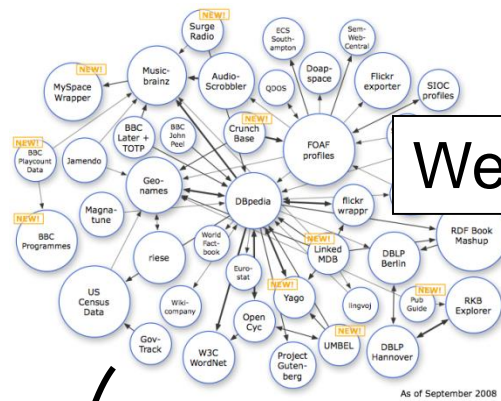


Semantic Web

Web of Data

Evolution of the Web: The Origins



Picture from [4]

Web of Data

Semantic Web

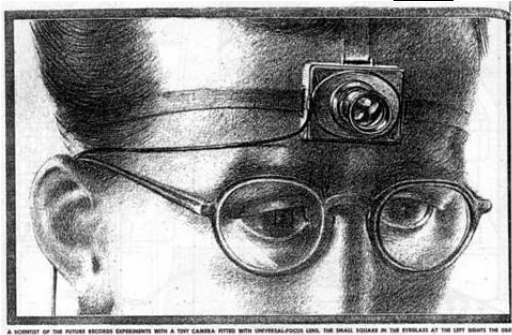


Web

Hypermedia

Semantic Annotations

Hypertext

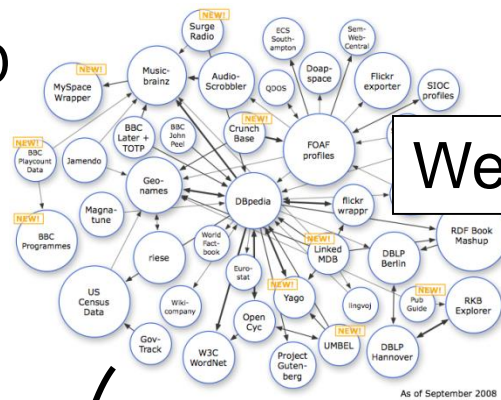


“As We May Think”, 1945

Picture from [3]

Evolutio

The Origins



Picture from [4]

Web of Data

Semantic Web

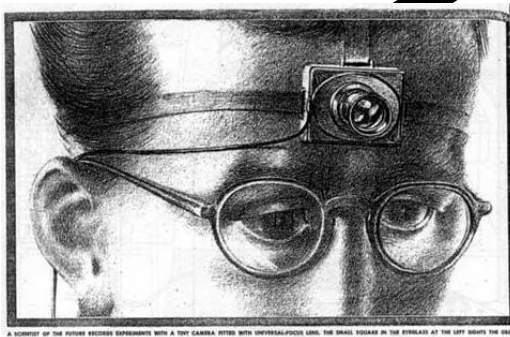


Web

Hypermedia

Hypertext

Soci (We Semantic Annotations



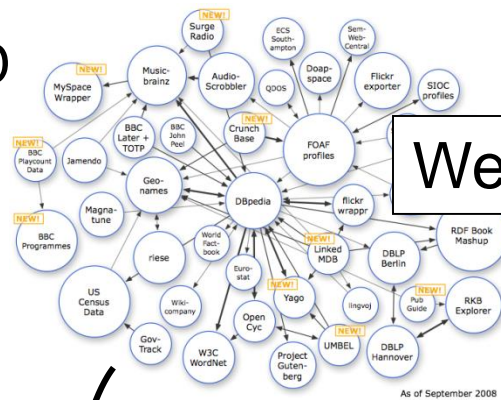
“As We May Think”, 1945

Picture from [3]

As We May Think (1945):

- Introduction of the Memex.
- Memex was envisioned to provide access to huge collections of text in which people could follow trails of links and notes.
- Memex is widely known as the precursor of the Hypertext movement.

