physician</span></i>
  <span rel="doc:address" href="http://www.marcsmith.com/Address">
    <span property="doc:street">Main Street</span>
    <span property="doc:number">14</span>
    <span property="doc:city">Smalltown</span>
  </span>
  …
</body>
</html>
```

RDFa Language Constructs

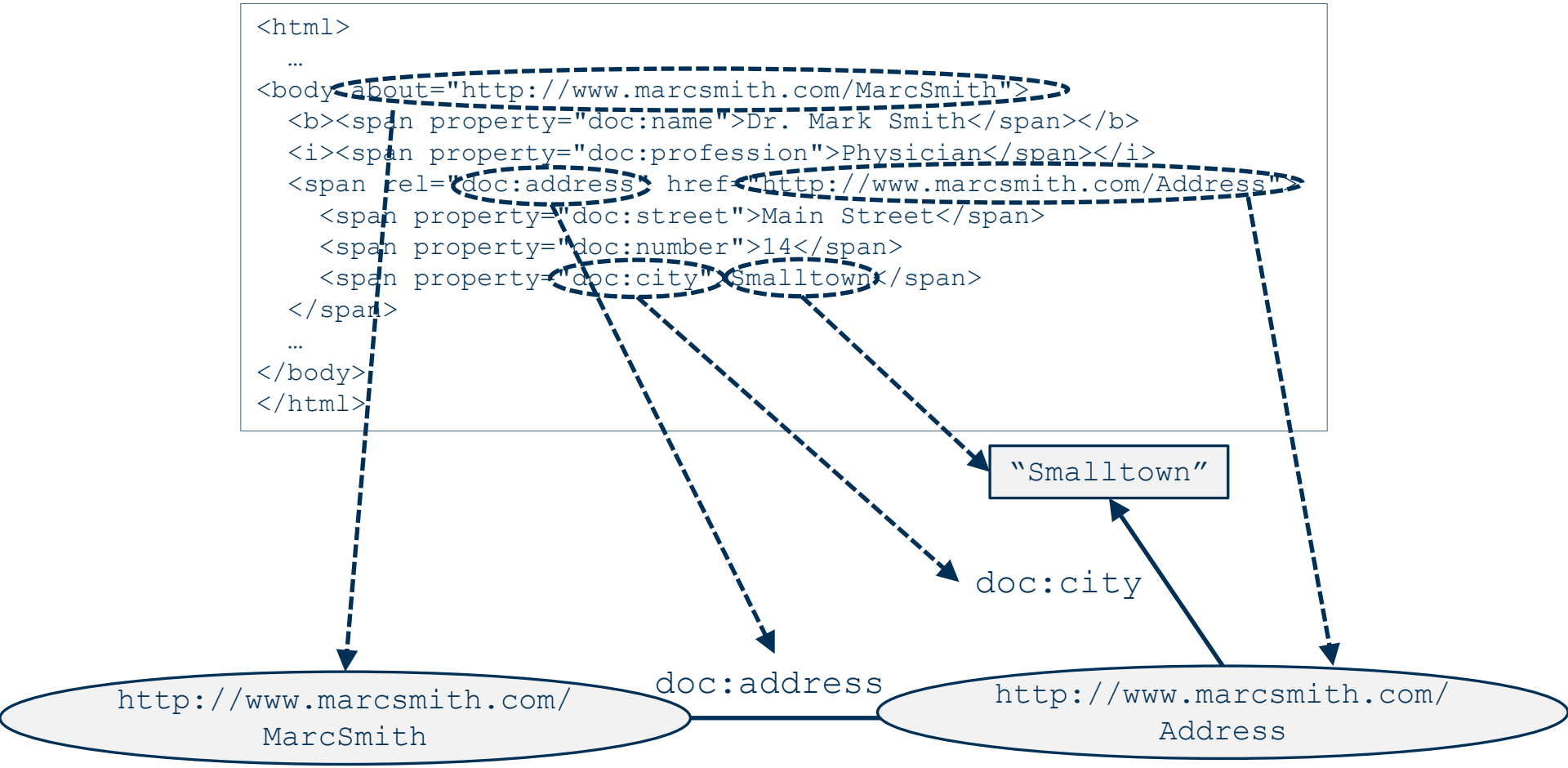
UNIVERSITY OF MANNHEIM
Data and Web Science Group

- `about="http://foo.bar/aSubject"`
  - Defines the subject of a page or section
- `property="http://foo.bar/aProperty"`
  - Defines a relation
  - Contents of the tag are interpreted as a literal
- `rel="http://foo.bar/aRelation"`
  - Defines a relation to another resource
- `href="http://foo.bar/aResource"`
  - Defines a relation's object
  - can be the subject of a resource again
- `typeof="http://foo.bar/aType"`
  - defines a resource's type

RDF in Attributes (RDFa)



RDF in Attributes (RDFa)



Alternative to RDFa: Microdata

- Adding structured information to web pages
  - By marking up contents
  - Arbitrary vocabularies are possible
  - Introduced with HTML5
- Similar to RDFa
  - W3C Standard since 2008, maintained by WHATWG since 2019


```
<div itemscope itemtype="http://schema.org/PostalAddress">
 <span itemprop="name">Data and Web Science Group</span>
 <span itemprop="addressLocality">Mannheim</span>,
