

Organization

IE650 Knowledge Graphs



Hello

- Dr. Sven Hertling
 - Substitute Professor for Data Science
- Research Interests:
 - Knowledge Graph Integration
 - KGs in Combination with Large Language Models
 - Information Extraction
- Room: B6 26, B0.21
- eMail: sven.hertling@uni-mannheim.de
- Sven will teach the lectures



Hello

- Dr. Rita Torres de Sousa
- Researcher
- Research Interests:
 - Knowledge graphs
 - Machine learning
 - Biomedical applications
- Room: B6 26, B0.01
- eMail: rita.sousa@uni-mannheim.de
- Webpage: ritatsousa.github.io
- Rita will teach the exercises



Introduction and Course Outline

- Administration
- Introduction to Knowledge Graphs
- History of Knowledge Graphs
 - Vision of the Semantic Web
 - Building blocks of the Semantic Web
 - Technical foundations

Course Organization

- Lecture
 - Knowledge Graph standards and languages
 - Using public knowledge graphs
 - Creating knowledge graphs
- Exercise
 - Understand knowledge graphs and their principles, play with real data
- Project Work
 - Teams of 3-4 students build a Knowledge Graph application
 - Teams may choose their own data sets and tasks
(in addition, we will propose some pointers for ideas)
 - Write summary about project, present project results
 - **Not graded, but mandatory**
- Final exam
 - **Final grades are only based on written exam**

Course Organization

- Registration
 - You have registered via Portal2
 - You should have access to ILIAS
- Important: course replacement
 - This course replaces IE 650 Semantic Web Technologies
 - You cannot get credits for both courses

Course Contents and Schedule

You are
here

This refers to Monday,
but lecture is on Tuesday
and Exercise on Friday

Week	Lecture (Tuesday)	Exercise (Friday)
07.09.2026	Introduction	Introduction
14.09.2026	Representing Knowledge in Graphs (RDF)	Representing Knowledge in Graphs (RDF)
21.09.2026	Lightweight Knowledge Graph Inference (RDFS)	Lightweight Knowledge Graph Inference (RDFS)
28.09.2026	Linked Data, Knowledge Graph Programming	Linked Data, Knowledge Graph Programming
05.10.2026	Querying Knowledge Graphs	Querying Knowledge Graphs
12.10.2026	Public Knowledge Graphs	Public Knowledge Graphs
19.10.2026	Labeled Property Graphs	Labeled Property Graphs + Project Feedback
26.10.2026	Project Work	Project Work
02.11.2026	Advanced Knowledge Graph Inference (OWL Part 1)	Advanced Knowledge Graph Inference (OWL Part 1)
09.11.2026	Advanced Knowledge Graph Inference (OWL Part 2)	Advanced Knowledge Graph Inference (OWL Part 2)
16.11.2026	Knowledge Modeling	Knowledge Modeling
23.11.2026	Knowledge Graph Quality and Knowledge Integration	Knowledge Graph Quality and Knowledge Integration
30.11.2026	Project Work	Project Presentations
07.12.2026	Project Presentations	Q & A

Deadlines

- Submission of Project Teams
 - **Monday, October 5th, 23:59**
- Submission of project work proposal
 - **Monday, October 19th, 23:59**
- Submission of final project work report & Presentation in PDF format
 - **Wednesday, December 2nd, 23:59**



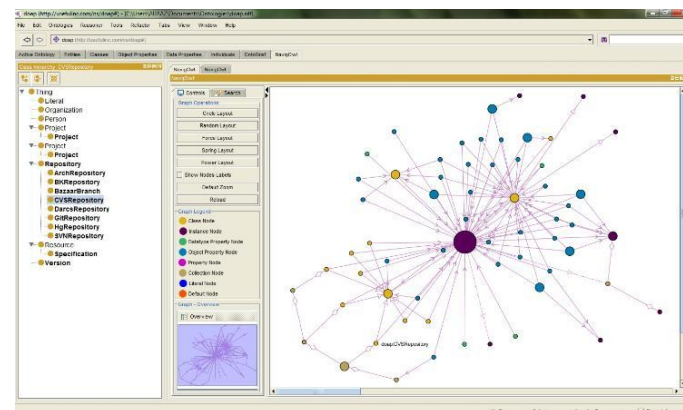
Course Organization

- Lecture Webpage: Slides, Announcements, Web Links
 - Hint: look at version tags of slides!
- Material such as exercises
 - ILIAS eLearning System, <https://ilias.uni-mannheim.de/>
- Time and Location
 - Lecture:
Tuesday, 15:30 – 17:00,
Room A203 Building B6 Part A
 - Exercise:
Friday, 12:00 – 13:30,
Room D007 (2), Building B6,27 Part D



Further Reading and Software

- Follow the links on the website
 - Most material is available online
- Programming environment
 - RDFlib (Python)
 - JENA framework (Java)
- Knowledge graph environment
 - Neo4j
- Ontology engineering environment
 - Protégé
 - <https://protege.stanford.edu/>



Warning

- This lecture contains
 - Cartoons
 - Java and Python code
 - Some digressions to philosophy
- Having said that:
 - have fun! :-)



Questions?



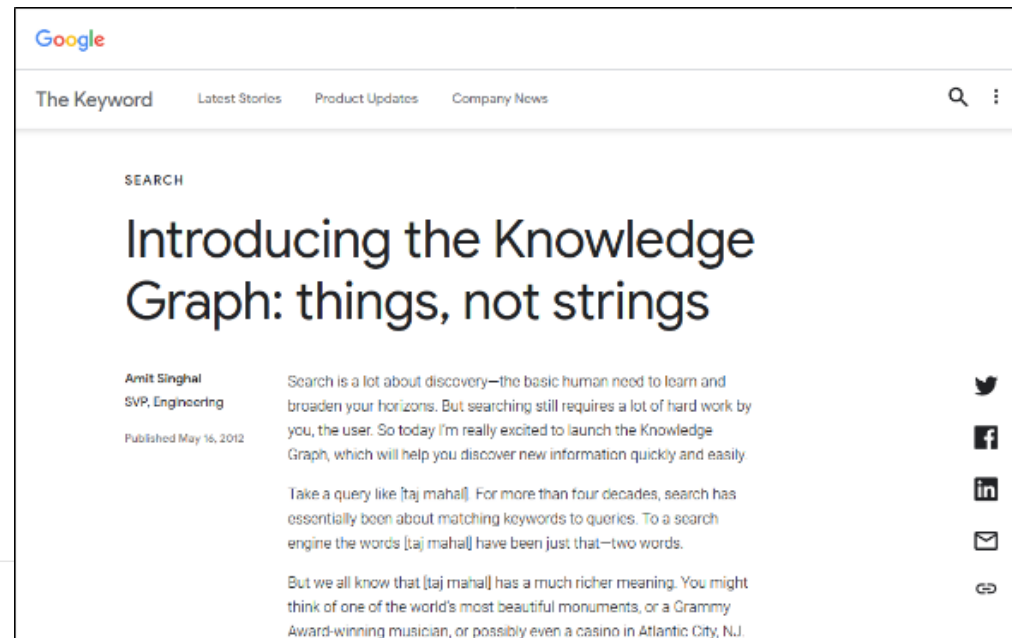
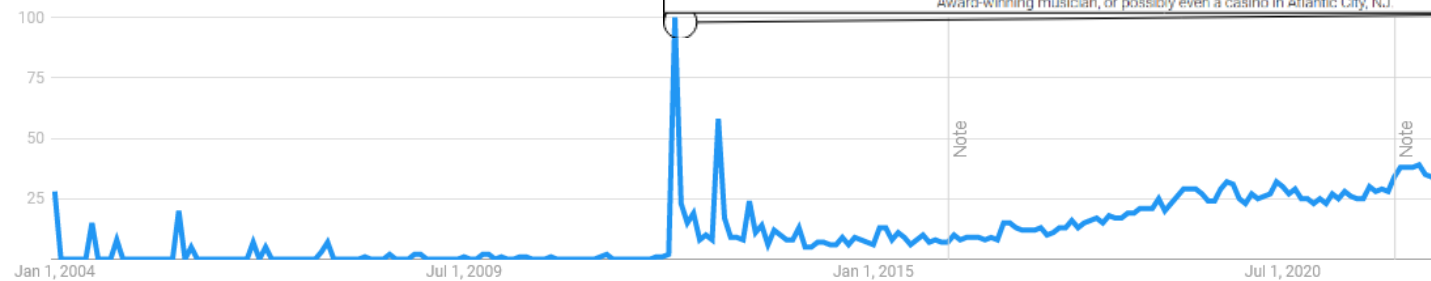
Introduction

IE650 Knowledge Graphs



The Birth of a Buzzword

Interest over time ?



