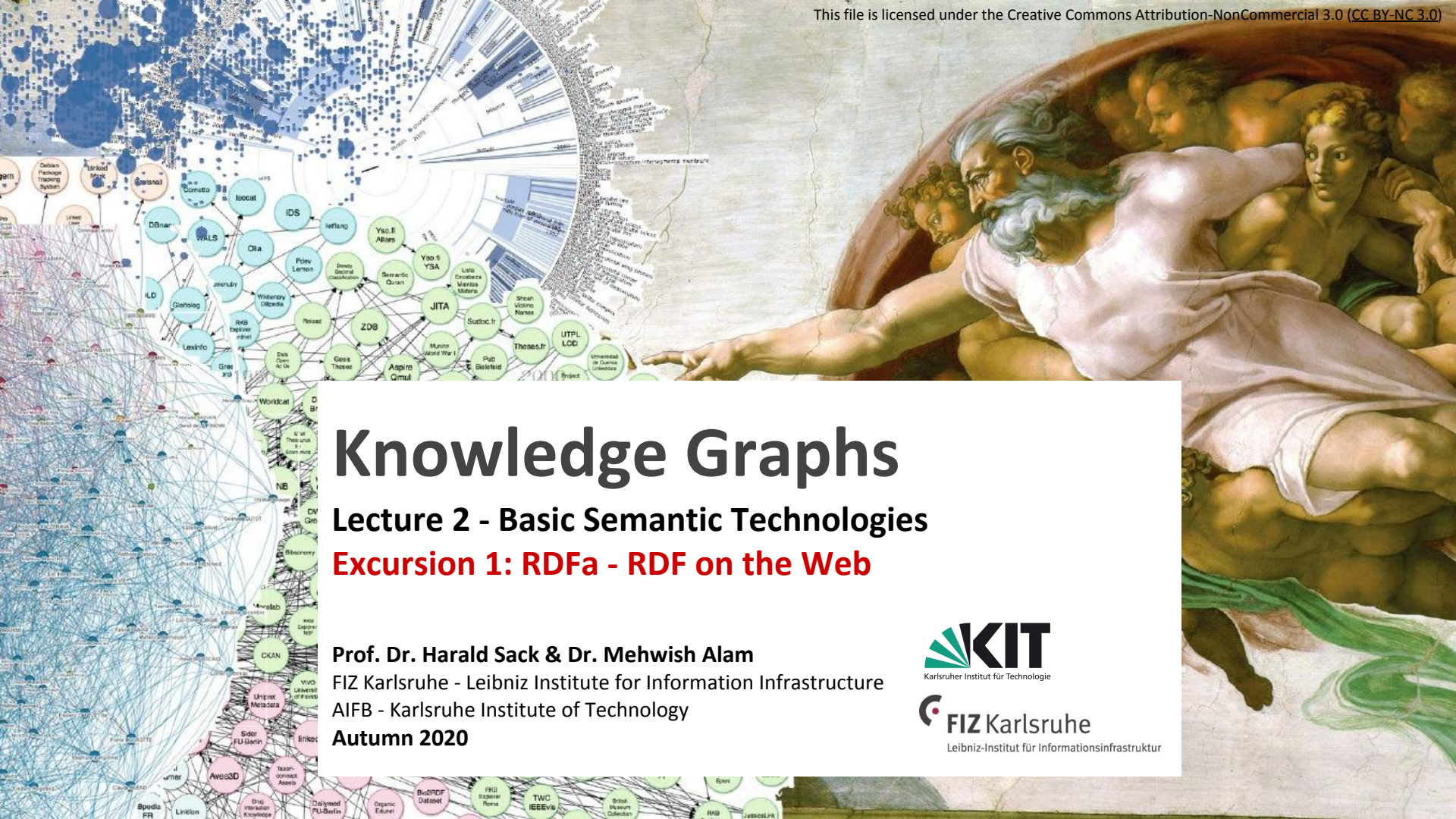


Knowledge Graphs

Lecture 2 - Basic Semantic Technologies

Excursion 1: RDFa - RDF on the Web

Prof. Dr. Harald Sack & Dr. Mehwish Alam
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure
AIFB - Karlsruhe Institute of Technology
Autumn 2020