 <span itemprop="postalCode">68131</span>
 <span itemprop="addressCountry">Germany</span>
</div>
```

Alternative to RDFa: Microdata

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Commonalities
  - Arbitrary classes/predicates are possible
  - Although Microdata is mainly used with schema.org
- Differences
  - Microdata is slightly less expressive
  - No URIs, only blank nodes
  - No cycles in the resulting RDF graph
  - No reification (see later)

RDFa and Microdata

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- In a specific <script> tag
- Note: unlike RDFa/Microdata, this is content duplication!

```
<script>
  "@context": "https://schema.org",
  "@type": "Organization",
  "address": {
    "@type": "PostalAddress",
    "name": "Data and Web Science Group",
```

```
_:0 a <http://schema.org/Organization> .
_:0 <http://schema.org/address> _:1 .
_:1 a <http://schema.org/PostalAddress> .
_:1 <http://schema.org/name> "Data and Web Science Group" .
_:1 <http://schema.org/addressLocality> "Mannheim" .
_:1 <http://schema.org/postalCode> "68131" .
_:1 <http://schema.org/adressCountry> "Germany" .
```

RDFa, Microdata, and MicroFormats

UNIVERSITY OF MANNHEIM
Data and Web Science Group

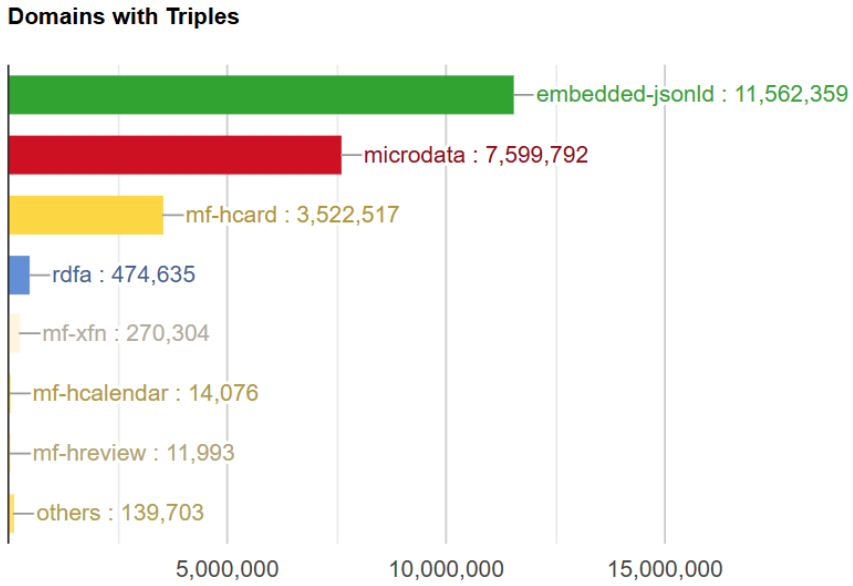
- MicroFormats: fixed vocabularies for persons, addresses, etc.