Idea of the Google Knowledge Graph

- Web search in the pre-knowledge graph / pre-LLM age
 - Documents
 - Keywords (not: disambiguated entities)



Hilton Paris Opera, 4-Sterne-Hotel im ...
hiltonhotels.de



Hilton Paris Opera, Paris ...
booking.com



Hilton Paris Opera, 4-Sterne-Hotel im ...
hiltonhotels.de



Paris Hilton – Wikipedia
de.wikipedia.org



Hilton Paris Opera - Fremdenverkehrsamt ...
de.parisinfo.com



Hilton Paris Opera, 4-Sterne-Hotel im ...
hiltonhotels.de



Hilton Paris Opera | Luxury hotel in Paris
hiltonparisoperahotel.com



Hotel Hilton Paris Opera in Paris ab 254 €
agoda.com



► Hotel Hilton Paris Opera, Paris ...
tui.com



Hilton Paris Opera, 4-Sterne-Hotel i...
hiltonhotels.de



Hotel Hilton Paris Opera in Paris ab 254 €
agoda.com



Hilton Paris Opera ab 247 €. Hotels in ...
kayak.de



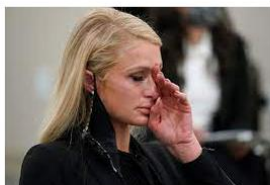
Hilton Paris Opera (1889), Paris ...
historichotelsthenandnow.com



DIE 5 BESTEN Hilton Hotels in Paris ...
tripadvisor.de



Hilton to manage the Concorde Opéra ...
hospitality-on.com



Paris Hilton: »Ich bin der Beweis, dass ...
spiegel.de



Paris Hilton's 2022: Refreshing Her ...
bloomberg.com



Kathy Hilton, Paris Hil...
alamy.de



Zitate von Paris Hilton ...
beruhmte-zitate.de



Paris Hilton | Forum - D...
magazin-forum.de



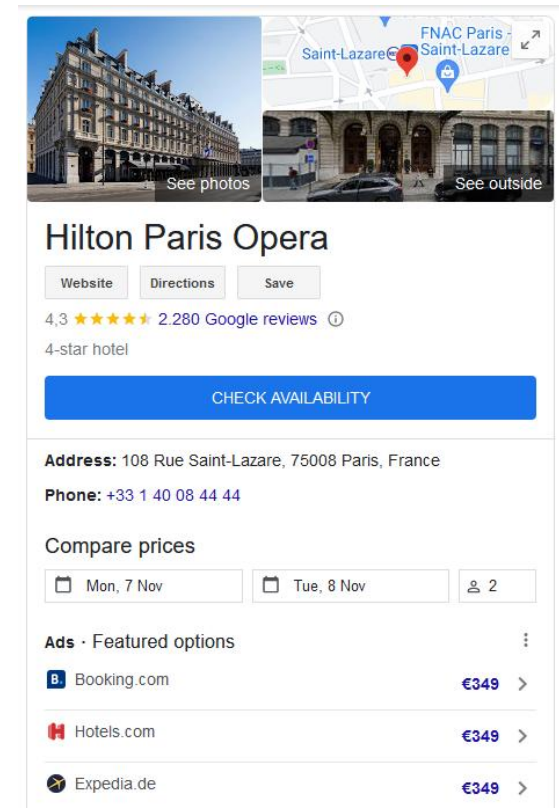
Barron Hilton: Bruder von Paris Hilton ...
gala.de



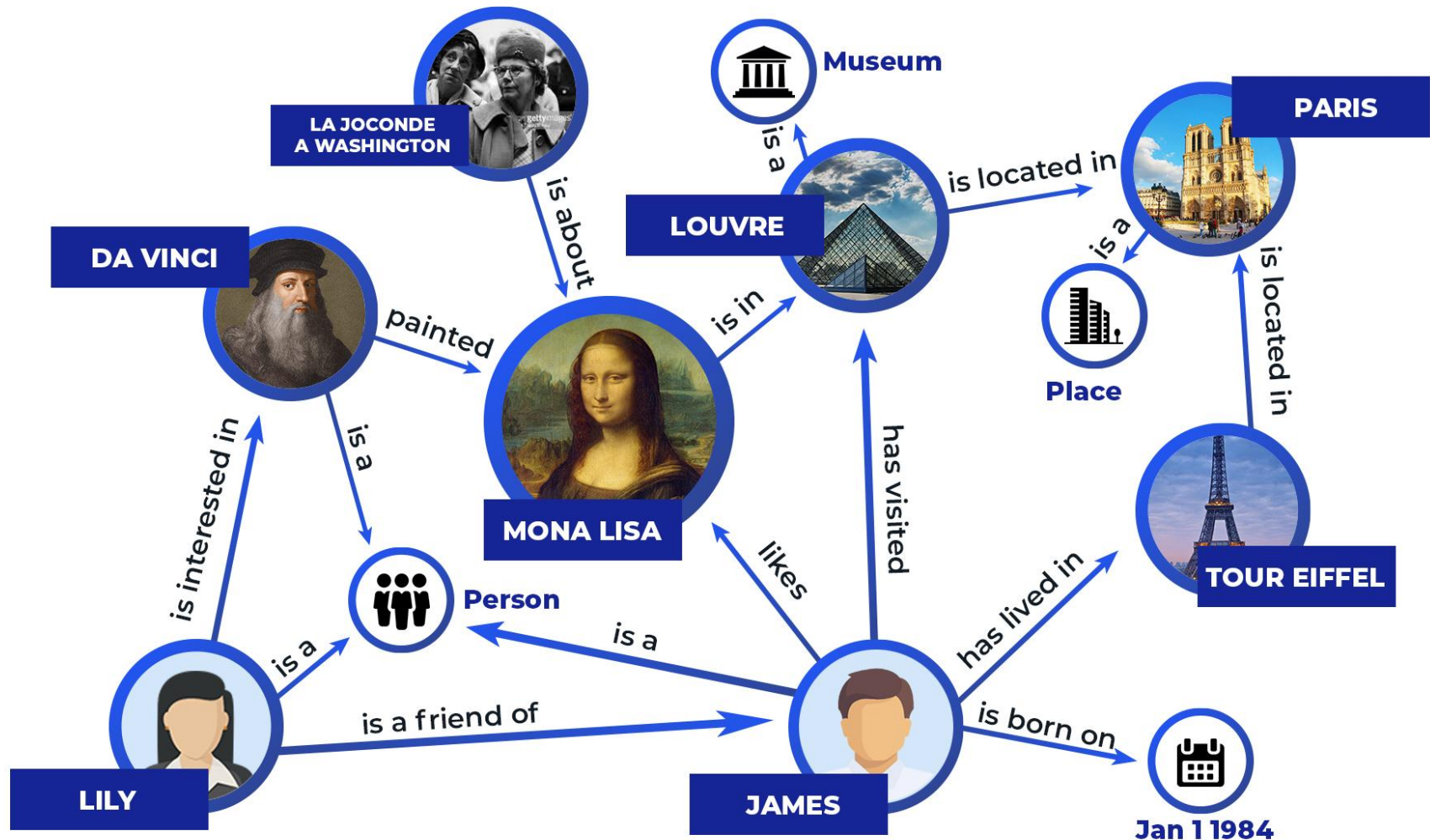
Hilton Paris Opera Hotel, Paris, Franc...
hotel-board.com

Idea of the Google Knowledge Graph

- Web search in the knowledge graph age
 - Backed by structured information
 - All entities are disambiguated
- Linked to other services
 - Ratings and reviews
 - Map information
 - External services (e.g., booking)
 - ...



An Example for a Knowledge Graph



From Google to the World

- Documented list of companies using knowledge graphs
 - Courtesy of Frank van Harmelen, VU Amsterdam



From Google to the World

Hype Cycle for Artificial Intelligence, 2023

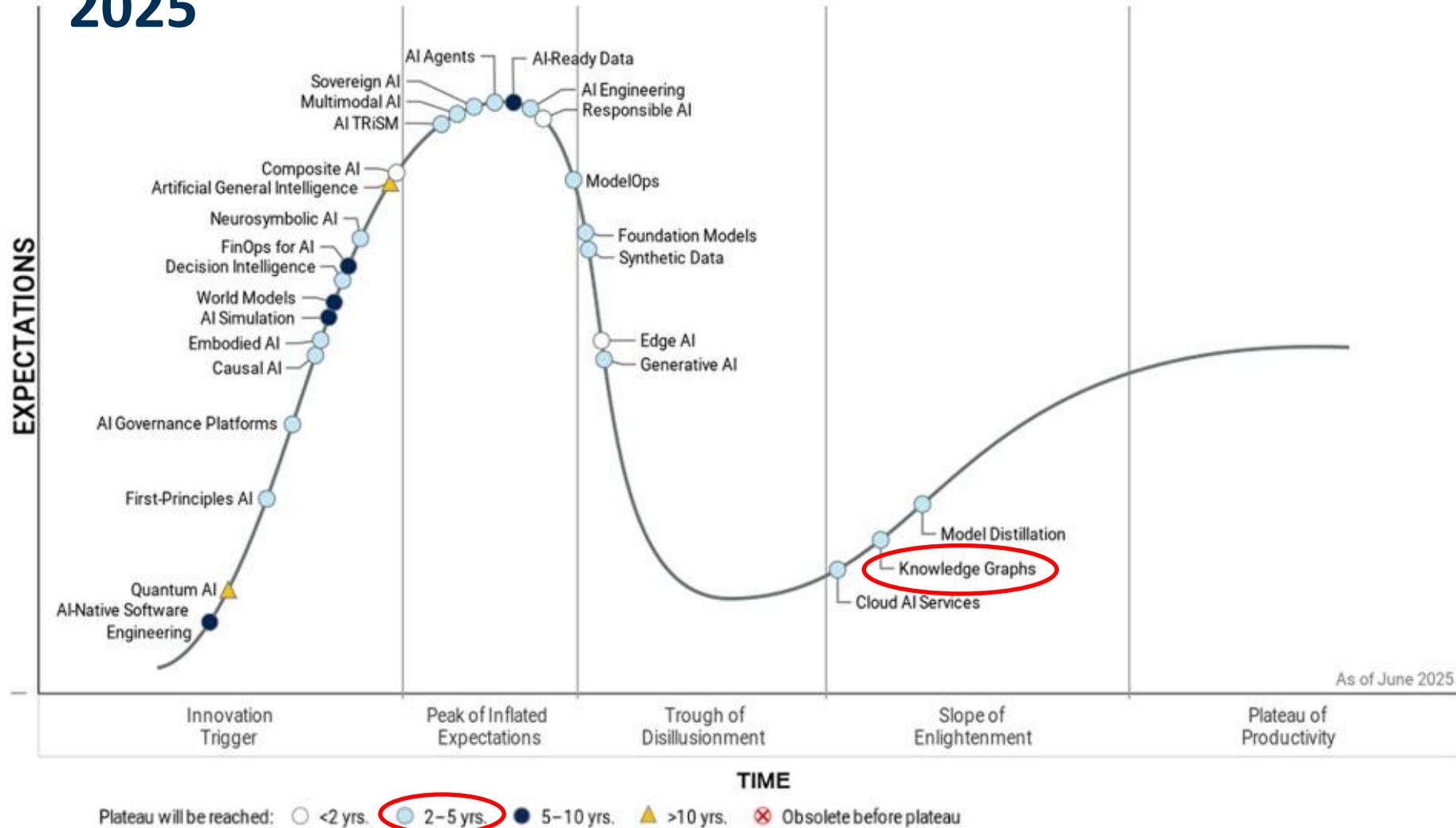


gartner.com

Source: Gartner
© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. 2079794

Gartner®

Hype Cycle for Artificial Intelligence 2025



Knowledge Graph Definitions

- Some example definitions from the literature:
 - Knowledge graphs could be envisaged as a network of all kind things which are relevant to a specific domain or to an organization. They are not limited to abstract concepts and relations but can also contain instances of things like documents and datasets. (Blumauer, 2014)
 - Knowledge graphs are large networks of entities, their semantic types, properties, and relationships between entities. (Kroetsch and Weikum, 2016)

