Semantic Web Technologien für Fachinformation

Prof. Dr. York Sure-Vetter
28. Juni 2017
Agenda

- Beginners guide to the Web, the Semantic Web and the Social Web
- Semantics in practice: the Semantic CorA approach
Trivia

Obama has 60 Millions more followers than Trump.
Trump twitters

https://twitter.com/realdonaldtrump

http://www.trumptwitterarchive.com/

234 tweets with "loser"

I feel sorry for Rosie’s new partner in love whose parents are devastated at the thought of their daughter being with @Rosie—a true loser.

— Donald J. Trump (@realDonaldTrump) December 14, 2011

222 tweets with “dumb” or “dumby”

You must admit that Bryant Gumbel is one of the dumbest racists around - an arrogant dope with no talent. Failed at CBS etc why still on TV?

— Donald J. Trump (@realDonaldTrump) August 21, 2013

204 tweets with "terrible"

I loved beating these two terrible human beings. I would never recommend that anyone use her lawyer, he is a total loser!


183 tweets with "stupid"

@MichelleMalkin You were born stupid!


156 tweets with "weak"


## Obama twitters

### [https://twitter.com/barackobama](https://twitter.com/barackobama)

<table>
<thead>
<tr>
<th></th>
<th>Four more years.</th>
<th>@BarackObama[13]</th>
<th>0.937</th>
<th>November 6, 2012</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td><a href="https://twitter.com/bAJE6Vom">pic.twitter.com/bAJE6Vom</a></td>
<td><img src="https://en.wikipedia.org/wiki/List_of_most_retweeted_tweets_on_Twitter" alt="Barack Obama" /></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|   | Thank you for everything. My last ask is the same as my first. I'm asking you to believe—not in my ability to create change, but in yours. | @POTUS44\[15\] | 0.882 | January 10, 2017 | [H] |

[13]: https://twitter.com/barackobama
[15]: https://twitter.com/potus
2017: Frank-Walter Steinmeier on Facebook

https://www.facebook.com/steinmeierbundespraesident/
The Web

1989  Web Idea
1994  Amazon founded
1998  Google founded
2001  .com Bubble
2004  Facebook founded
2008  Obama becomes president due to Web strategy
2009  SAP defines Web strategy
2010  Instagram
2011  Google+
2012  Coursera
2017  Bundespräsident on FB

Advertising & Buy/sell

Live and Work
The inventor of the World Wide Web

Sir Tim Berners-Lee
(Turing Award winner 2017)
The first Website ... ever!

- The „first Website“, such as shown by Tim Berners-Lee at the Hypertext 91 surfaces conference:

  http://info.cern.ch/

- First Website:
  http://info.cern.ch/hypertext/WWW/TheProject.html
The first Website ... ever!

The World Wide Web project

WORLD WIDE WEB

The World Wide Web (W3) is a wide-area hypermedia[1] information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an executive summary[2] of the project, Mailing lists[3], Policy[4], November’s W3 news[5], Frequently Asked Questions[6].

What’s out there?[7] Pointers to the world’s online information, subjects[8], W3 servers[9], etc.

Help[10] on the browser you are using

Software A list of W3 project components and their current state. (e.g. Line Mode[12], X11 Viola[13], NeXTStep[14], Servers[15], Tools[16], Mail robot[17], Library[18])

Technical[19] Details of protocols, formats, program internals etc

<ref.number>, Back, <RETURN> for more, or Help:  
Computers – some 30 years back (1987)
Search (like it’s 1997!)

http://web.archive.org/web/19981111183552/google.stanford.edu/
Web 1.0: Vision of the Web

- Is a space in which

  - **Resources** are identified by Uniform Resource Identifiers (URIs)
  - **Protocols** support interaction between agents (HTTP)
  - **Formats** represent information resources (HTML)
Web 1.0: Uniform Resource Identifiers (URI)

- **Uniform Resource Identifier**
  - Resources may be anything that can be linked to or spoken of
    - Resources can be abstract or physical objects
    - Resources can contain a reference to another resource

- **Identifiable, but not necessarily retrievable**
  - (e.g. protected access)

- A single global system of identifiers
  - Each URI ideally identifies a single resource in a context-independent manner

- URIs act as names and addresses

- URIs require institutions
  - E.g. the registry that handles domain names
Web 1.0: Uniform Resource Identifiers (URI)

- Structure:
  - URI = scheme "::" hier-part [ "?" query ] [ "#" fragment ]

- Example:
  - foo://example.com:8042/over/there?name=ferret#nose

```
scheme    authority    path    query    fragment
```
Web 1.0: HTTP

- **HTTP**: A protocol that is basically stateless, a transaction consisting of:
  - **Connection**
    - The establishment of a connection by the client to the server - when using TCP/IP port 80 is the well-known port, but other non-reserved ports may be specified in the URL;
  - **Request**
    - The sending, by the client, of a request message to the server;
  - **Response**
    - The sending, by the server, of a response to the client;
  - **Close**
    - The closing of the connection by either both parties.
Web 1.0: HTML

- **HTML**: A representation format
  - Idea: Decoupling of content and representation
  - Cues for graphical presentation of content

```html
<p>Prof. Dr. York Sure-Vetter</p>
</tr>
<tr>
<td class="row-left">
<p><span class="ein-em">
Professor
</span><br />
Tel.: +49 (721) 608 44283
<br />
Email: york.sure-vetter@kit.edu</p>
<br />
<a href="/web/Wissensmanagement" title="Wissensmanagement">Wissensmanagement</a>
```
**Difference Web 1.0 and Web 2.0**

- **Web 1.0 focused on one-to-many information publishing**
  - Very static information

- **Web 2.0 focuses on many-to-many information publishing**
  - Collaborate and share Information like e.g. in wikis, facebook, blogs, etc.