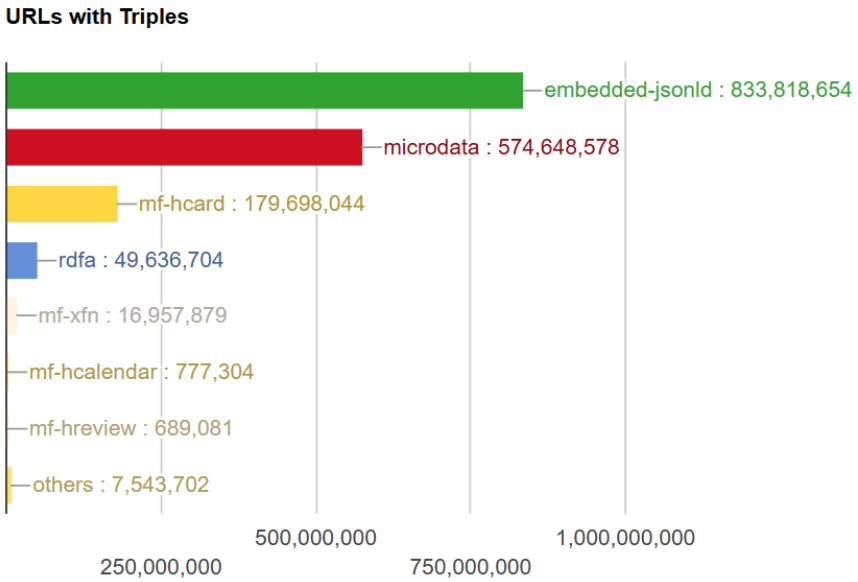
```
<div class="h-card">
  <span class="p-name">Alice</span>
  <span class="p-job-title">Developer</span>
</div>
```

RDFa, Microdata, and MicroFormats

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- MicroFormats: fixed vocabularies for persons, addresses, etc.
- WebDataCommons: Large-Scale Extraction of RDFa, MicroFormats, Microdata, JSON-LD from the Web



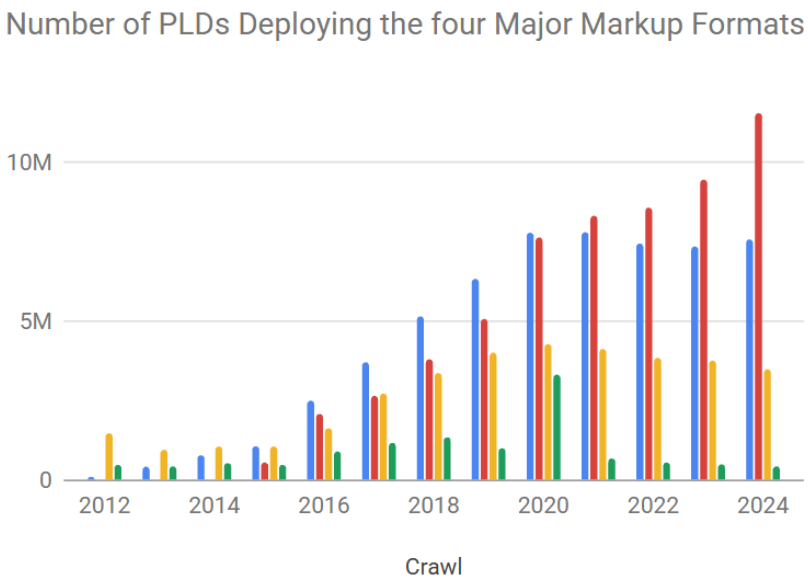