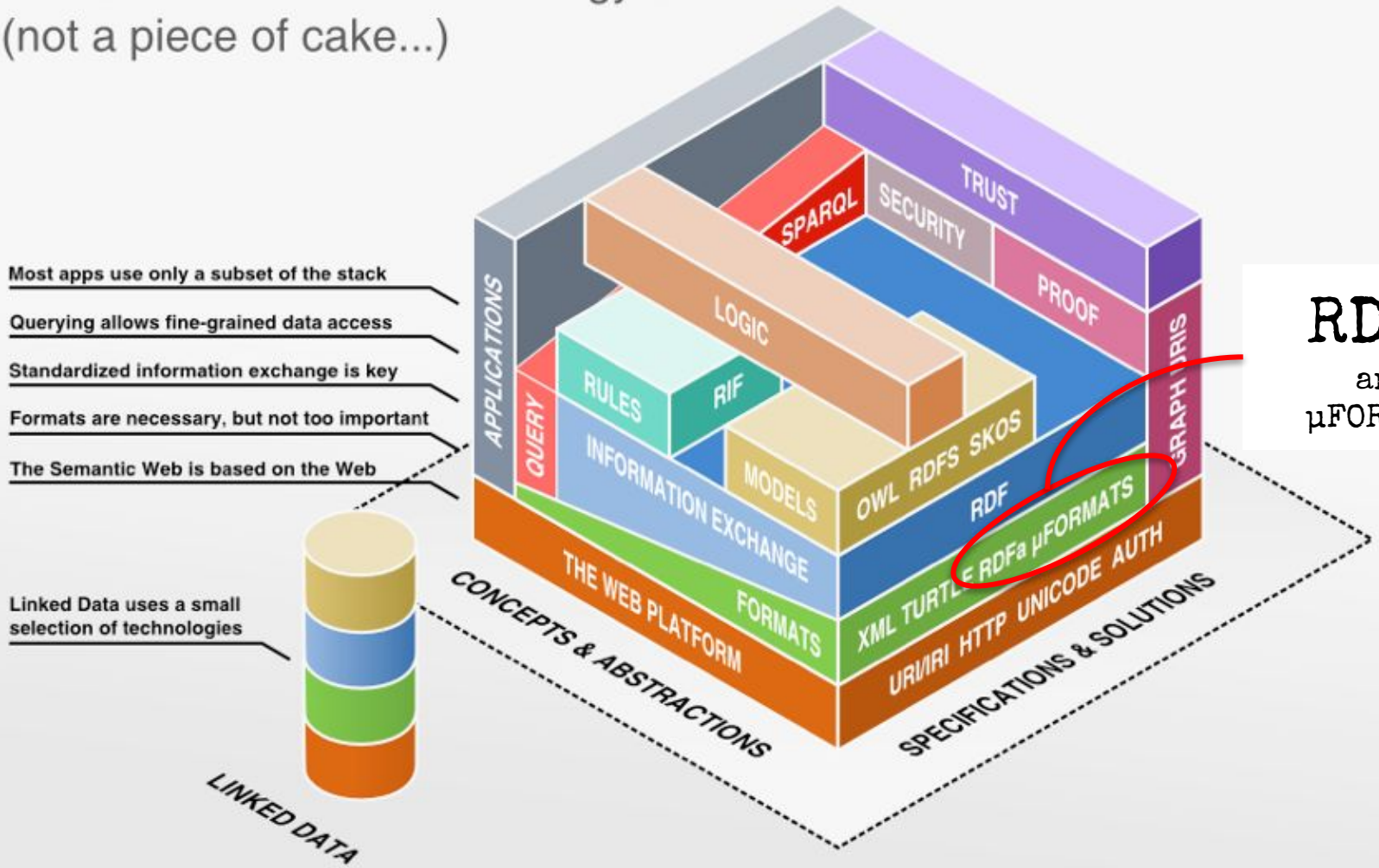
Karlsruher Institut für Technologie

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Knowledge Graphs

Lecture 2: Basic Semantic Technologies

The Semantic Web Technology Stack
(not a piece of cake...)



