

## Microformats

Randy Schauer  
CMSC 691M

## What are Microformats?

- *“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 behaviors and usage patterns (e.g. XHTML, blogging).”*  
– Microformats.org

## Semantic Web vs. semantic web

	Semantic Web	semantic web
Philosophy	Build a common data format for expressing the meaning of data. Use ontologies to help machines to understand web content.	Humans first, machines second. Encode existing Web content with special tags.
Language	RDF, RDFS, OWL	Microformats
Format	Well formed RDF documents	Valid XHTML documents
Semantic	Defined by the underlying ontology model	Loosely defined. No formal semantic model.

From Harry Chen's Geospatial Semantic Web Presentation on 3/27/2007

## Advantages

- Encode explicit information to aid machine readability
- Uses standard XHTML with a set of common class-names. They are freely available for anyone to use.
- Simple to both describe and use
- Offers a way to go beyond just presentation in XHTML.

## Disadvantages

- Do not address implicit knowledge representation, ontological analysis or logical inference.
- Since there is no equivalent to namespaces, name conflicts are destined to occur with increasing frequency.
  - In order to try to reduce possible conflicts, there is a single repository for microformats; this is not scalable.
- Hard to validate microformats against plain CSS embedded in XHTML.

## Why use Microformats?

- In short, microformats are the convergence of a number of trends:
  - a logical next step in the evolution of web design and information architecture.
  - a way for people and organizations to publish richer information themselves, without having to rely upon centralized services.
  - an acknowledgment that (outside of specialist areas) "traditional" metadata efforts have either failed or taken so long to garner any adoption, that a new approach was necessary.
  - simple effort which has appealed to many frustrated with previous complex efforts.
  - a way to use (X)HTML for data.

## Microformat Specifications

- hCalendar
  - A calendaring and events format based on the iCalendar standard.
- hCard
  - A format for representing people, companies, places, and organizations using a 1:1 representation of the properties and values of the vCard standard.

## Microformat Specifications

- rel-license
  - A format for indicating content licenses in hyperlinks.
- rel-nofollow
  - A format for indicating that the associated hyperlink should not receive additional weight or ranking by user agents which perform link analysis upon web pages (e.g. search engines).
- rel-tag
  - A format for indicating that the destination of that hyperlink is an author-designated "tag" for the current page.

## Microformat Specifications

- **VoteLinks**
  - A format expressing three new values for the rev attribute of the <a> (hyperlink) tag in HTML:
    - vote-for
    - vote-abstain
    - vote-against
- **XFN (XHTML Friends Network)**
  - A simple way to represent human relationships using hyperlinks.
- **XMDP (XHTML Meta-Data Profiles)**
  - A format for defining HTML meta data profiles.
- **XOXO (Extensible Open XHTML Outlines)**
  - A simple, open outline format.

## Microformat Drafts

- **adr**
  - A format for marking up address information.
- **geo**
  - A format for marking up WGS84 geographic coordinates (latitude; longitude).
- **hAtom**
  - A format for content that can be syndicated, primarily but not exclusively weblog postings.
- **hResume**
  - A format for publishing resumes and CVs.
- **hReview**
  - A format for embedding reviews of products, services, businesses, events, etc.

## Microformat Drafts

- **rel-directory**
  - A format for indicating that the destination of the hyperlink is a directory listing containing an entry for the current page.
- **rel-enclosure**
  - A format for indicating files to cache.
- **rel-home**
  - A format for indicating that the destination of that hyperlink is the homepage of the current site.
- **rel-payment**
  - A format for indicating that the destination of that hyperlink provides a way to show/give support for the current page.

## Microformat Drafts

- **Robots Exclusion**
  - The Robot Exclusion Profile is a reworking of the Robots META tag (and less-standard extensions) as a microformat.
- **xFolk**
  - A format for publishing collections of bookmarks.

## Pre-Microformat XHTML

`<a href="/myEvent.vcs">`

`<b>Microformats: What the Hell Are They and Why Should I Care?</b>`

`<p>Ryan King will explain why microformats are important and how you can mark up specific kinds of content in ways that make it easier for the right people to find your stuff.</p>`

`<small>September 25th, 2005, 5-6PM in the`

`<i>Balder Room</i></small>`

`</a>`

## XHTML with Microformat tags

`<div class="vcalendar vevent">`

`<span class="summary">Microformats: What the Hell Are They and Why Should I Care?</span>`

`<p class="description">Ryan King will explain why microformats are important and how you can mark up specific kinds of content in ways that make it easier for the right people to find your stuff.</p>`

`<abbr class="dtstart" title="20050926T050000-0700">September 25th, 2005, 5</abbr>—`

`<abbr class="dtend" title="20050926T060000-0700">6PM</abbr> in the`

`<span class="location">Balder Room</span>`

`</div>`

## Tagging

- *"If one Web site links to another, the link doesn't carry any information about why the sites are linked. But what if it did?" – Tantek Celik*
- Technorati tracks the most successful form of microformats for humand-generated content... tagging.
- Tags on Technorati-tracked blogs grew from 0 to 20 million between January and June 2005.

■ Animation from CMU Art & Computer Science Research Group

## Technorati State Of The Blogosphere - Q4 2006

- The blogosphere is now 70 million weblogs wide.
- 1.5 million posts per day.
- Tracking 230 million posts with tags or categories.
- 35% of all February 2007 posts used tags.
- 2.5 million blogs posted at least one tagged post in February.

## Native Firefox Support Coming

- Native handling of Microformats expected in Firefox 3
  - <http://ejohn.org/blog/microformats-in-firefox-3/>
- Items Mozilla believes are still needed:
  - A standard for parsing Microformats had to be clearly defined.
  - An excellent implementation of that standard needed to be implemented.
  - And an important player needed to adopt the use of that tool.