Evolutio



Picture from [4]

Web of Data

Semantic Web



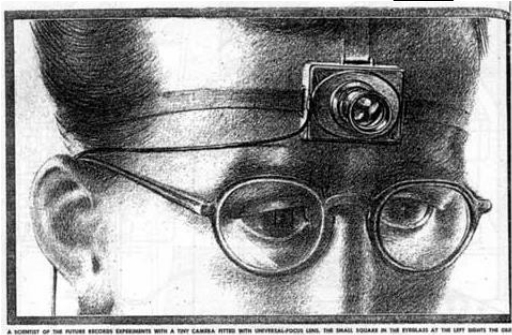
Semantic Annotations

Web

Hypermedia

Hypertext

?



“As We May Think”, 1945

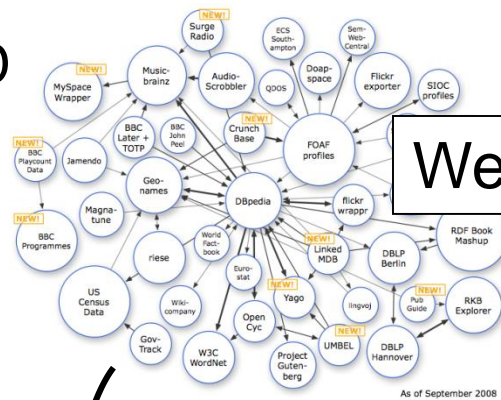
Picture from [3]

Hypertext:

- Term coined 1965 by Ted Nelson
- Definition: A hypertext is an organisation of objects in a highly connected fashion
- Characteristic elements: Nodes (e.g., text parts) and hyperlinks (logical connections between nodes)
- Further people: John Lickleder, Douglas Englbart

Evolutio

: Hypermedia



Picture from [4]

Web of Data

Semantic Web

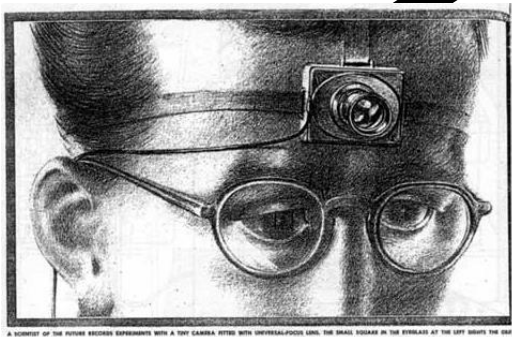


Web

Hypermedia

Hypertext

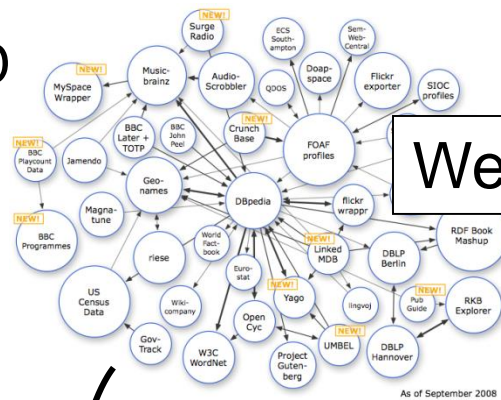
Semantic Annotations



“As We May Think”, 1945

Picture from [3]

Evolutio



Picture from [4]

Web of Data

Semantic Web



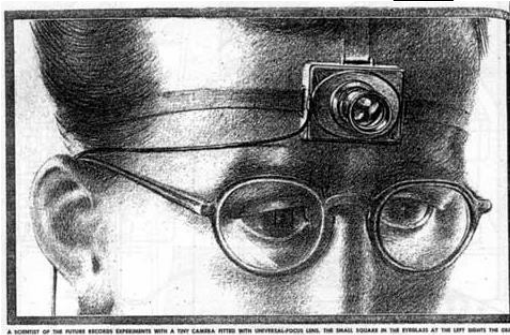
Semantic Annotations

Web

Hypermedia

Hypertext

?