RDFa
and
µFORMATS

[1]

RDF(S) Semantics

- In principle there are three ways to embed structured data with explicit semantic annotations within HTML documents.


Domain specific microformats (µFormat)


Generic RDFa

HTML5 Microdata (including schema.org)

Microformats

- Microformats (µformats) emerged about 2005.
- (X)HTML Markup to express (limited) semantics in an HTML document
  - designed to solve simple, specific problems
  - designed for humans first, machines second
  - used in web pages to describe a specific type of information, as e.g. a person, an event, a product, a review, etc.
- Applications can easily extract data from HTML documents.
- In general, Microformats use the `class` attribute in **HTML tags** (most times `<span>` or `<div>` tags) and assign brief and descriptive names to entities and their properties.

Microformats

- Simple Example: HTML marked up with hCard microformat

```
<div class="vcard">
 <span class="fn">Joe Blow</span>
 <span class="title">Senior Manager</span>
 <span class="org">The Example Company</span>
 <span class="adr">
    <span class="street-address">Hauptstr. 123, </span>
    <span class="postal-code">14482</span>
    <span class="locality">Potsdam</span>
  </span>
 Tel.<span class="tel">604-555-1234</span>
 <img class="photo" src="www.example.com/JoeBlow.jpg" />
</div>
```

Microformats - Pros and Cons

- Microformats can **easily be transcoded to RDF** via XSLT.
- New microformat vocabularies **first must be consolidated** by the community, while a new XSLT stylesheet must always be developed for extraction.
- By **using more than one** microformat vocabulary in a single (X)HTML document the processing **complexity increases** rapidly.
- **Conflicts** with used (X)HTML attributes might be possible.

Embedding RDF in HTML Attributes

[2]

RDFa

- RDFa = **RDF** in HTML **a**ttributes
- enables generic RDF annotation in (X)HTML documents by reusing existing (X)HTML attributes.

- RDFa 1.0 based on XHTML (W3C Recommendation 2008)
- RDFa 1.1 based on HTML5 (W3C Recommendation June 2012)
  - **RDFa Lite 1.1**
  - RDFa 1.1

RDFa Lite 1.1

- RDFa reuses existing (X)HTML attributes (e.g. `href`, `src`) and introduces new HTML attributes
  - `vocab`,
  - `typeof`,
  - `property`,
  - `resource`,
  - `prefix`

RDFa Lite 1.1

- First we need a **vocabulary** to talk about things.

```
<p vocab="http://xmlns.com/foaf/0.1/">
   My name is Harald Sack and you can
   give me a ring via 1-800-555-0527.
</p>
```

RDFa Lite 1.1

- Then we have to define the **type of thing** we are talking about.

```
<p vocab="http://xmlns.com/foaf/0.1/"
   typeof="Person">
   My name is Harald Sack and you can
   give me a ring via 1-800-555-0527.
</p>
```

RDFa Lite 1.1

- Now we can define all **properties** of the thing we are talking about.

```
<p vocab="http://xmlns.com/foaf/0.1/" typeof="Person">
   My name is
   <span property="name">Harald Sack</span>
   and you can give me a ring via
   <span property="phone">1-800-555-0527</span>.
   <img property="img"
        src="http://fiz-karlsruhe.de/images/harald.jpg" />
</p>
```

RDFa Lite 1.1

- We can create **identifiers** for the things we are talking about.

```
<p vocab="http://xmlns.com/foaf/0.1/"
   resource="#harald" typeof="Person">
   My name is
   <span property="name">Harald Sack</span>
   and you can give me a ring via
   <span property="phone">1-800-555-0527</span>.
   <img property="img"
        src="http://fiz-karlsruhe.de/images/harald.jpg" />
</p>
```

- `resource` refers to the base URI of the web page.