Ehrlinger and Wöβ : Towards a Definition of Knowledge Graphs. 2016

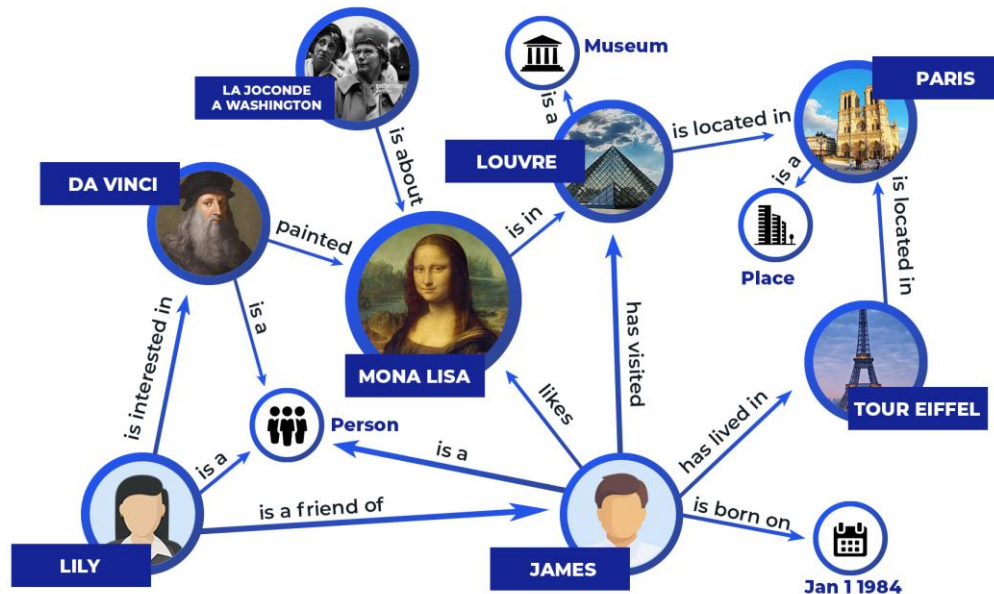
Knowledge Graph Definitions (ctd.)

- Some example definitions from the literature:
 - [...] systems exist, [...], which use a variety of techniques to extract new knowledge, in the form of facts, from the web.
These facts are interrelated, and hence, recently this extracted knowledge has been referred to as a knowledge graph.
(Pujara et al., 2013)
 - A Knowledge Graph
 - (1) mainly describes instances and their relations in a graph,
 - (2) defines possible classes and relations in a schema or ontology,
 - (3) allows for interlinking arbitrary entities with each other, and
 - (4) covers various domains.
(Paulheim, 2017)

Ehrlinger and Wöβ : Towards a Definition of Knowledge Graphs. 2016

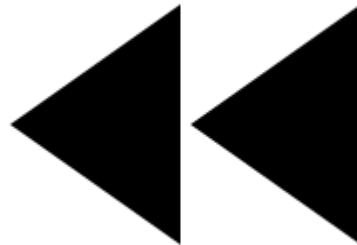
Knowledge Graph Definitions (ctd.)

- Common ground so far:
 - There are entities and relations that are connected and form a graph
 - There is a set of entity and relation types
 - Those are often referred to as a schema or ontology
 - We will get back to this



Rewinding the Time Machine

- Google claim “things, not strings”
 - Entities instead of words in documents
 - Relations between entities explicitly modeled
 - Accessible to humans and machines



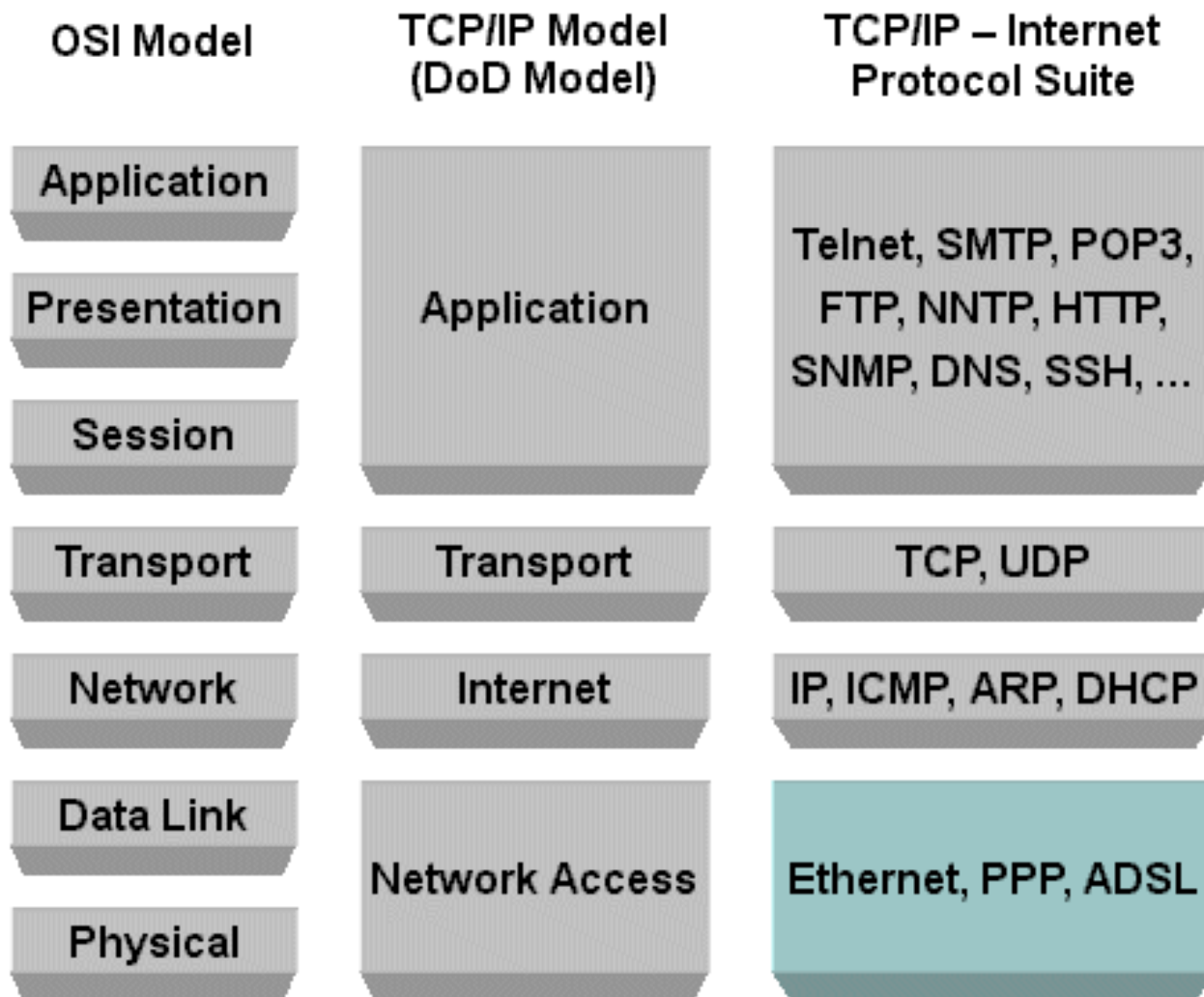
The Vision of the Semantic Web (2001)

- 2001 article by Tim Berners-Lee, Jim Hendler, and Ora Lassila:
 - „The Web is the killer app of the Internet.
The Semantic Web is another killer app of that magnitude.“



Berners-Lee et al. (2001): The Semantic Web. In: Scientific American, Mai 2001.

Web vs. Internet?