“As We May Think”, 1945

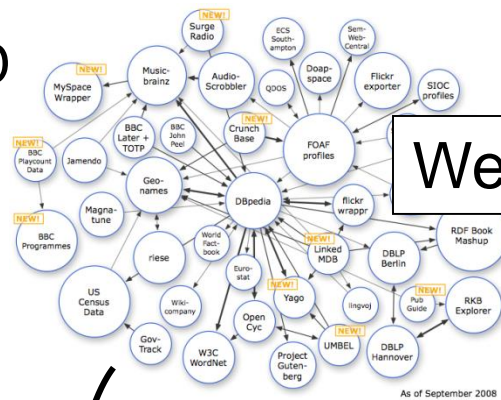
Picture from [3]

Hypermedia:

- Evolution of the hypertext idea
- Novelty: Multimedia aspects; i.e., multimedia resources might be part of interlinked structure

Evolutio

dia: the Web



Picture from [4]

Web of Data

Semantic Web



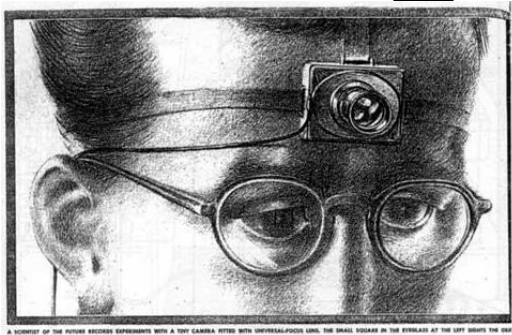
Semantic Annotations

Web

Hypermedia

Hypertext

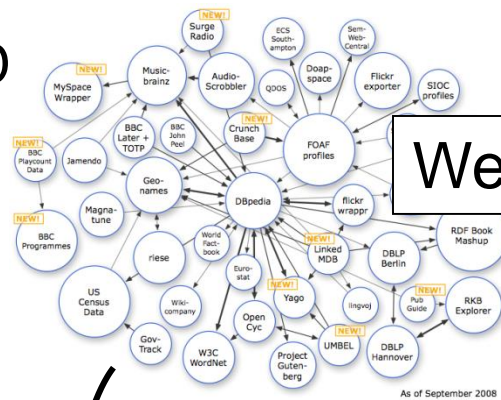
?



“As We May Think”, 1945

Picture from [3]

Evolutio



Picture from [4]

Web of Data

Semantic Web



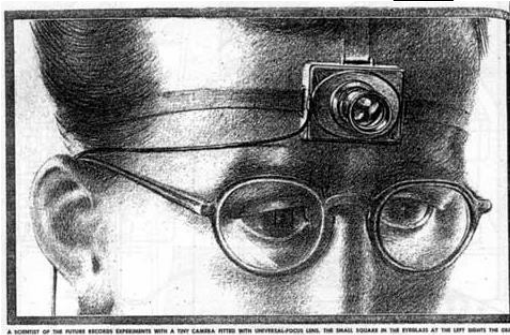
Semantic Annotations

Web

Hypermedia

Hypertext

?



“As We May Think”, 1945

Picture from [3]

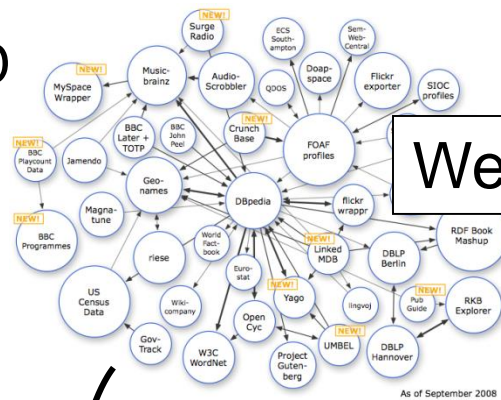
Web:

- Exemplary hypermedia system
- Proposed by Tim-Berners-Lee in 1990



Evolutio

The Semantic Web



Picture from [4]