RDFa Lite 1.1

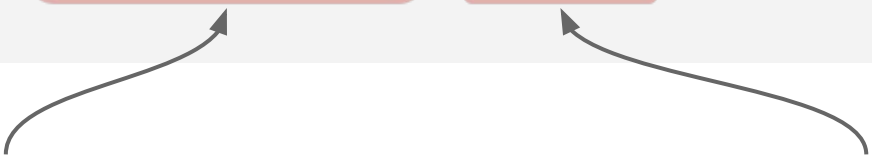
- And if the vocabulary is not sufficient to describe all properties, we can use additional vocabularies by using prefixes.

```
<p vocab="http://xmlns.com/foaf/0.1/"
   prefix="ov: http://open.vocab.org/terms/"
   resource="#harald" typeof="Person">
   My name is
   <span property="name">Harald Sack</span>
   and you can give me a ring via
   <span property="phone">1-800-555-0527</span>.
   My favorite beverage is
   <span property="ov:preferredBeverage">Espresso</span>.
</p>
```

RDFa 1.1

- With full RDFa 1.1 you can add additional functionality:
  - Separate **content** from presentation

```
<p vocab="http://purl.org/dc/terms/">
   <h2 property="title">Syllabus SWT 2012/13</h2>
   <h3 property="creator" resource="#me">Harald</h3>
   Creation Date:
   <span property="created" content="2012-10-28"> 28.10.2012 </span>
</p>
```



RDFa 1.1

- With full RDFa 1.1 you can add additional functionality:
  - Use **datatypes** from XML Schema Definition

```
<p vocab="http://purl.org/dc/terms/">
   <h2 property="title">Syllabus SWT 2012/13</h2>
   <h3 property="creator" resource="#me">Harald</h3>
   Creation Date:
   <span property="created" datatype="xsd:gYear"> 2012 </span>
</p>
```

RDFa 1.1

- Distinguish two different sorts of RDF triples:
  - Triple with **resource** as object
  - Triple with **literal** as object

	Subject	Property	Object
Object is **Literal**	`resource / about`	`property`	**`content`** or #PCDATA
Object is **Resource** (URI)	`resource / about`	`property/rel`	**`href`** or **`resource`**

- **`about`** / **`rel`** only for compatibility with RDFa 1.0

RDFa Tools

https://www.w3.org/2012/pyRdfa/

RDFa 1.1 Distiller and Parser

Warning: This version implements RDFa 1.1 Core, including the handling of the Role Attribute. The distiller can also run in XHTML+RDFa 1.0 mode (if the incoming XHTML content uses the RDFa 1.0 DTD and/or sets the `version` attribute). The package available for download, although it may be slightly out of sync with the code running this service.

Distill by URI | Distill by File Upload | Distill by Direct Text Input

Distill by URI

URI:

Output Format: Turtle

Returned content: Only core triples

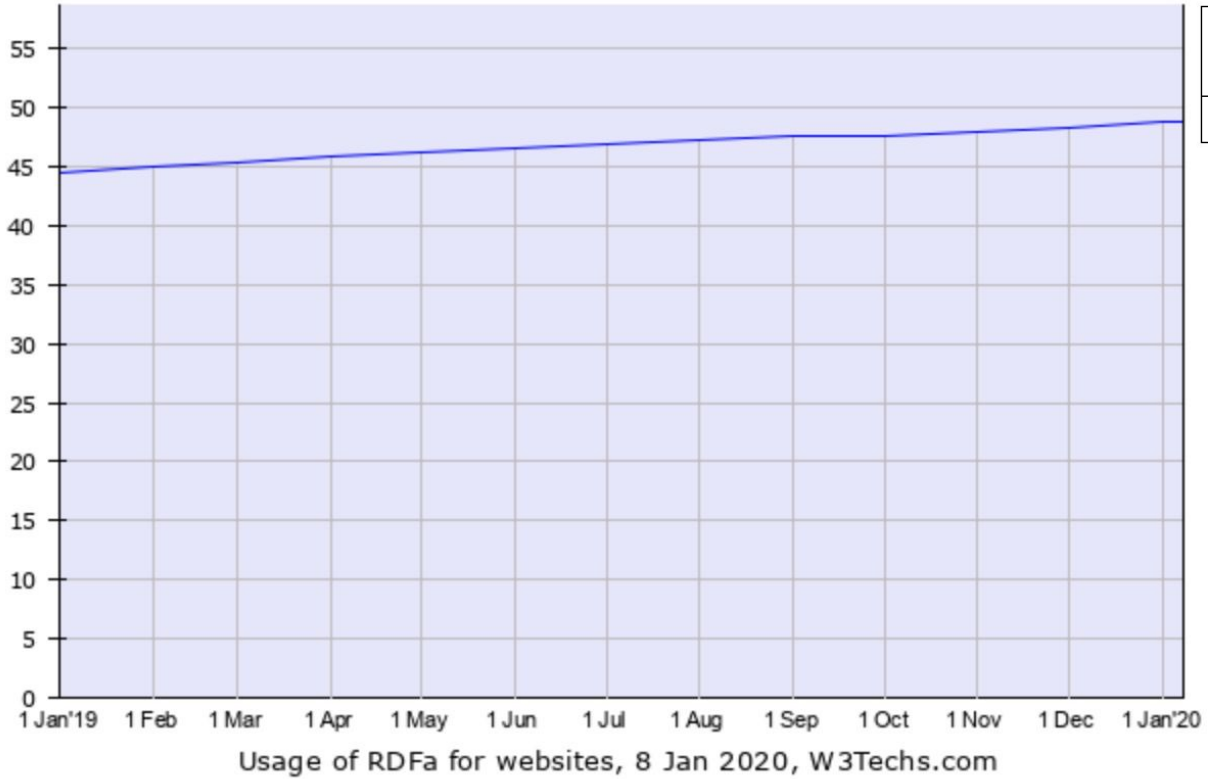
Expand vocabularies: No

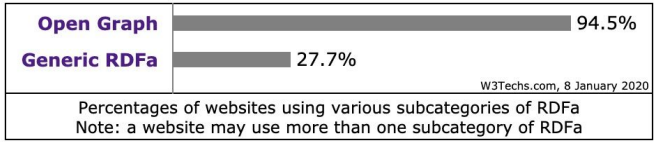
Generate warnings for non RDFa 1.1 Lite usage: No