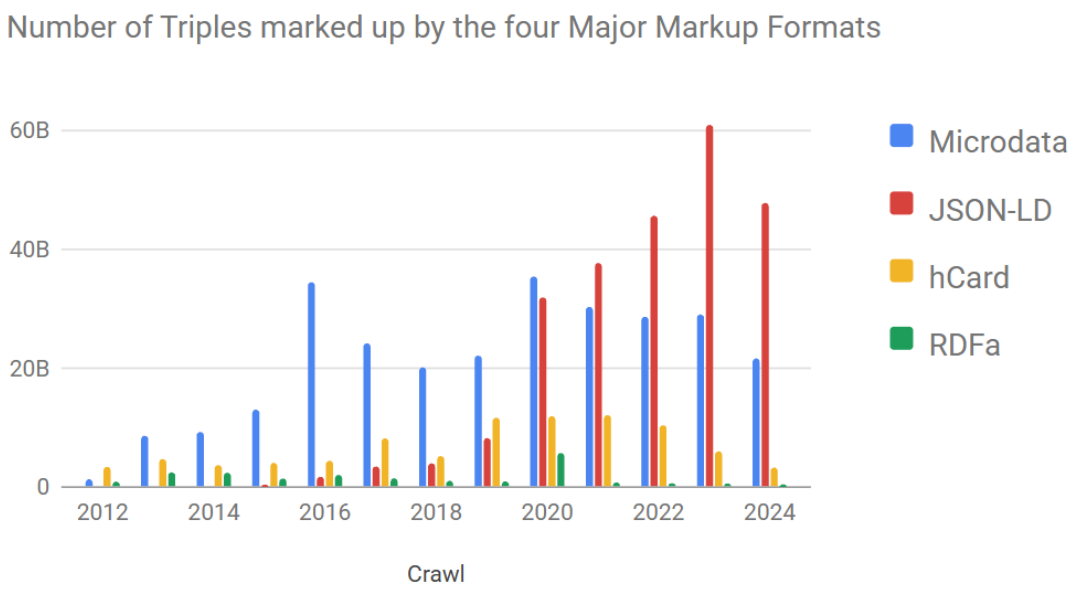
http://webdatacommons.org/structureddata/

RDFa, Microdata, and MicroFormats

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- MicroFormats: fixed vocabularies for persons, addresses, etc.
- WebDataCommons: Large-Scale Extraction of RDFa, MicroFormats, Microdata, JSON-LD from the Web

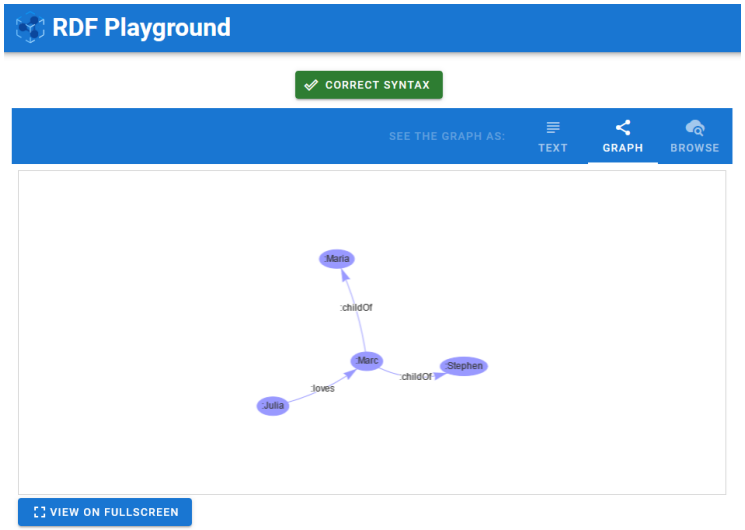




http://webdatacommons.org/structureddata/

RDF Tools

- Storage: relational databases, graph databases, ...