Web of Data

Semantic Web

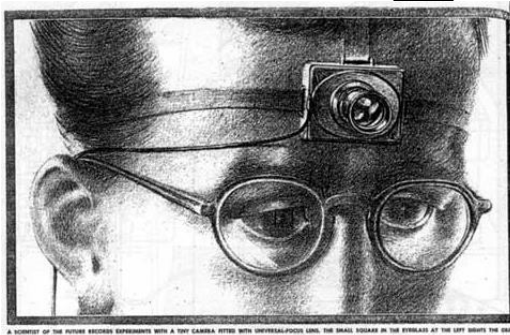


Web

Semantic Annotations

Hypermedia

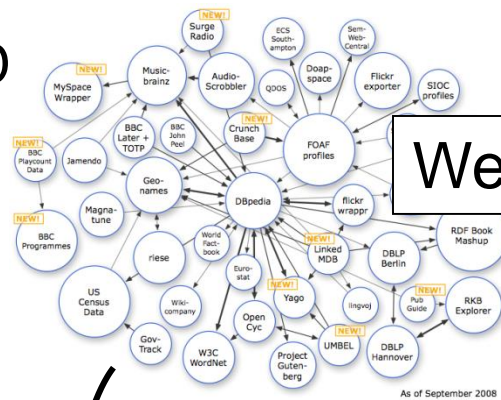
Hypertext



“As We May Think”, 1945

Picture from [3]

Evolutio



Picture from [4]

Web of Data

Semantic Web



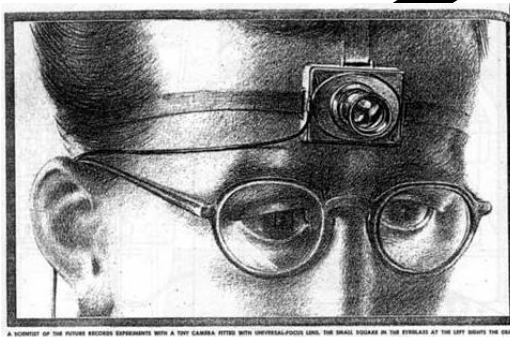
Semantic Annotations

Web

Hypermedia

Hypertext

?



“As We May Think”, 1945

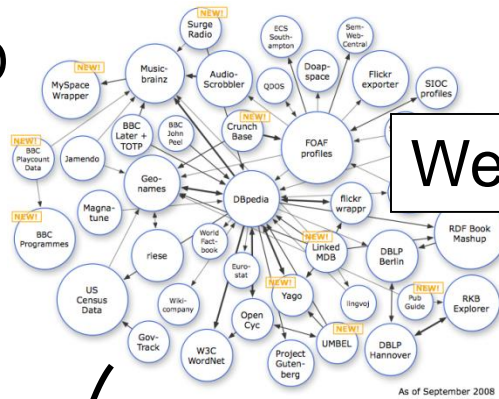
Picture from [3]

Semantic Web:

- Vision advocated by Tim Berners Lee.
- Contents have well-defined meaning.
- Backbone: formal ontologies allowing agents to draw automatic conclusions.

Evolutio

Web 2.0



Web of Data

Semantic Web



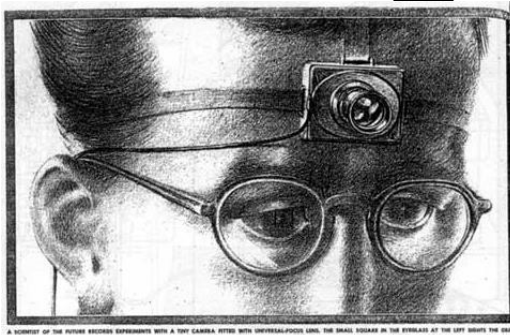
Semantic Annotations

Web

Hypermedia

Hypertext

?



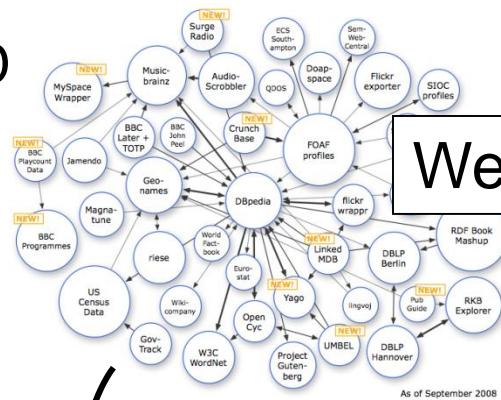
“As We May Think”, 1945

Picture from [3]

Picture from [4]

Evolutio

Semantic Annotations



Picture from [4]

Web of Data

Semantic Web



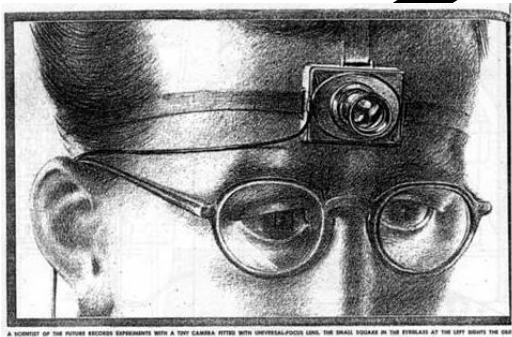
Semantic Annotations

Web

Hypermedia

Hypertext

?