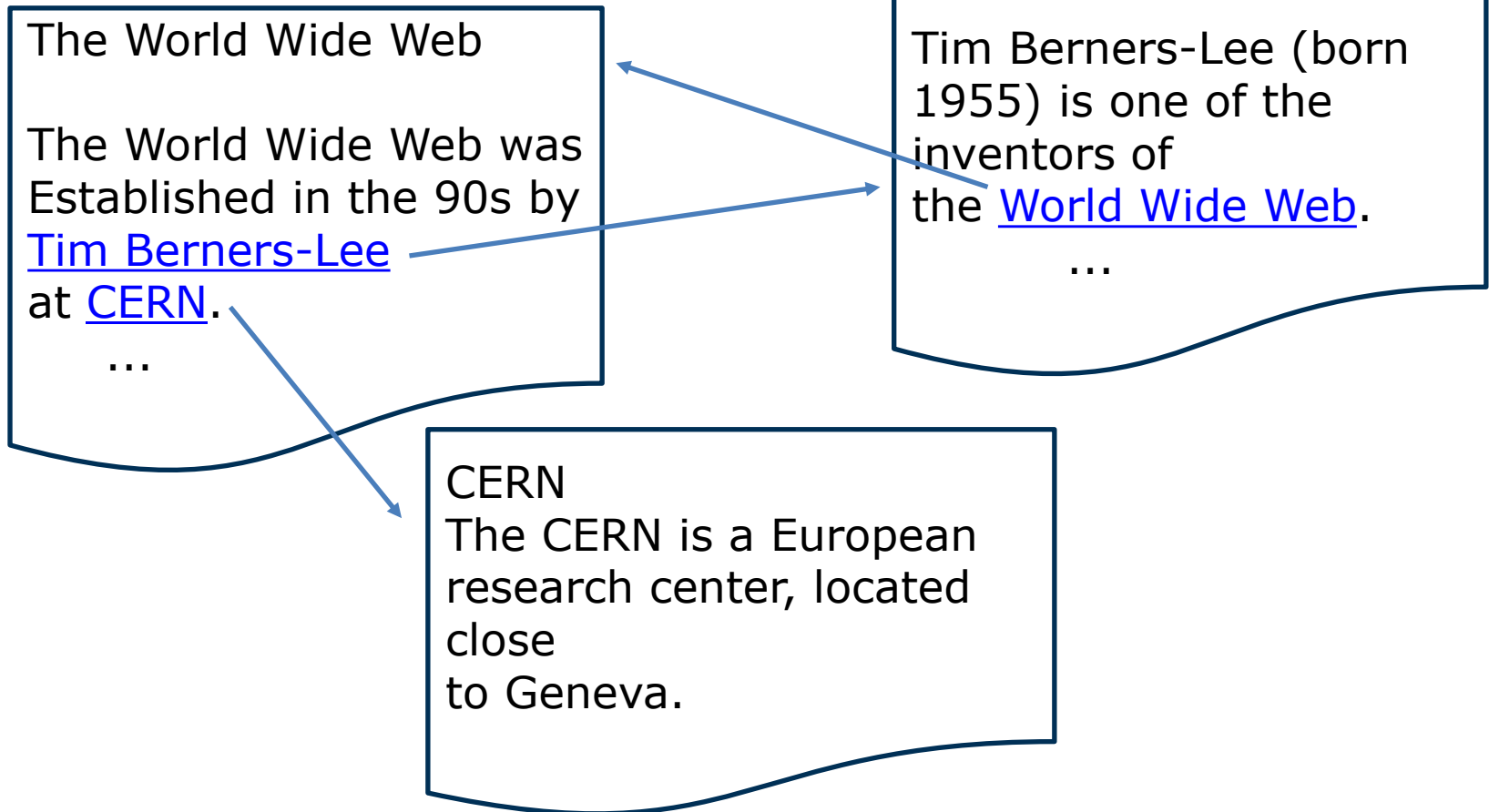


The “Classic” Web

- a.k.a. “World Wide Web”, “Document Web”
- Uses HTTP protocol and URLs
- HTML as a markup language
 - plus CSS, JavaScript, ...
 - plus a few other, more or less standardized formats (GIF, JPEG, Flash, ...)
- Browser as a universal client

The “Classic” Web

- Hypertext: linked documents



The “Classic” Web

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

in the eyes
of a human

Dr. Mark Smith
Physician
Main St. 14
Smalltown
Mon-Fri 9-11 am
Wed 3-6 pm

in the eyes
of a
browser /
computer

Print in bold: „hmf298hmmhudsa“
Print in italics: „mj2i9ji0“
Print normal: „fdsah
02hfadsh0um2m0adsmf0ihm
asdfjköfdsa298ndsfmij32mio
lk2mjpoimjiofdpmsajiomjm“

in the eyes
of an LLM

Dr. Mark Smith Physician Main St.
14 Smalltown Mon-[MASK]

Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - „Mark Smith“
 - „Physician in Smalltown“
 - „Doctor in Smalltown“
 - „Doctor in Smalltown with opening hours on Wednesday afternoon“
 - „Somebody in Smalltown who can fix a broken leg“

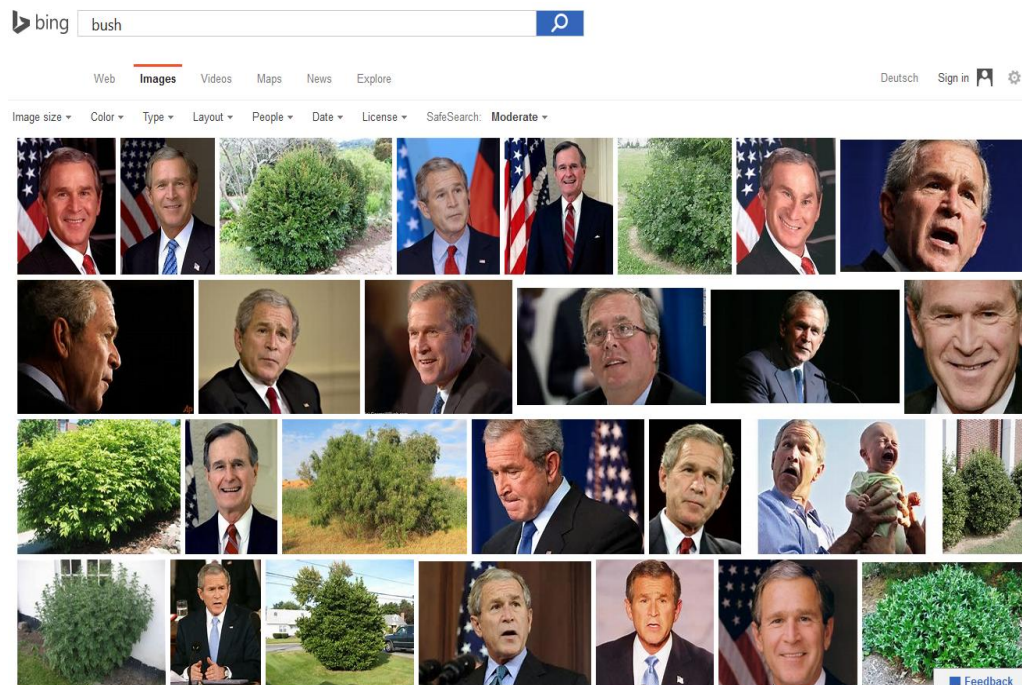
```
<html>
...
  <b>Dr. Mark Smith</b>
  <i>Physician</i>
  Main St. 14
  Smalltown
  Mon-Fri 9-11 am
  Wed 3-6 pm
...
</html>
```

Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - جامعة مانهايم (University of Mannheim)

Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words



Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words
 - Ambiguity of natural language

“One morning I shot an elephant in my pajamas.

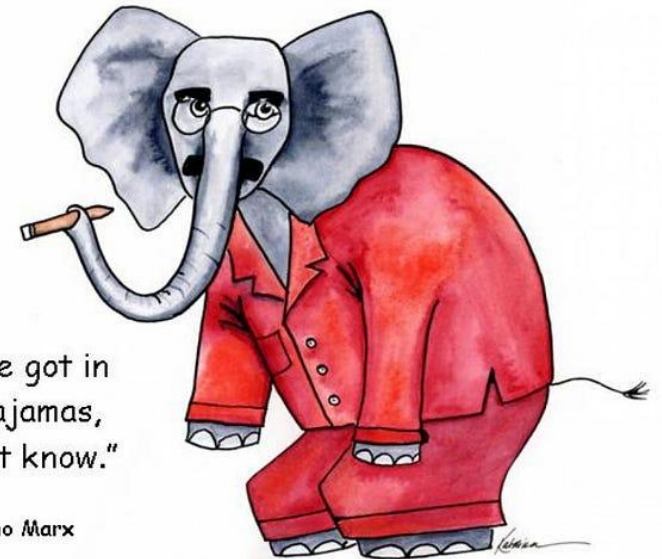
Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words
 - Ambiguity of natural language

“One morning I shot an elephant in my pajamas.

How he got in
my pajamas,
I don't know.”

Groucho Marx



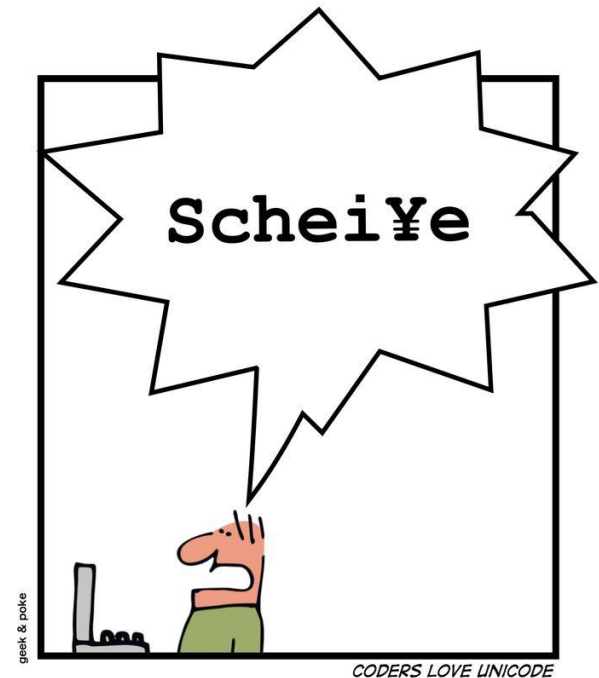
Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words
 - Ambiguity of natural language

“Include your children when baking cookies”

Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words
 - Ambiguity of natural language
- Processing information
 - Formats and encodings



Problems of the “Classic” Web Using Search Engine

- Finding information
 - Keyword based search instead of natural language questions
 - Different natural languages
 - Synonyms, homonyms and polysemous words
 - Ambiguity of natural language
- Processing information
 - Formats and encodings
- Making use of information
 - Distributed across pages
 - e.g., a book's author on the publishers site, address on his/her personal page

Problems of the “Classic” Web Using LLMs

- Generate an image of a analog watch which shows 3:35



DALL-E



Lexica



Midjourney

Problems of the “Classic” Web Using LLMs

- Bias in training data not obvious



eBay.de · Auf Lager
Fashion Men's Busin...