![Web 1.0 vs Web 2.0 Diagram](image-url)
Web 2.0: e.g. Wikis

- Everyone can edit content on the Web
- ... without access to the server
Semantic Web: RDF and OWL

- Semantic descriptions of the content
- Machine readable and executable

Prof. Dr. York Sure-Vetter teaches Web Science (WS2017-18) is affiliated with KIT.
Semantic MediaWiki (SMW)

Max Völkel, Markus Krötzsch, Denny Vrandecic, Heiko Haller, Rudi Studer: Semantic Wikipedia. WWW 2006: 585-594
VRE Semantic CorA

- www.semantic-cora.org
- Partners: DIPF, KIT, Uni Göttingen

Main Objectives:
- Development of a Virtual Research Environment (VRE)
- Exemplary realization in the research field History of Education
- Enhancing quantitative + qualitative analysis
- Infrastructural distribution into other disciplinary contexts

Technological Platform:
- MediaWiki (MW) + extension Semantic MediaWiki (SMW)
Credits

Thanks to Dr. Basil Ell for the slides about Semantic CorA
Step 1: Importing Research Data

Scripta Paedagogica Online

http://bbf.dipf.de/retro-digit0.htm/digitale-bbf/scripta-paedagogica-online

25 lexica
22,000 articles

+ manual import of
62 undigitized lexica

Life-Cycles & Interactions

DL  VRE  VRE  BIBO  SKOS  FOAF  PRISM  DC

OAI
Step 2: Enriching Research Data

* Semantically annotate parts of images
* Create data about entities
* Link entities
* Open coding

Queryable network

Life-Cycles & Interactions

SW  VRE

Linked Open Data

http://linkeddata.org/static/images/lod-datasets_2009-07-14_cropped.png

Semantic network metadata, entities

Imported research objects, scanned pages
Step 3: Data Cleansing

- Imported data may contain errors
- Inform the provider & other consumers and provide patches

Life-Cycles & Interactions

DL → VRE

VRE → DL
Step 4: Exploring and Analyzing

Reference analysis

Visualization and exploration

Fine-grained image analysis
Step 5: Export and Sharing

- Selectively export content using imported vocabularies:
  - Corrected data
  - Annotations / Semantic network
  - Aggregated data
  - Evidence data related to research hypothesis
Next Steps

- Support **more resource types** besides lexica and images:
  - Video, audio, interview transcripts

- Vision:
  - future **ecologies** of small to large research projects and data providers
  - **data flows** between the participants
  - leads to **enriched and improved data**
  - providing benefits in a **multitude of collaborations**
Summary

Activities:
1. Importing Research Data
2. Enriching Research Data
3. Data Cleansing
4. Exploring and Analyzing
5. Export and Sharing

Life-Cycles & Interactions:
Trivia – revisited

Obama has 60 Millions more followers than Trump.

The World Record in Retweeting is currently being held by Carter Wilkerson.
2017: New World Record in Retweets

Top tweets

The following table lists the top ten most retweeted tweets on Twitter, the account that tweeted it, the total rounded to the nearest thousand, and the date it was originally tweeted. The notes include the details surrounding the tweet.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Tweet</th>
<th>Account name</th>
<th>Retweets (millions)</th>
<th>Date tweeted</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><strong>HELP ME PLEASE. A MAN NEEDS HIS NUGGS</strong></td>
<td>@carterjwm[2]</td>
<td>3.641</td>
<td>April 5, 2017</td>
<td>[A]</td>
</tr>
<tr>
<td></td>
<td><img src="https://en.wikipedia.org/wiki/List_of_most_retweeted_tweets_on_Twitter" alt="Image of chicken nuggets" /></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

“Carter Wilkerson tweeted [Wendy's](https://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet) asking how many retweets he needed for a year of free chicken nuggets. When Wendy's replied with "18 Million", he accepted the challenge. The tweet received unlikely support from several major companies including [Microsoft](https://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet), [Amazon](https://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet) and [Google](https://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet), prompting the hashtag #NuggsForCarter and propelling the tweet to become the most retweeted of all time within 34 days, at which point Wendy's gave him a year of free nuggets.”

See also: [http://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet](http://www.npr.org/sections/thetwo-way/2017/05/09/527597422/quest-for-free-chicken-nuggets-inspires-twitters-most-retweeted-tweet)
Vielen Dank für Ihre Aufmerksamkeit!