- Validation: validating parsers checking consistency
- Visualization: mostly graph based visualization
- Reasoning: inference over graphs (see next lectures)
- Programming: APIs (see next lectures)

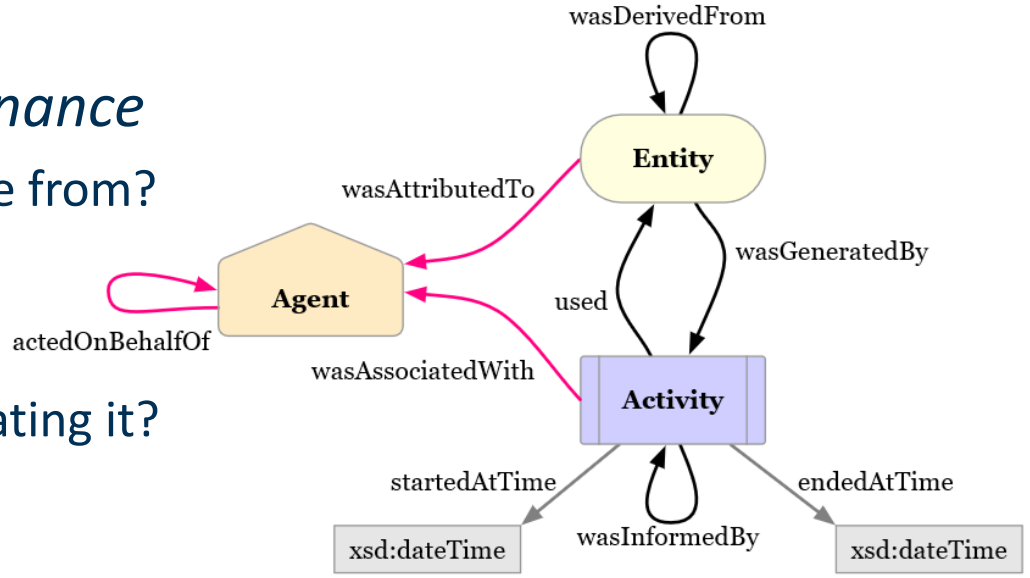


http://rdfplayground.dcc.uchile.cl/

Metadata for RDF

- Recap: Dublin Core was designed as Metadata for the Web
- Knowledge graphs may have metadata as well

- Most prominently: *provenance*
  - Where does the data come from?
  - Who created it?
  - When was it created?
  - What was the process creating it?
  - …



https://www.w3.org/TR/prov-o/

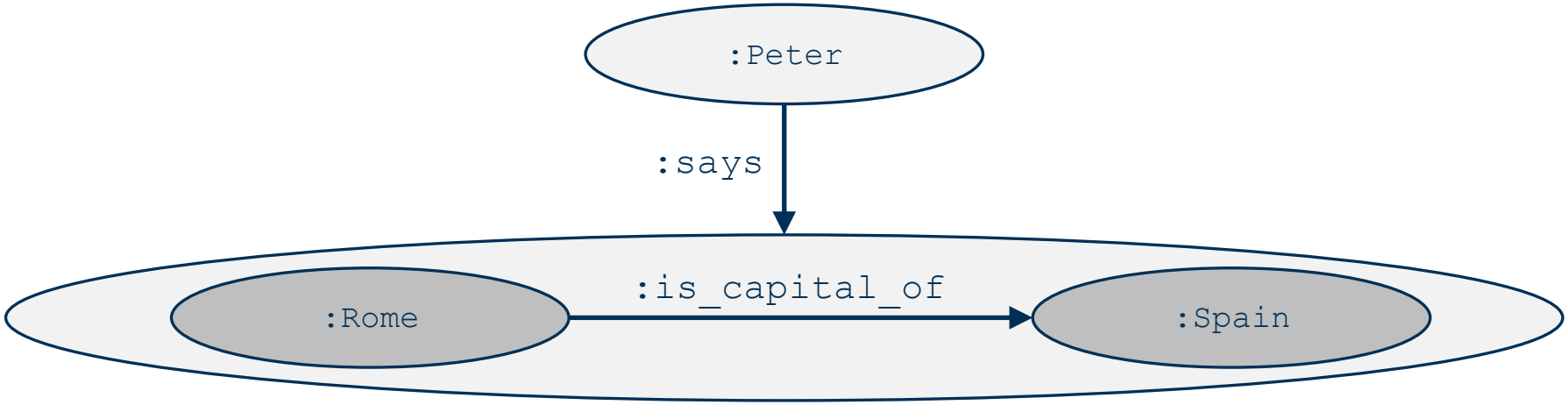
Reification

UNIVERSITY OF MANNHEIM
Data and Web Science Group

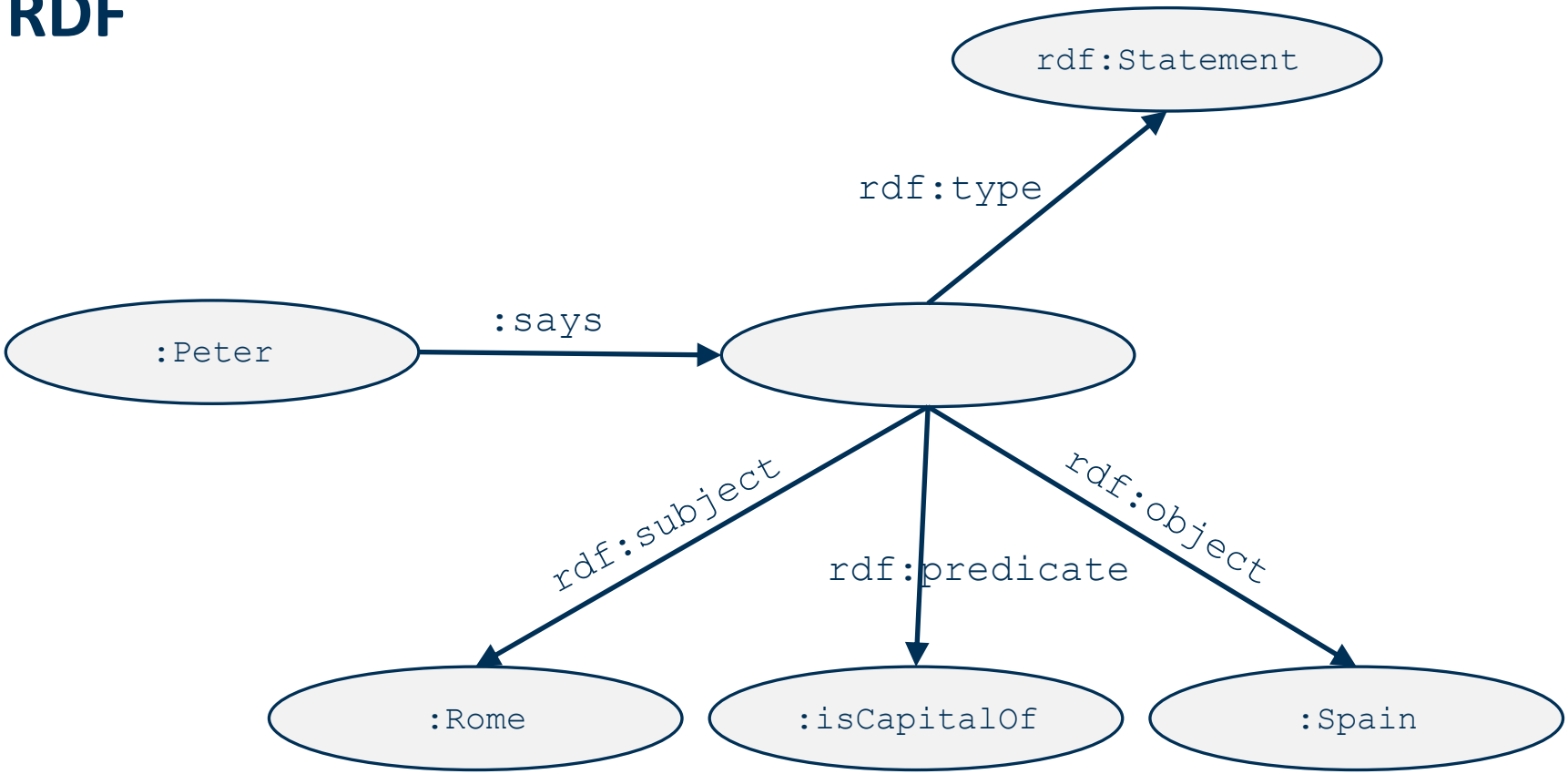
- Latin *res* ("Thing"), *facere* ("make")
  - An Explication
  - making a statement, an opinion etc. the subject of a statement
- In RDF: Statements about statements

- "Peter says that Rome is the capital of Spain."
- Implementation:
  - RDF Statements are considered resources themselves
  - Can be subject or object of other statements
- Reification can have multiple levels
  - “Peter says that Wikipedia states that Rome is the capital of Spain.”

Reification in RDF



Implementing Reification as Standard RDF



Encoding Reification in Turtle

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Variant 1: Named Statement (with URI)

```
:triple1 rdf:type rdf:Statement ;
         rdf:subject :Rome ;
         rdf:predicate :isCapitalOf ;
         rdf:object :Spain .
:Peter :says :triple1 .