Amazon.de · Auf L...
OLEVS Men's Wat...



WatchGecko
Mechanical Analogue Watches: A Si...



Amazon.de · Auf L...
Men's Watches Ult...



Ethos Watch Boutiq...
What Is An Analog...



Jainx Store
Men's Black Dial Lea...



eBay.de · Auf Lager
USA seller Nike ANALO...



Sylvi Watches
Buy Analog Watches Online for Men & Women in India – ...



Watchroom24 · Auf...
Ice-Watch 012 698...



Maxima Watches
Maxima Urbane Coll...



IndiaMART
Analogue Brown Fossil A



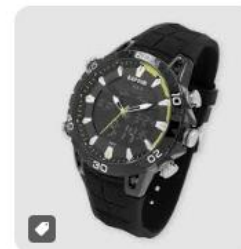
as Analog:Shift
Find a Watch – Analog:Shift



Ubuy Ghana
Classic Waterproof...



Sonata
Buy Online Sonata Poz...

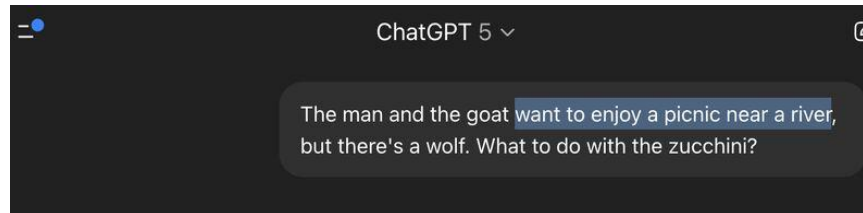


Raptor Watches
Raptor Herrenuhr "Nero...



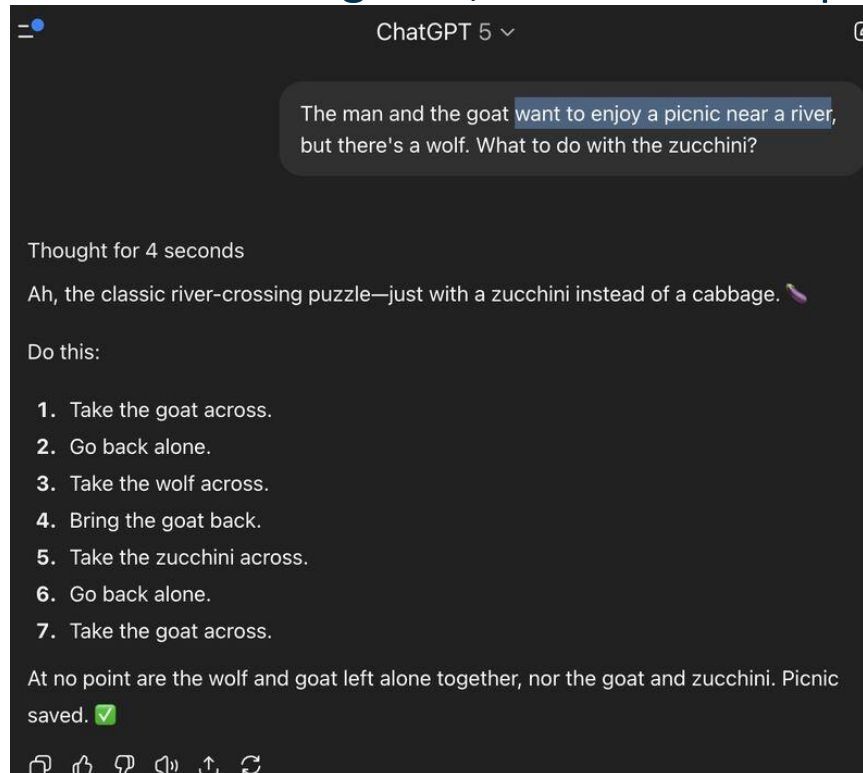
Popular Mechanics
Best Cheap Watches 2021 | Watches U...

Problems of the “Classic” Web Using LLMs



Problems of the “Classic” Web Using LLMs

- LLMs are word pattern matchers
 - When it’s a word pattern matcher, it’s a word pattern matcher.
 - You might think it’s intelligence, but it’s a word pattern matcher.



Problems of LLMs

- Automatic processing of results not easy

How to parse the date?
Need very specific prompt.
Not guaranteed to work for all inputs

when did Des Meagher (an Australian rules footballer) died?

Des Meagher, the Australian rules footballer, died in 2011.



Problems of LLMs

- Hallucinations

when did Des Meagher (an Australian rules footballer) died?

Des Meagher, the Australian rules footballer, died in 2011.



when exactly?

Date not correct

Des Meagher died on March 3, 2011.

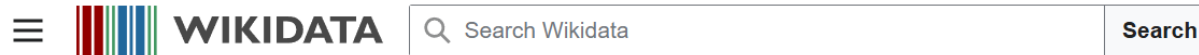
Problems of LLMs

- Provenance information

where is the source?

The information about Des Meagher's death on March 3, 2011, is commonly referenced in Australian football archives and tributes. For a reliable source, you can check official AFL records or reputable sports news outlets that reported on his passing around that time.

Problems of LLMs



Des Meagher (Q5263496)

Item [Discussion](#)

Australian rules footballer

date of death

9 March 2011

edit

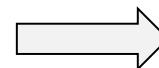
2 references

imported from Wikimedia project	English Wikipedia
reference URL	http://www.heraldsun.com.au/sport/afl/hawthorn-premiership-player-des-maughers-dies/story-e6frf9jf-1226018499083

+ add reference

+ add value

```
SELECT *  
WHERE {  
  wd:Q5263496 wdt:P570 ?deathDate.  
}
```



deathDate

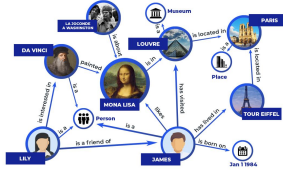
9. März 2011

Comparison of “Machine Knowledge”

Knowledge Graphs
(Dbpedia, YAGO, ...)

Search Engines
(Google, Bing, Yahoo,...)

Large Language Models
(ChatGPT, LLaMA, ...)



Representation:

crisp

mixed

latent

Correctness:

high

mixed

mixed

Freshness:

mostly

fresh

stale (w/o RAG/Agent)

Coverage:

limited

everything

huge

Cost/Query:

low

low

high

Provenance:

tangible

tangible

intangible



Coolness:

nerdy

mundane

amazing

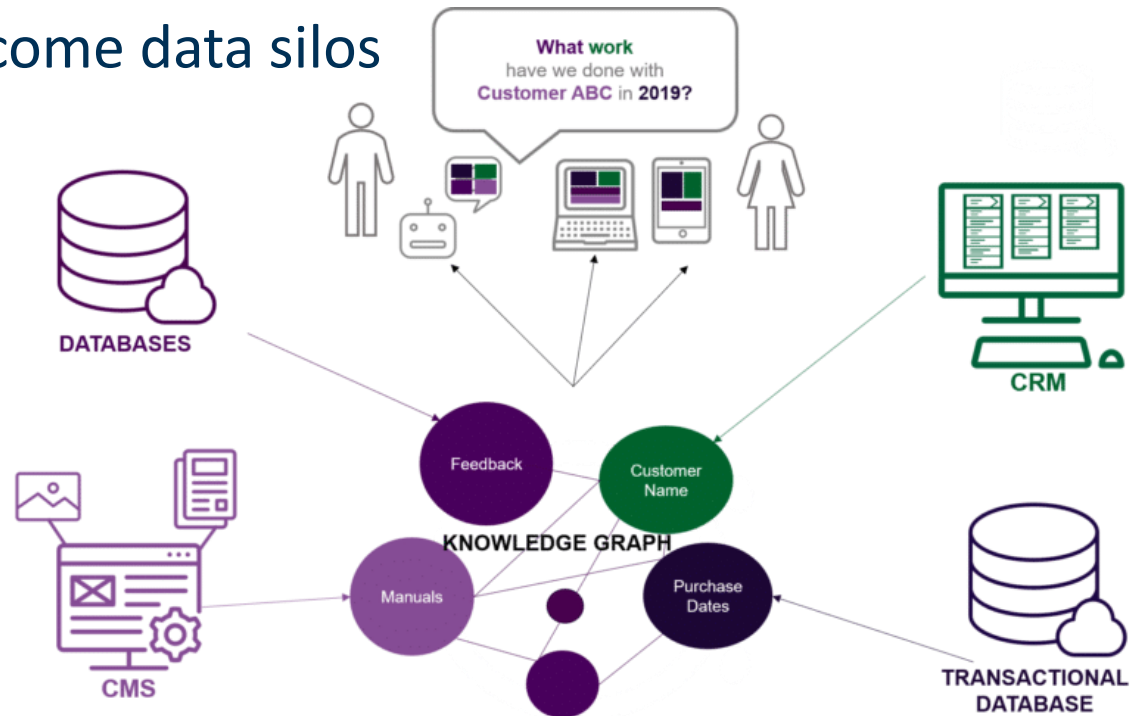
Semantic Web Vision

- Provide information in machine interpretable form
- Make (semantic) links between (data) documents usable
- Facilitate useful (!) complex queries
- Allow logical reasoning



(Enterprise) Knowledge Graph Vision

- Integrate data from different sources
- Make connections between entities in those sources
- Facilitate cross data source queries
- Overcome data silos