▸ More (non-standard) options

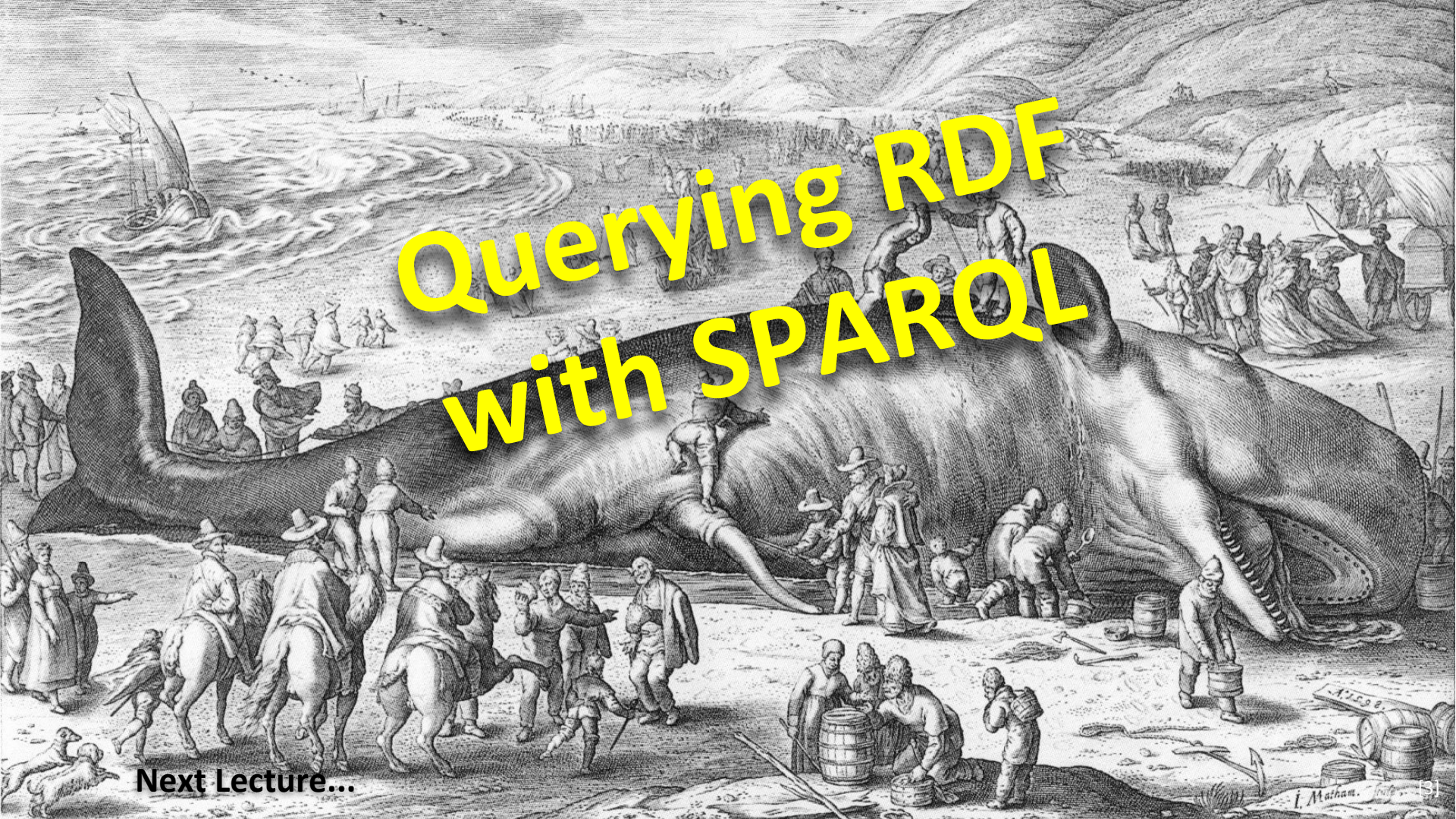
Go!

RDFa Usage





https://w3techs.com/technologies/details/da-rdfa


Querying RDF
with SPARQL
Next Lecture...

Picture References:

- [1] Benjamin Nowack, *The Semantic Web - Not a Piece of cake…*, at bnode.org, 2009-07-08 , [CC BY 3.0]
  http://bnode.org/blog/2009/07/08/the-semantic-web-not-a-piece-of-cake
- [2] Giuseppe Acrimboldo, Vertumnus, a portrait depicting Rudolf II, Holy Roman Emperor painted as Vertumnus, the Roman God of the seasons, c. 1590-1. Skokloster Castle, Sweden. [Public Domain]
  https://commons.wikimedia.org/wiki/File:Vertumnus_%C3%A5rstidernas_gud_m%C3%A5lad_av_Giuseppe_Arcimboldo_1591_-_Skoklosters_slott_-_91503.tiff
- [3] Hugo de Groot, Stranded Whale at Dutch sea coast in February 1598, [Public Domain]
  https://commons.wikimedia.org/wiki/File:Hugo-de-Groot-Nederlandtsche-jaerboeken_MG_0190.tif