“As We May Think”, 1945

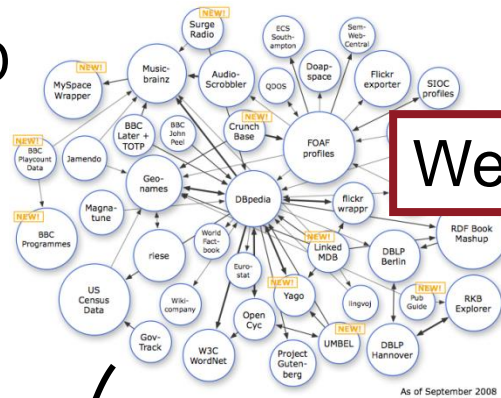
Picture from [3]

Semantic Annotations:

- Annotations are generated for the existing Web
- Generation automatic, semi-automatic, or manually based on human input
- See following lecture.

Evolutio

Web of Data



Web of Data

Semantic Web



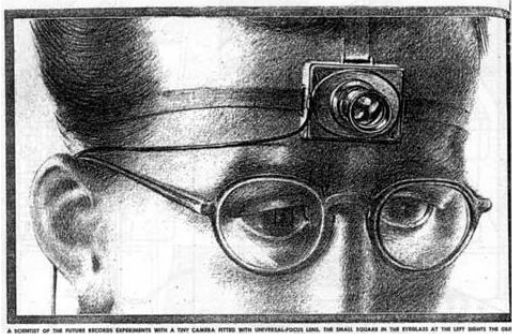
Semantic Annotations

Web

Hypermedia

Hypertext

?



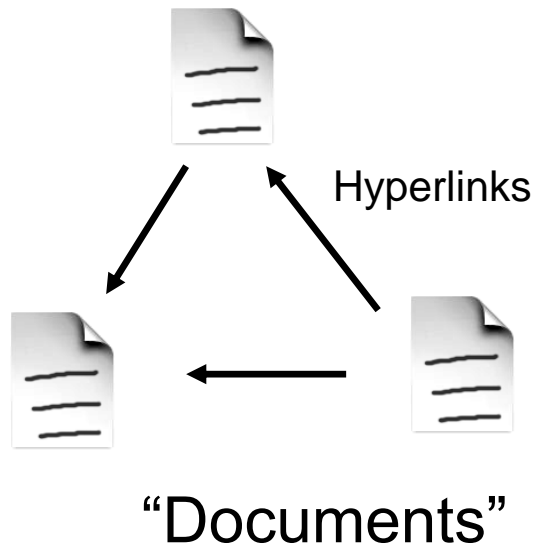
“As We May Think”, 1945

Picture from [3]

Picture from [4]

Motivation: From a Web of Documents to a Web of Data

- Web of Documents



- Fundamental elements:

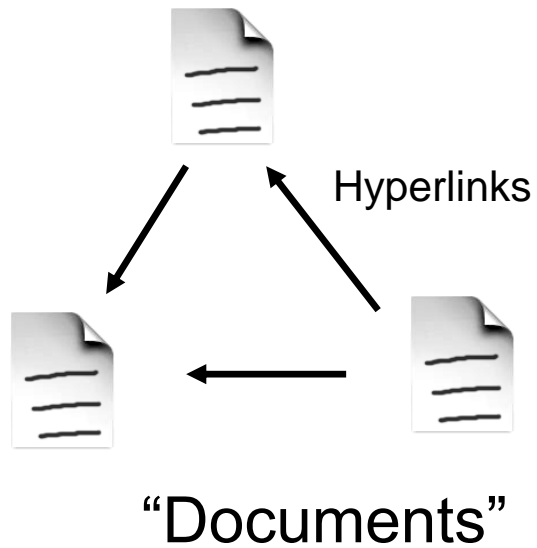
1. Names (URIs)
2. Documents (Resources) described by HTML, XML, etc.
3. Interactions via HTTP
4. (Hyper)Links between documents or anchors in these documents

