# Open Data in HTML

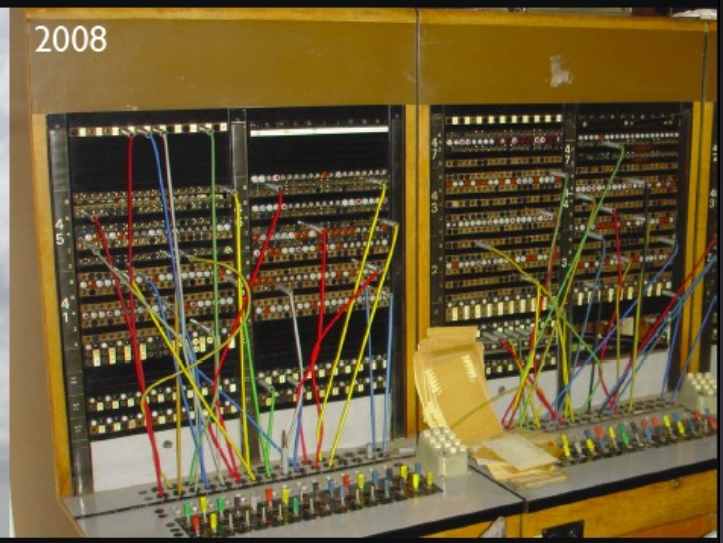
Elias Torres

IBM

*XTech 2007*

Alex Faaborg's browser timeline.






## Book

Bookmarks

Back and Next

You find the information

## Radio

Live Bookmarks

RSS Reader Integration

You subscribe to the information

## Switchboard

Microformat Detection

Application Integration

You move the information around

# Principles

- Metadata Vocabulary
  - Independence
  - Modularity/Reusability
  - Evolvability
- Don't Repeat Yourself (DRY)
- In-Context Metadata (Cut/Paste,DnD)

# Examples

- An event on a web page can be directly imported into a user's desktop calendar.
- A license on a document can be detected so that the user is informed of his rights automatically.
- A photo's creator, camera setting information, resolution, and topic stored side by side with the picture.

# Business Examples

- A web-based HR resume database.
- A customer (CRM) tracking online application that can expose customer information to mashup with online sales.
- Structured scientific annotations on web content that could be copy and pasted into personal notepad.
- Project information such as title, members, milestones, etc (DOAP).

# RDFa

```
<div>




licensed under a
<a
  href="http://cc.org/licenses/by/2.5/">
Creative Commons License</a>.
</div>
```

# RDFa

```
<div xmlns:cc="http://web.resource.org/cc/">

licensed under a
<a rel="cc:license"
   href="http://cc.org/licenses/by/2.5/">
Creative Commons License</a>.
</div>
```

Use @rel for URL properties

# RDFa

```
<div xmlns:cc="http://web.resource.org/cc/"
     xmlns:dc="http://purl.org/dc/1.1/">

This document was written by
<span property="dc:creator">Ben Adida</span> and is
licensed u der a
<a rel="
  href=
Creati
</div>
```

Use @property for literal properties

# RDFa

```
<div xmlns:cc="http://web.resource.org/cc/"
     xmlns:dc="http://purl.org/dc/1.1/"
     about="photo.jpg">
This photo was taken by
<span property="dc:creator">
licensed under a
<a rel="cc:license"
  href="http://cc.org/licenses/by/2.5/">
Creative Commons License</a>.
</div>
```

Use @about to change the subject of your property

# Different from Microformats

- If you needs go beyond micro-data (e.g. name, address, events) you can build your own vocabulary independently from the microformats.org community (e.g. proteins, geological data)

  and ...