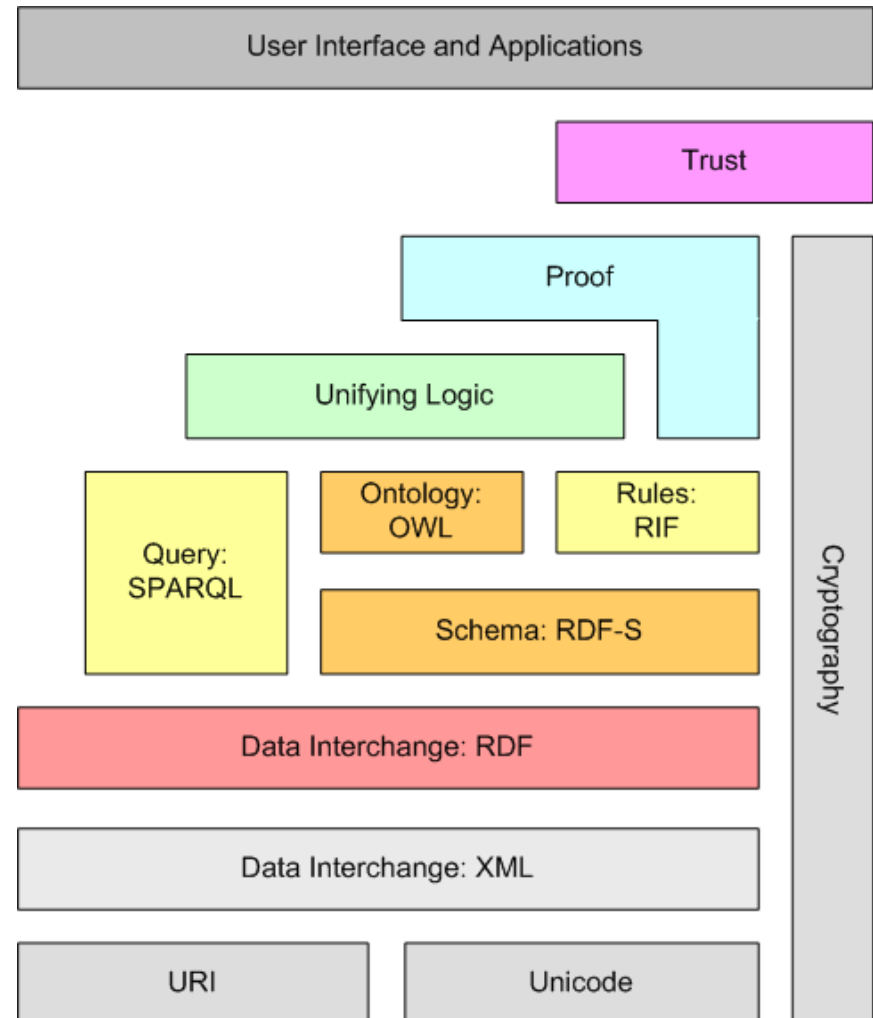
Semantic Web – Architecture



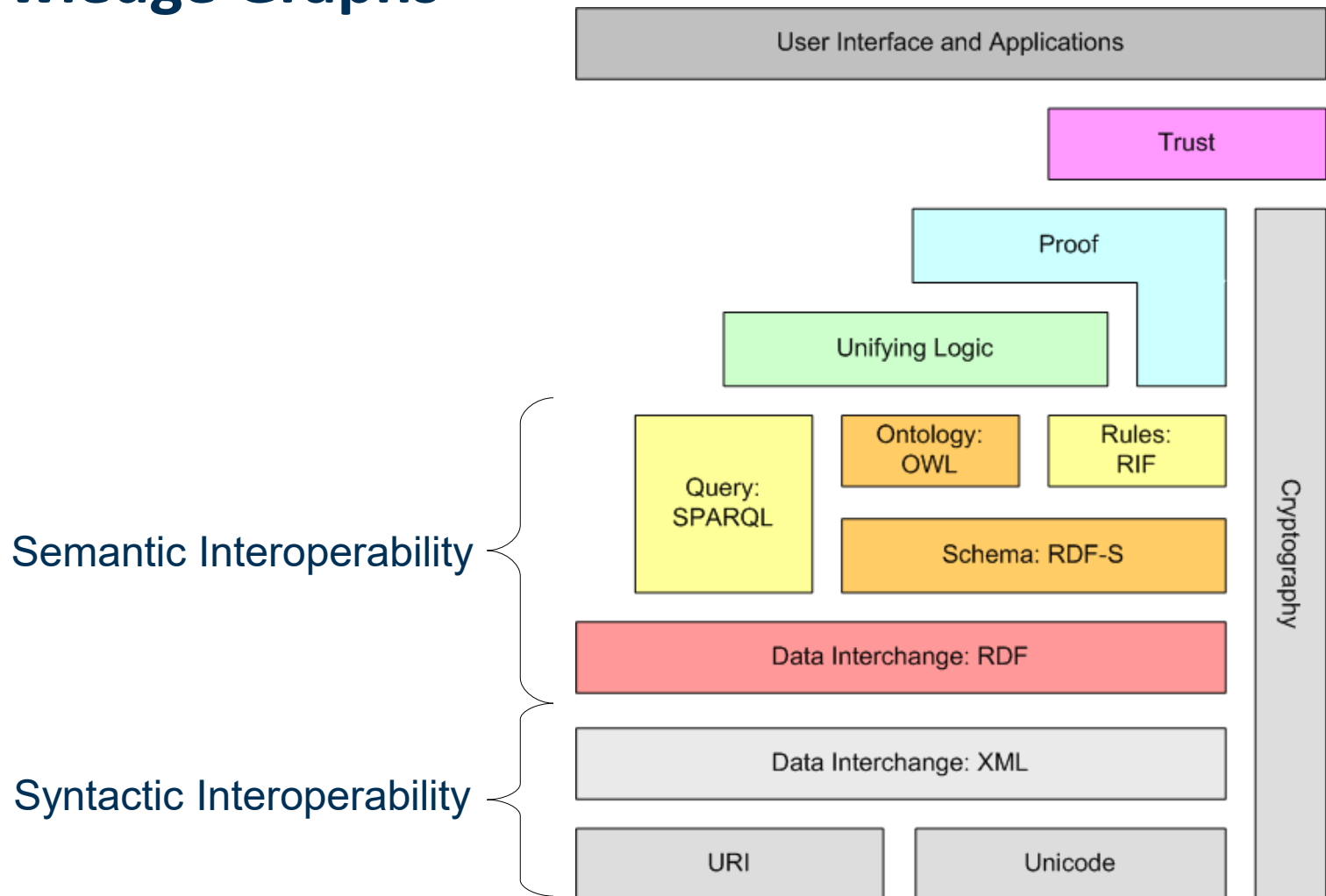
here be dragons...

Knowledge Graph Technologies
(This Course)

Technical
Foundations
(Today)



Data Interoperability with Knowledge Graphs



Short break

- Drink something

AND / OR

- Discuss with your neighbours
 - What are other tools that you use for knowledge retrieval
What are their pros/cons?
 - Which other courses do you visit this semester?

AND / OR

- Solve puzzle

$$8809 = 6$$

$$7111 = 0$$

$$2172 = 0$$

$$6666 = 4$$

$$1111 = 0$$

$$3213 = 0$$

$$7662 = 2$$

$$9313 = 1$$

$$0000 = 4$$

$$2222 = 0$$

$$2581 = ?$$

Syntactic Interoperability: Character Sets

- ASCII („American Standard Code for Information Interchange“) ISO 646 (1963), 127 characters, 95 of which are printable:

!"#\$%&'()*+,-./0123456789:;<=>?

@ABCDEFGHIJKLMNOPQRSTUVWXYZ[\]^_

`abcdefghijklmnopqrstuvwxyz{|}~

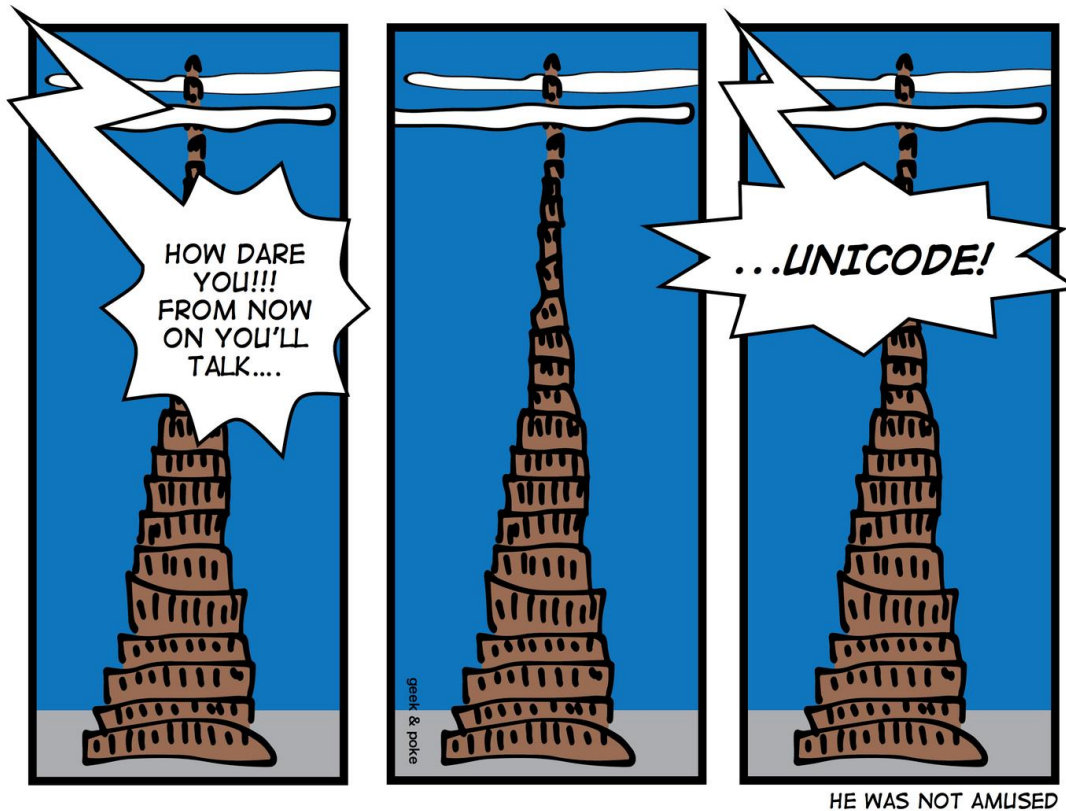
وللحب علامات يتفقها الف
فأولها إدمان النظر، والعج
سرايرها، والمعبرة لضمائرها
ولا يطرف، يتنقل بتنقل
ن مال، كالهرباء مع الشمس

- Extension to 8 Bit: ISO 8859-1 to -16(1998)
 - covers major European languages
 - most well known: 8859-1 („Latin-1“)
- The Web, however,
speaks many more languages...