- Shortcomings:

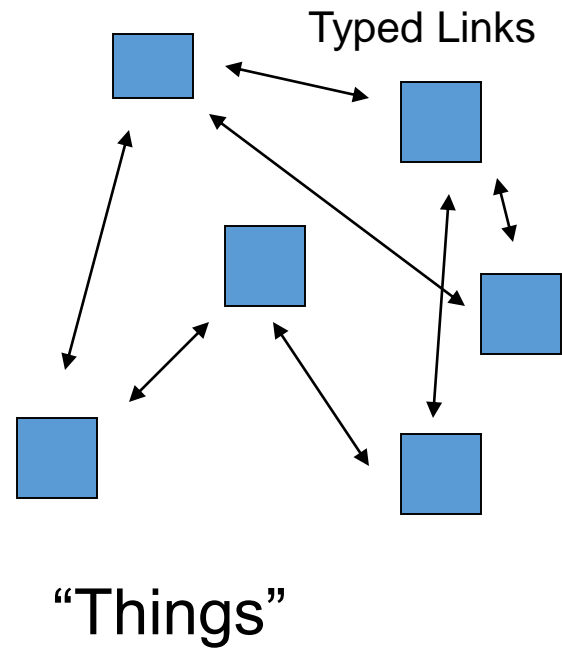
- Untyped links
- Web search engines fail on complex queries

Motivation: From a Web of Documents to a Web of Data

- Web of Documents



- Web of Data

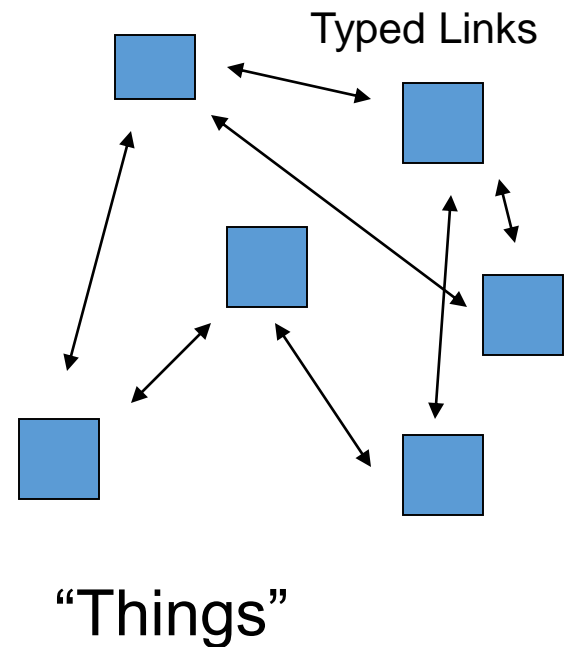


Motivation: From a Web of Documents to a Web of Data

- Characteristics:

- Links between arbitrary things (e.g., persons, locations, events, buildings)
- Structure of data on Web pages is made explicit
- Things described on Web pages are named and get URIs
- Links between things are made explicit and are typed

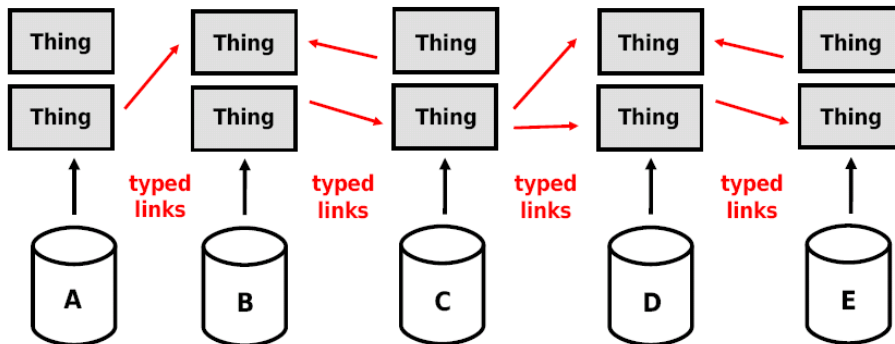
- Web of Data



Vision of the Web of Data

- The Web today

- Consists of data silos which can be accessed via specialized search engines in an isolated fashion.
- One site (data silo) has movies, the other reviews, again another actors.
- Many common things are represented in multiple data sets