```

- Variant 2: Unnamed Statement (Blank Node)

```
:Peter :says [
       a rdf:Statement ;
       rdf:subject :Rome ;
       rdf:predicate :isCapitalOf ;
       rdf:object :Spain .
] .
```

Wrap Up

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- RDF is a language for describing arbitrary things
  - interpretation: set of statements or directed graph
  - Notations: RDF/XML, Turtle
- Special language constructs
  - Blank nodes
  - Reification and its variants
- Semantics
  - Non-unique name assumption
  - Open world assumption
- Embedding in HTML is possible
- Large set of tools is available

A Critical Look in the Rear View Mirror

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- Is RDF more powerful than XML?

- XML is a markup language for information
- In XML, arbitrary elements and attributes can be defined
- XML tag names are meaningless for a computer

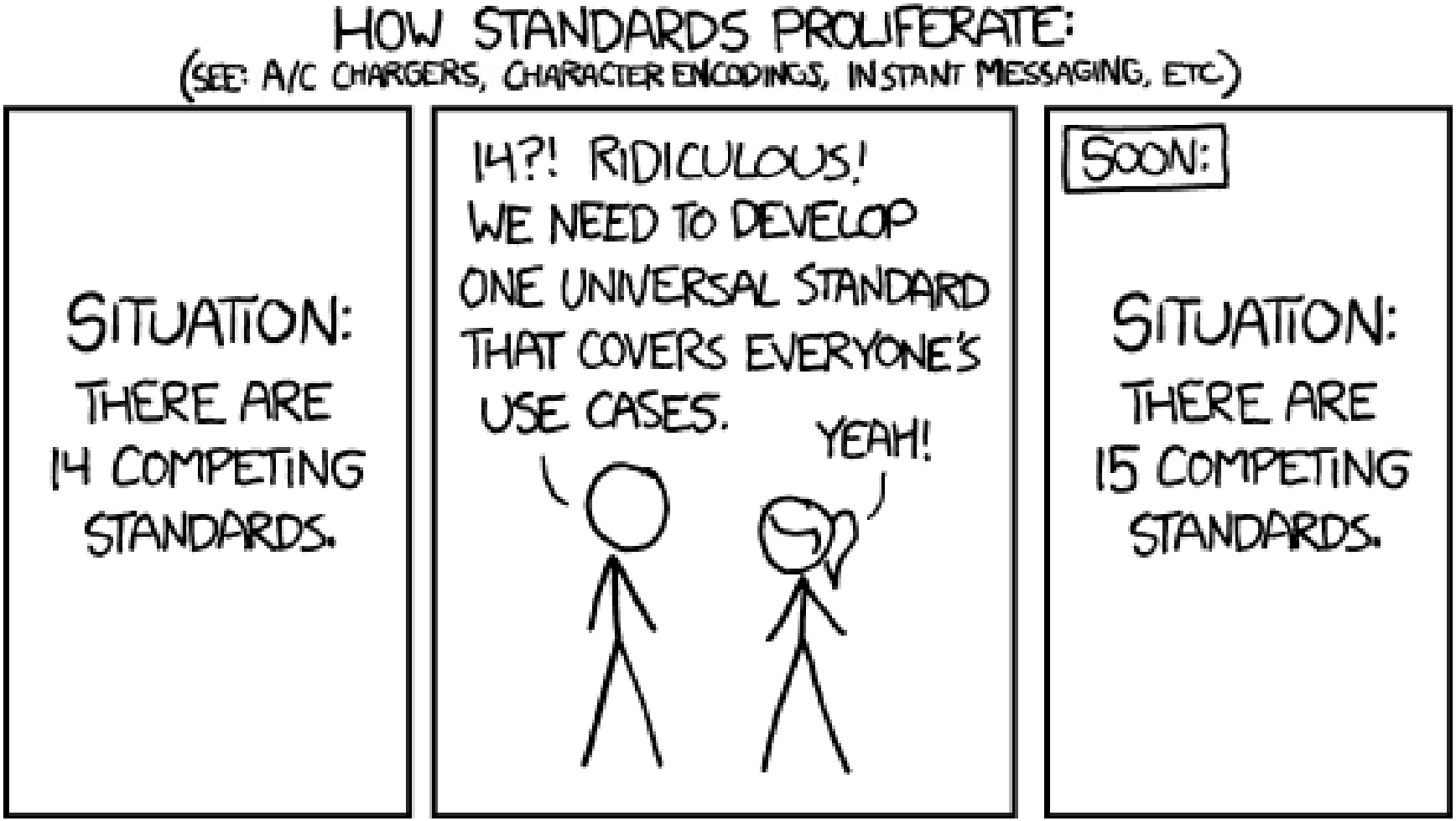
- RDF is a markup language for information
- In RDF, arbitrary classes and predicates can be defined
- RDF class and predicate names are meaningless for a computer

A Critical Look in the Rear View Mirror

UNIVERSITY OF MANNHEIM
Data and Web Science Group

- So, why did we spend an entire lecture on RDF?