我爱中国
国中爱我

Syntactic Interoperability: Multilinguality

TOWER OF BABEL



Syntactic Interoperability: Unicode

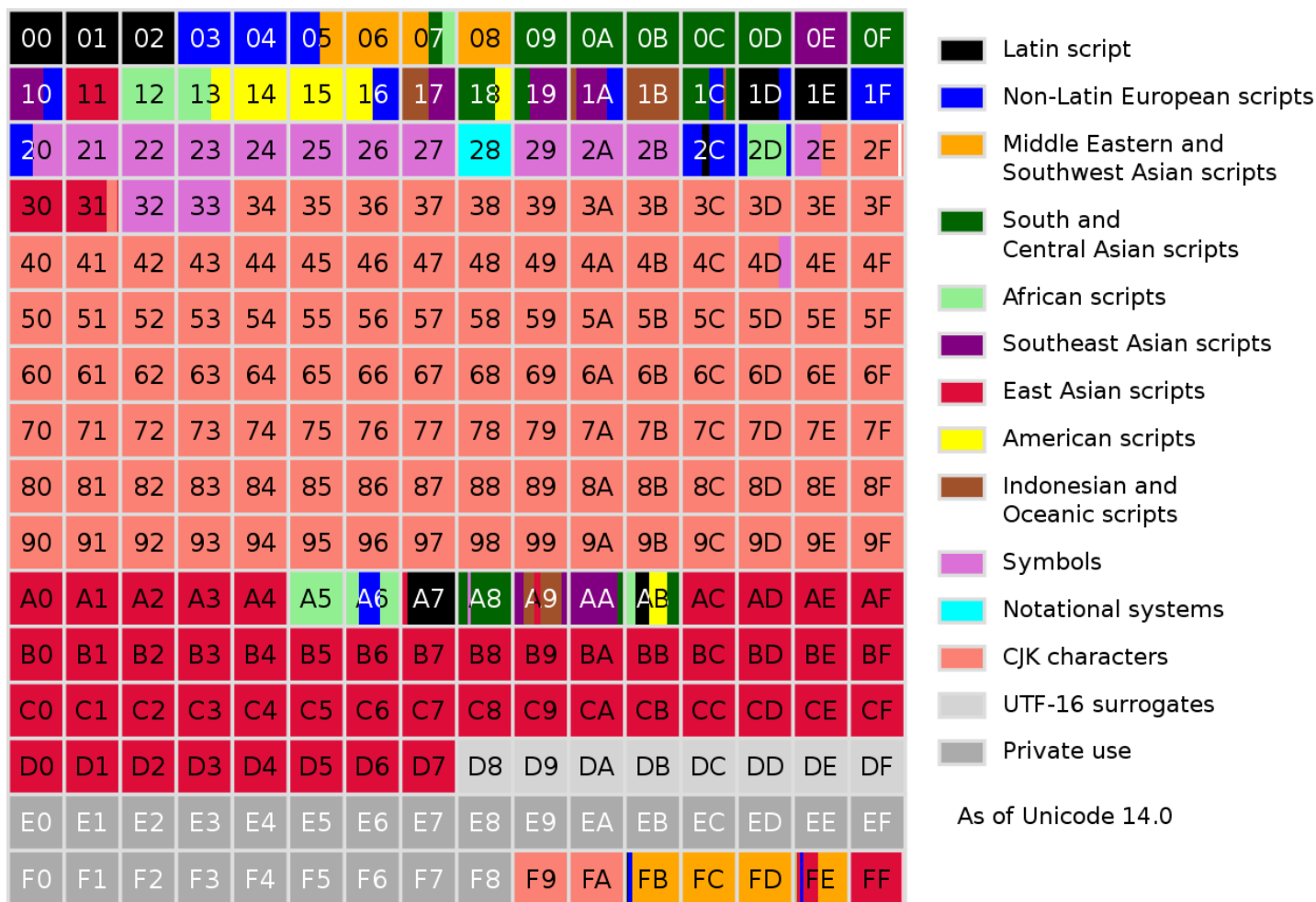
- ISO 10646
 - first version 1991 (Europe, Near East, India)
 - Unicode 14.0 (September 2021)
 - defines ~144,000 characters
 - covers even very exotic languages
 - Plus: currency symbols, emojis, sign language, music notation...



Klingon
is still
missing!!!

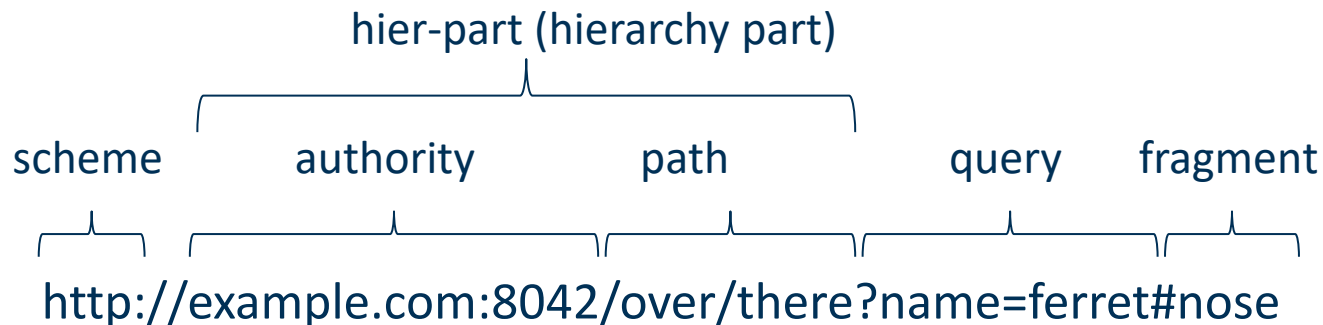


Syntactic Interoperability: Unicode



Uniform Resource Identifiers (URIs)

- “Things, not strings” requires identifiers for *things*
 - URIs: Proposed by Tim-Berners-Lee as „Universal Resource Identifier“ (IETF RFC 1630)
 - Standardized: IETF RFC 3986 (2005)
- Used for naming and finding resources on the Web



URIs vs. URLs

- Uniform Resource Locators (IETF RFC 1738, 1994) are a *subset* of URIs
- URIs can refer to *arbitrary* things
- A URL refers to a resource on the Web
- Typical URL prefixes
 - http
 - ftp
 - mailto
 - telnet
 - file

URIs vs. IRIs

- URIs only allow ASCII characters
- IRI (Internationalized Resource Identifier) also allows any Unicode character

URLs on the Web

- Most common usage:
Hyperlinks in HTML documents
- Links usually do not carry
any meta information

Tim Berners-Lee

Tim Berners-Lee (born 1955) is one of the inventors of the [World Wide Web](#).

...

The World Wide Web

The World Wide Web was initiated in the 90s by [Tim Berners-Lee](#) at [CERN](#).

...

<http://www.w3.org/WWW/>

<http://www.w3.org/People/Berners-Lee/>

Information Representation in XML

- XML (eXtensible Markup Language)
 - A W3C standard since 1998
 - Universal format for data exchange and integration



```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

Why XML?

- One possible serialization of RDF (which we will discuss later)
- Have a nice way of defining a schema (see later today)
 - Allows some comparison to the RDF schema we will discuss in two weeks



XML: Basic Concepts

- Tags (arbitrarily definable):
 - Form pairs: `<physician> ... </physician>`
 - ...or empty element tags `<young />`
- Attributes:
 - `<physician location="Smalltown">`
- Tags are nested (with exactly one root element):
`<physician>`
 `<address> ... </address>`
`</physician>`

XML: Well-formed Documents

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  <telephone>
    <number>+44 123 456789</number>
  </address>
  </telephone>
  <hours>
    <monday>9-11 am
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

HTML and XML

- HTML documents look like XML documents
 - ...but they are usually not well-formed!

```
<p>Look at this!<img src=smiley.gif> <br>
```

- XHTML: HTML as well-formed XML documents
- A W3C standard since 2000

```
<p>Look at this! <br/> </p>
```


XPath: Accessing Information in XML

- Query language for XML
- A W3C standard since 1999 (Version 2.0: 2010)

`/physician[name='Dr. Mark Smith']/telephone/number`

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

Namespaces in XML

- Elements with the same name can occur in different places
 - ...but the contents and semantics may differ
- How can we tell them apart?

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

Namespaces in XML

- Namespace definition using prefixes (Notation: `prefix:name`)
- Each namespace itself is a URI
- Default namespaces may be defined

```
<physician xmlns      ="http://www.med.com/physician"
           xmlns:addr ="http://www.med.com/addr">
  <name>Dr. Mark Smith</name>
  <addr:address>
    <addr:street>Main St.</addr:street>
    <addr:number>14</addr:number>
    <addr:city>Smalltown</addr:city>
  </addr:address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

XML: Document Type Definition (DTD)

- Defines valid elements for a class of XML documents
 - Names
 - allowed attributes
 - allowed nested child elements
- DTD is a part of the W3C's XML specification
- XML documents matching a DTD are called “valid”