- The Web of Data is envisioned as a global database

- consisting of objects and their descriptions
- in which objects are linked with each other
- with a high degree of object structure
- with explicit semantics for links and content
- which is designed for humans and machines

BUILDING THE WEB OF DATA BY PUBLISHING
STRUCTURED DATA ON THE WEB

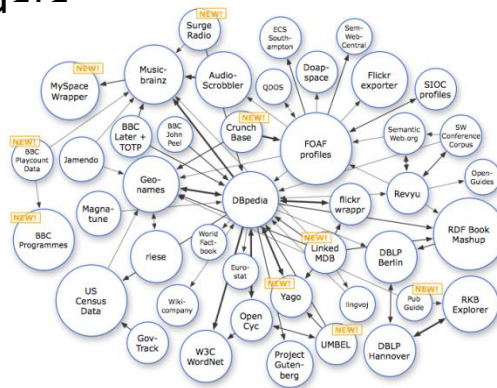
How to “Build” the Web of Data?

- Publish structured data by

1. using Web (2.0) APIs
(will be discussed in the Lecture on “Service Web”)
2. embedding structured information (Microformats, RDFa,
3. linking data



[5]



As of September 2008

[4]



[6]



[2]



[7]

2.1 EMBEDDING STRUCTURED INFORMATION IN WEB PAGES



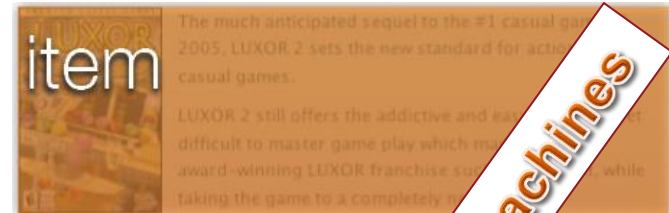
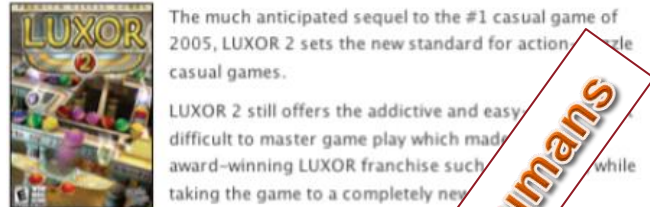
What are Microformats?

- An approach to add meaning to HTML elements and to make data structures in HTML pages explicit.
- “Designed for humans first and machines second, microformats are a set of simple, open data formats built upon existing and widely adopted standards. Instead of throwing away what works today, microformats intend to solve simpler problems first by adapting to current behaviours and usage patterns (e.g. XHTML, blogging).” [6]

What are Microformats? /2

- Are highly correlated with semantic (X)HTML / “Real world semantics” / “Lowercase Semantic Web” [9].
- Real world semantics (or the Lowercase Semantic Web) is based on three notions:
 - Adding of simple semantics with microformats (small pieces)
 - Adding semantics to the today’s Web instead of creating a new one (evolutionary not revolutionary)
 - Design for humans first and machines second (user centric design)
- A way to combine human with machine-readable information.
- Provide means to embed structured data in HTML pages.
- Build upon existing standards.
- Solve a single, specific problem (e.g. representation of geographical information, calendaring information, etc.).
- Provide an “API” for your website.
- Build on existing (X)HTML and reuse existing elements.
- Work in current browsers.
- Follow the DRY principle (“Don’t Repeat Yourself”).
- Compatible with the idea of the Web as a single information space.

