

---

# **Understanding the web: publishing and processing microformats**

**Anca Paula Luca  
Summer <Web /> 2007**

---

---

# **1. Microformats definition**

---

# 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.

or, more simple...

- Simple conventions for embedding semantics in HTML to enable decentralized development.
  - emerged from patterns of correspondence between semantics and markup.

# Now, seriously, what are microformats?

---

```
<li class="vevent ev-item">
```

```
  <a class="url summary"
  href="http://microformats.org/wiki/events/2007-05-16-XTech">
```

```
    Microformats: The Nanotechnology of the Semantic Web
```

```
  </a><div>
```

```
    <abbr class="dtstart" title="2007-05-16T14:00:00+0100"> May
    16th, 2:00 </abbr> –
```

```
    <abbr class="dtend" title="2007-05-16T14:45:00+0100">
    2:45pm </abbr>
```

```
    <div class="location">Novotel Paris Tour Eiffel, Paris</div>
```

```
  </div>
```

```
</li>
```

# Now, seriously, what are microformats?

That can be styled to look like this:

Howard  
microformats

ook from  
resentation  
ne is  
it  
ations

## Upcoming events

[Subscribe...](#) ([More events on the wiki](#))

### [Microformats Skillswap](#)

May 10th, 6pm–9pm

28 Kensington Street, Brighton

### [Microformats: The Nanotechnology of the Semantic Web](#)

May 16th, 2:00–2:45pm

Novotel Paris Tour Eiffel, Paris

# Why microformats?

---

- Why semantics?
  - to understand each other
    - e. g. an address format.
- Why semantics embedded in (X)HTML?
  - “lower case semantic web” aka “real world semantics”
    - no theoretical discussions and correctness proofs
    - is not “yet another technology to learn”
    - easy to understand and use by anyone familiar with (X)HTML
  - evolutionary not **revolutionary**
    - build on existing technologies instead of inventing new ones

# The microformats principles

---

- solve a specific problem
- start as simple as possible
- design for humans first, machines second
- reuse building blocks from widely adopted standards
- modularity / embeddability, loose connections
- enable and encourage decentralized development, content, services

# Microformats examples (1)

---

- rel-tag

```
<a href="/tag/web20/" rel="tag" class="category">web2.0</a>
```

- geo

```
<span class='geo' style='display:none'>  
  <span class='latitude'>55.764792</span>  
  <span class='longitude'>-4.853038</span>  
</span>
```

- XFN (XHTML Friends Network)

- <a href="http://scobleizer.wordpress.com/" rel="met acquaintance colleague">Robert Scoble</a>
- <a href="http://theyanking.com/" rel="friend colleague co-worker met">Ryan King</a>

# Microformats examples (2)

---

- hCard

```
<div class="vcard nosave" >
  
  <strong class="fn">Anca Luca</strong>
  <span class="adr">
    <a class="country-name"
href="/place/Romania">Romania</a>
  </span>
  <a href="http://ancaluca.blogspot.com" class="url">
ancaluca.blogspot.com </a>
</div>
```

- ... more at <http://microformats.org>

---

# **2. Publishing microformats**

---

# Publishing semantics

---

- Why publish microformats?
  - because we want to understand each other
- When to publish microformats?
  - when we find ourselves using a pattern
  - when we have suitable data
- When NOT to publish microformats?
  - when we only need a “buzzword” - use another!
  - when we don't have suitable data

# Publishing for parsing

---

*"Always code as if the person who will maintain your code is a maniac serial killer who knows where you live!"*

*Microsoft Visual Studio's Tip of the day*

- Keep it as simple as possible
- Don't "adapt" semantics to fit your information
- Don't interfere with microformats markup
- In social websites, trust the user
- Read the parsing guidelines, too!

(More at: <http://microformats.org/wiki/implementation-guidelines>)

# Examples: dos

---

- Cork'd (<http://corkd.com>)
  - modeled entire website in terms of hCards and hReviews
- Last.fm (<http://www.last.fm/>)
  - hCards for users, hCalendars for events, friend messages marked with hAtom
- Revyu (<http://revyu.com/>)
  - reviews well marked with hReview

# Examples: don'ts

---

- Yahoo! Tech (<http://tech.yahoo.com/>)
  - review details page not marked as hReview
- old Eventful (<http://eventful.com/>)
  - user-added links for an event didn't get a class="url"
- old Upcoming (<http://upcoming.yahoo.com/>)
  - all attending users got a class="url", interfering with the event url
- LinkedIn (<http://www.linkedin.com/>)
  - all contacts have a default, non-changeable rel="contact"

---

# **3. Parsing microformats**

---

# Do we have to?

---

- No, but it would be cool
- Semantics can help automate some human activities:
  - detecting information
  - using information
- Microformats are a good, wide spread source for semantics.
- Because semantics are specified with the standards, human assistance is minimal.

# Using microformats

---

- Collect data
  - e.g. get .vcf and .ics files from hCards and hCalendars
  - examples: Tails Export extension
- Access services
  - e.g. show an adr marked item on a map, using map services
  - examples: Operator extension (maps, tags, calendars)
- Intelligent agent
  - e.g. infer a friends graph from XFN relations

# Scenario (1)

---

- Suppose you use an intelligent agent to whom you only tell your homepage
- You access your homepage
  - the agent learns other data from the XFN owner marked hCard
  - the agent learns the XFN marked friend list
- You access a friend's homepage
  - the agent learns his data, his friend list, and some events he has published with hCalendar

# Scenario (2)

---

- You access an event he's going to: a concert
  - the agent learns more data about the event
  - “More of your friends are going! Wanna go, too?”
- “Hmm, yeah, cool!”
  - the agent shows you the event location (from “geo” or “adr” data) on a map
  - “It's pretty far, would you like to go there by car or by plane?”
- “By car”
  - using event data and your location data, the agent prints you nice directions to the place, from a map

# Scenario (3)

---

- "It's a two days event, and you have some friends who live in the city, will you stay with them, or I should look for a hotel?"
- "Who? Those guys? Nooooo, look for a hotel!"
  - the agent uses hotel reservation services to get a hotel reservation two only streets far from the concert location
- "Ok, thank you! Now could you store this in my online calendar?"
  - "Yes, sir!"

---

# **4. Conclusions**

---

# Conclusions

---

- Microformats are a method to publish semantically marked data
- Use them properly to make information widely accessible
- Process them to build semantics-aware, intelligent agents

---

**Thank you!**

---

---

**Questions?**

---