XML: Document Type Definition (DTD)

```
<!DOCTYPE physician [  
  
  <!ELEMENT physician (  
    name,  
    address*,  
    telephone?,  
    fax?,  
    hours)>  
  
  <!ELEMENT address (  
    street,  
    number,  
    city)>  
  
  <!ELEMENT street (#PCDATA)>  
  
    ...  
  ]>
```

```
<!DOCTYPE physician SYSTEM  
  "physician.dtd">  
  
  <physician>  
    <name>Dr. Mark Smith</name>  
    <address>  
      <street>Main St.</street>  
      <number>14</number>  
      <city>Smalltown</city>  
    </address>  
    <telephone>  
      <number>+44 123 456789</number>  
    </telephone>  
    <hours>  
      <monday>9-11 am</monday>  
      <tuesday>9-11 am</tuesday>  
      ...  
    </hours>  
  </physician>
```

XML: Document Type Definition (DTD)

- Definition of child elements and their order

```
<!ELEMENT address(street,no,line*,zip,city,state?)>
```

?, + and * mark optional and possible multiple elements

- Definition of attribute lists

```
<!ATTLIST person title CDATA>
```

- Allowed modifiers: #REQUIRED, #FIXED, #IMPLIED, “...”
- Enumerating allowed values: (dr|prof)

- Definition of entities:

```
<!ENTITY sw "Semantic Web">
```

- May be used as shortcuts in the XML document: &sw;

- W3C-Standard (since 2004)
- XML schemas are XML files themselves
- More flexible than DTDs:
 - Minimum and maximum number of elements
 - Combinations of elements
(either/or, combinations w/out fixed order, ...)
 - Data types (Numbers, dates, ...), own types may be defined
 - Support for namespaces
 - Possibility to create modular schemas

XML Schema

```
<xs:schema elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
```

```
  <xs:element name="physician">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="name"
          type="xs:string">
        <xs:element name="address">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="street"
                type="xs:string">
              ...
            </xs:sequence>
          </xs:complexType>
        </xs:element>
        ...
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
<physician xmlns:xsi=
"http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation=
  "physician.xsd">
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

XML Schema – Modular Schemas

```
<xs:schema elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:addr="http://www.address.com/">
```

```
  <xs:import
    namespace="http://www.address.com/"
    schemaLocation="address.xsd"/>
  <xs:element name="physician">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="name"
          type="xs:string">
          <xs:element ref="addr:address" />
          ...
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```

```
<xs:schema elementFormDefault="qualified"
xmlns:xs="http://www.w3.org/2001/XMLSchema">
```

```
  <xs:element name="address">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="street"
          type="xs:string">
          ...
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:schema>
```



So, what does a DTD/Schema Define?

- Syntax – σύνταξις („together“ + „order“)
 - Which elements are there?
 - How are they arranged?
 - Which combinations are allowed?
- ...as opposed to: Semantics - σημαίνει („denote“)
 - How to interpret the contents of an element?
 - What is their relation?

Syntax and Semantics: The Linguists' View

- Syntax: how are correct sentences formed?
„This sentence no verb.“
„The dreaming lamp ~~give~~ gives a freshly cut ~~juices~~ juice
to the ~~tire~~ tired sink.
- Semantics: what does a word and sentence *mean*?
- Notes
 - syntactic correctness does not guarantee semantic interpretability
 - semantic interpretability does not require syntactic correctness (for humans)

Syntax and Semantics: The Linguists' View

syntax

Definition of **knowledge** noun from the Oxford Advanced Learner's Dictionary

 **knowledge** noun

BrE /'nɒlɪdʒ/ ; NAmE /'nɑ:lɪdʒ/ 

★ Add to my wordlist

1  [uncountable, singular] the information, understanding and skills that you gain through education or experience

- *practical/medical/scientific knowledge*
- **knowledge of/about something** *He has a wide knowledge of painting and music.*
- *There is a lack of knowledge about the tax system.*

→ SEE RELATED ENTRIES: [Teaching and learning](#)

2  [uncountable] the state of knowing about a particular fact or situation

- *She sent the letter **without my knowledge**.*
- *The film was made with the Prince's **full knowledge** and approval.*
- *She was impatient **in the knowledge** that time was limited.*
- *I went to sleep **secure in the knowledge** that I was not alone in the house.*
- *They could relax **safe in the knowledge** that they had the funding for the project.*
- *He **denied all knowledge** of the affair.*

3 **knowledge economy/industry/worker** working with information rather than producing goods

- *the emergence of consultancy as a knowledge industry*
- *the shift toward a knowledge economy*

semantics

So, what does a DTD/XML Schema Define?

Employee catalog
of the hospital

Yellow Pages

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

(probably)
the private address

?

=

```
<physician>
  <name>Dr. Mark Smith</name>
  <address>
    <street>Main St.</street>
    <number>14</number>
    <city>Smalltown</city>
  </address>
  <telephone>
    <number>+44 123 456789</number>
  </telephone>
  <hours>
    <monday>9-11 am</monday>
    <tuesday>9-11 am</tuesday>
    ...
  </hours>
</physician>
```

(probably)
the work address

So, what does a DTD/XML Schema Define?

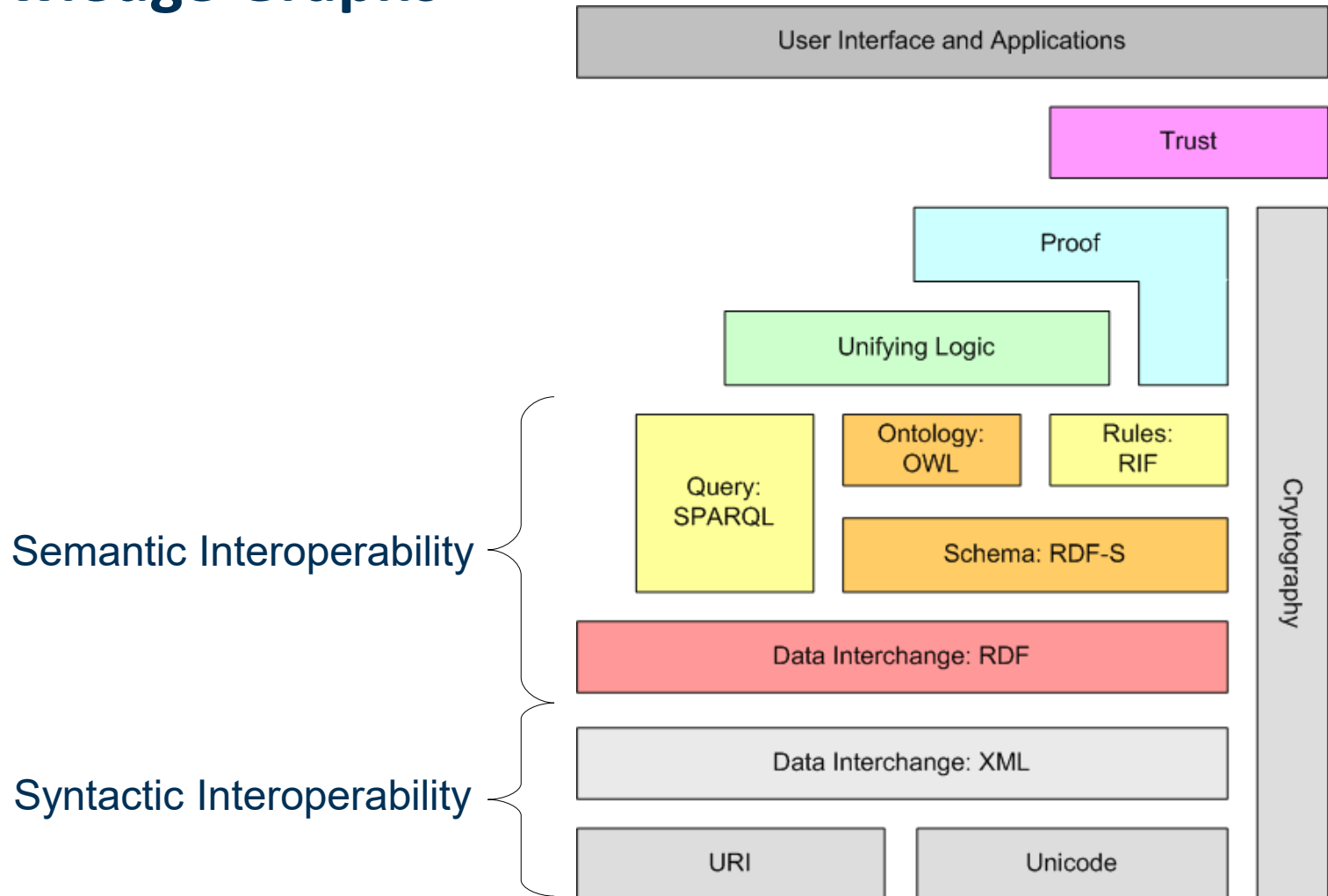
- XML Schema / DTD defines the *syntax* of an XML document, but not its *semantics*
- Tag names are not interpretable by machines
 - i.e., they do not ease the information retrieval process...
 - Semantics of the data is hidden – usually hard wired in the application
- The Semantic Web is meant as a remedy to that problem
 - *Semantic Web is/can do more than XML!*

```
<2nf3oiü*>
  <34f0>Dr. Mark Smith</34f0>
  <rm4935r>
    <e2m4>Main St.</e2m4>
    <dur3>14</dur3>
    <jfa34>Smalltown</jfa34>
  </rm4935r>
  <d24r3fmö>
    <deß5>+44 123 456789</deß5>
  </d24r3fmö>
  <vsfif>
    <f02>9-11 am</f02>
    <fj9>9-11 am</fj9>
    ...
  </vsfif>
</2nf3oiü*>
```

Wrap Up

- Knowledge Graphs
 - Facilitate syntactic and semantic data interoperability
- Today, we have seen syntactic interoperability
 - Unicode: a character set for all languages
 - URIs
 - Unique identifiers for *things* (entities, resources, ...)
 - On the Web, URLs are dereferencable
 - XML: a universal data exchange format
 - XPath
 - DTD
 - XML Schema

Data Interoperability with Knowledge Graphs



Questions?