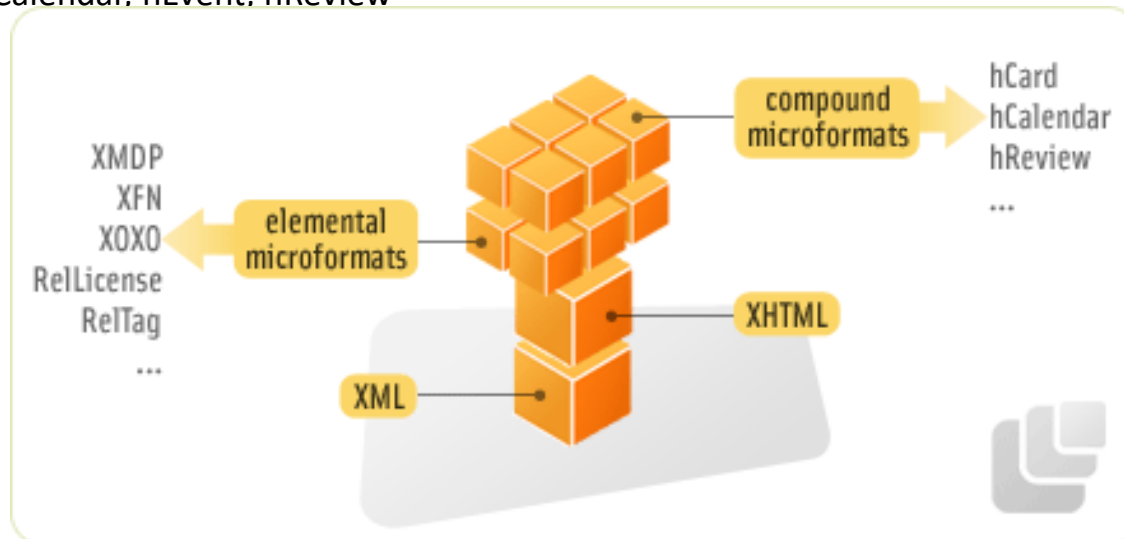
Microformats Illustrated



Example adapted from Chris Griego

Design Patterns

- Microformats can be seen as design patterns that make structure and semantics of data explicit.
- **Elemental microformats** (consist of just one tag)
 - Rel-home links to homepage `<link href="http://technorati.com" rel="home" />`
 - Rel-License links to content license `cc by2.0`
 - Others: rel-tag, rel-encluse, xfn-tags
- **Compound microformats** (more complex structures)
 - Often based on existing standard
 - E.g. hCard, hCalendar, hEvent, hReview



Picture from [6]

Syntax

- Microformats use existing HTML attributes to embed structured data types in an HTML document and to indicate the presence of metadata
- Rel/rev-attribute is used for elemental microformats, e.g.,
`semantics`
expresses that the current page is “tagged” with “semantics”
- Class-attribute is used for compound microformats, e.g.
`23.4444.33`
expresses that a given data block contains geo-coordinates (longitude/latitude)

Expressive Power

- Microformats extends the expressive power of HTML.
- Expressive power is limited as microformats are only designed to use pre-defined vocabularies to mark up content in Web pages using different HTML attributes.

Usage: Compound Microformat hCard

- hCard is a simple format for representing people, companies, organizations, and places, using a 1:1 representation of the properties and values of the vCard standard (RFC2426).

BEGIN: **VCARD**

VERSION: 3

FN: Dieter Fensel

ORG: STI Innsbruck

...

URL: <http://www.sti-innsbruck.at>

TEL: +43 512 507 9872

END: VCARD

Usage: Compound Microformat hCard: hCard /2

- hCard is a simple format for representing people, companies, organizations, and places, using a 1:1 representation of the properties and values of the vCard standard (RFC2426).

```
<div class="vcard">  
  <span class="fn">Dieter Fensel</span>  
  <a class="org url" href="http://www.sti-innsbruck.at">STI  
  Innsbruck</a>  
  <a class="email" href="mailto:dieter.fensel@sti2.at">mail me</a>  
  Phone: <div class="tel">+43 512 9872</div>  
</div>
```

Drawbacks of Microformats

- Only a fixed set of microformats exist.
- No way to connect data elements.
- Fixed vocabulary, not extendable and customizable.
- Separate parsing rules for each microformat needed.

Questions?