A Critical Look in the Rear View Mirror



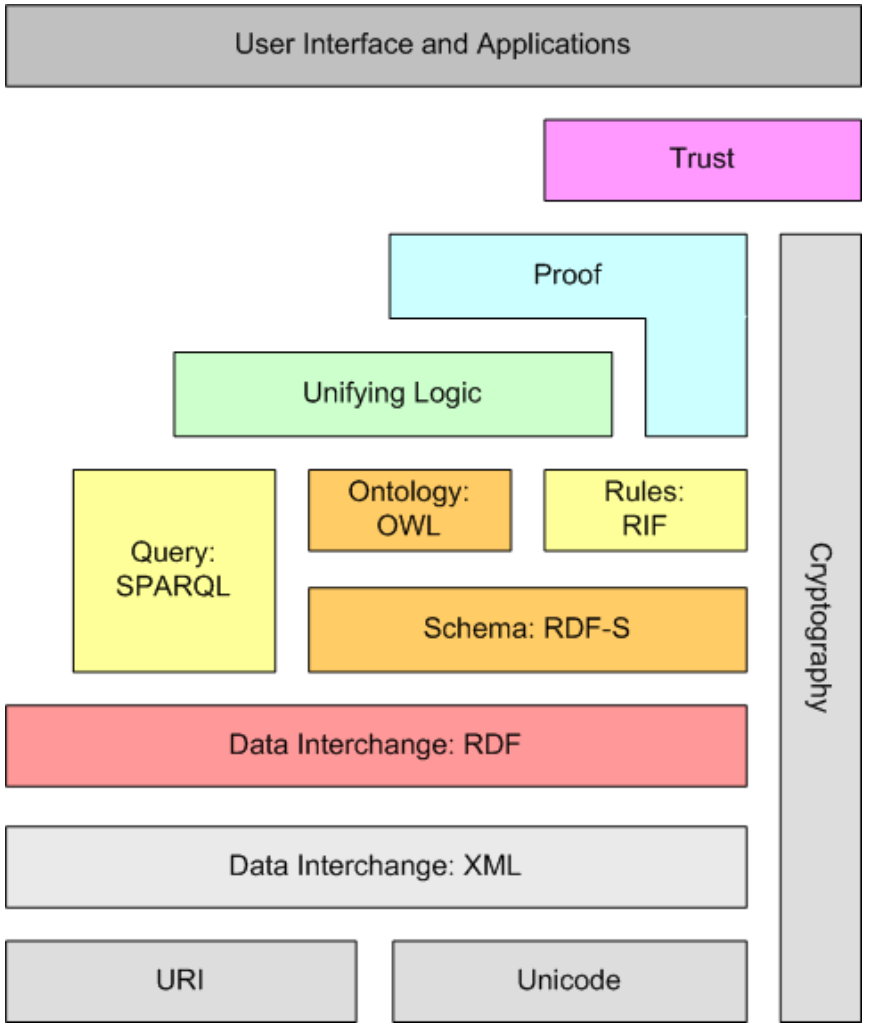
The Semantic Web Stack

UNIVERSITY OF MANNHEIM
Data and Web Science Group

here be dragons...

Knowledge Graph Technologies (This lecture)

Technical Foundations



Berners-Lee (2009): *Semantic Web and Linked Data* http://www.w3.org/2009/Talks/0120-campus-party-tbl/

Questions?