- If you need to combine properties from multiple vocabularies (e.g. FOAF + vCard, Atom + vEvent, Resume + DOAP) you can avoid conflicts (e.g. class="name" and class="name" vs. rel="foaf:name" and rel="doap:name"

# More Differences

- RDFa's syntax specification can be used for many vocabularies. Therefore you only need to write the extraction mechanism and validator once.
- There's a precise way to identify objects in pages: @about.
- Proper use of xml:base, xml:lang and lang attributes in HTML.
- Ability to specify machine readable content and data types e.g. xsd:date.
- Flexible ways to indicate if the author intends or not to preserve markup in content e.g.

```
<div property=”dc:description”>
    E = mc<sub>2</sub></div>
```

# Latest with RDFa

- It's not only for XHTML2
- Coming: XHTML1.1 + Modularization

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN"
"http://www.w3.org/MarkUp/DTD/xhtml+rdfa-1.dtd">
```

Coming soon to a W3C Validator!

- Maybe someday: HTML5
- Updated RDFa Primer and Syntax Documents
- ODF 1.2 (metadata) re-using RDFa constructs

# Implementations

- PHP: RDFa Monkey.
- Java: RDFa Extractor, TopBraid, SweetWiki.
- Python: RDFLib, rdfadict.
- Ruby: ruby-rdfa.
- JavaScript, XSLT and more...

# Good News

- RDFa works in most HTML versions when using in-browser JavaScript to extract metadata.
- A simple and general RDFa parser can be written in less than 100 lines of JavaScript.

# FOAF Card


Ben Adida's Card - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Ben Adida, FOAF Card
This page contains my contact information and contacts.
Ben Adida (call me Ben)
homepage | photo
32 Vassar Street
MIT CSAIL Room 32G-694
Cambridge 02139
USA
People I know
Danny Weitzner
TimBL
Eric Miller

# What if we could...

```
var model = RDFa.parse(document);
var person = model.byType("foaf:Person");
with(person) {
    alert("Name:" + givenname + " " + family_name);
    alert("Homepage:" + homepage);
}
alert("Friends: " + person.knows.$all);

... person.office.address.street;
... person.office.address.city;
... person.office.address.postalCode;
```

# Future Scenarios

- Enhancing Internet/Intranet search engines to parse RDFa to support queries at the property level (e.g. doap:name = “queso”)
- Enhancing Atom Feeds with metadata

```
<entry xmlns="http://www.w3.org/2005/Atom">
   <title>Atom-Powered Robots Run Amok</title>
   <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
   <updated>2003-12-13T18:30:02Z</updated>
   <content type=”xhtml”>
    <div xmlns:cc="http://web.resource.org/cc/"> licensed under a
    <a rel=”cc:license” href=”http://cc.org/licenses/by/2.5/”>Creative
        Commons License</a>.
    </div>
   </content>
</entry>
```

- Business mashups, rich internet applications exchanging structured data.

# A geo Example

```
<h1 property="rdfs:label">Genova</h1>

<div xmlns:geo=
"http://www.w3.org/2003/01/geo/wgs84_pos#">

<span property="geo:lat">44.41667</span>,
   <span property="geo:long">8.94972</span>

</div>
```

# A vCard Example

```
<div xmlns:vcard="http://www.w3.org/2006/vcard/ns#"
  about="http://redmonk.com/who#james"
  class="vcard:VCard">

<h3 property="vcard:n rdfs:label">James
  Governor</h3>

Email: <a rel="vcard:email"
  href="mailto:jgovernor@redmonk.com">
  jgovernor@redmonk.com</a>

Phone: <span property="vcard:tel">+44 (0)20 7481
  4900</span>

IM: <span
  property="foaf:aimChatId">jjgatremonk</span> on
  AOL.

Blog: <a rel="vcard:url" href="
  http://www.monkchips.com/">MonkChips.com</a>

</div>
```

# Operator Action

```
Operator.actions.yahoo_maps_rdfa = {
      description: "Find with Yahoo! Maps (a)",
      icon: "http://www.yahoo.com/favicon.ico",
      scope: {
          semantic: {
            "RDFa" :  {
               property : "http://www.w3.org/2003/01/geo/wgs84_pos#lat",
               defaultNS  : "http://www.w3.org/2003/01/geo/wgs84_pos#"
             }
          }
      },
     doAction: function(geo, type) {
        var url;
        if (type != "RDFa") return;
        if (geo.lat && geo["long"])
         return "http://maps.yahoo.com/?lat="+geo.lat+"&lon="+geo["long"];
        }
    };
```

# Questions?

Thank You.

Excerpts from this presentation are based on the work of Ben Adida, Fabien Gandon, Alex Faaborg, W3C RDFa Task Force, IBM Lotus and Operator creator: Michael Kaply.