



[Lady Gaga | Free Music, Tour Dates, Photos, Videos](#) +1

www.myspace.com/ladygaga - Cached

Lady Gaga's official profile including the latest music, albums, songs, music videos and more updates.

Judas	4:10	Judas
The Edge Of Glory	5:21	Born This Way
Born This Way	4:20	Born This Way
You And I	5:07	Born This Way

The Semantic Web

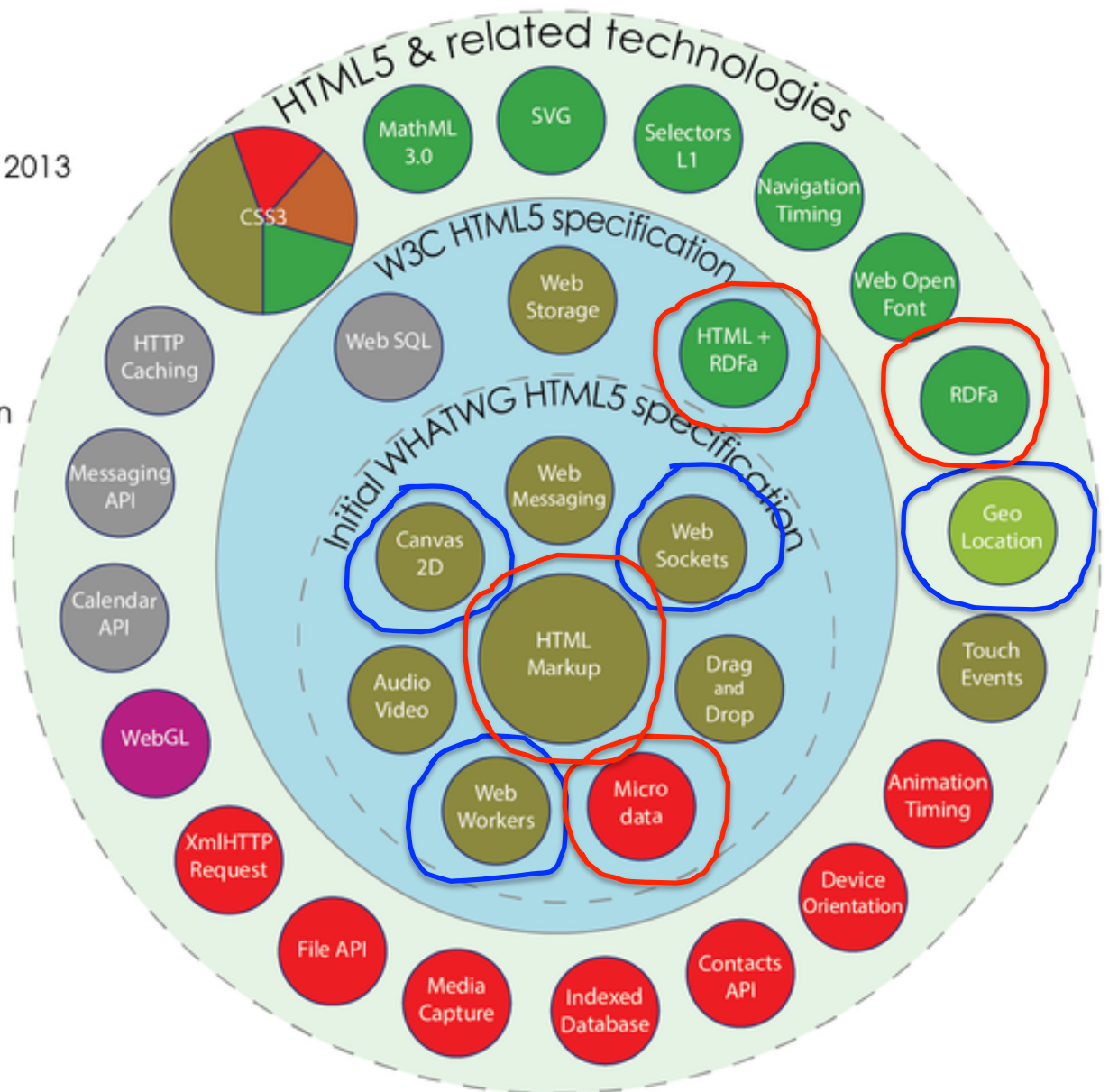
Overview

- HTML5 semantic markup
 - Why?
 - Common tags
- Custom markup
 - Microdata
 - RFDa
 - Microformats

HTML5

Taxonomy & Status on January 20, 2013

- W3C Recommendation
- Proposed Recommendation
- Candidate Recommendation
- Last Call
- Working Draft
- Non-W3C Specifications
- Deprecated



by Sergey Mavrody (CC) BY · SA

Semantic web: why?

- **The problem:**
 - Web pages are hard for computer to parse
 - Everything in a `<div>` with a `id/class`
 - Is the `<div>` with ID "menu" a navigation sidebar or a menu of a restaurant?
 - What part of the page is:
 - A blog post? The date of the post?
 - Header of the web site? Footer?
 - Figure? Caption of that figure?
 - Web pages can be hard for even humans to use
 - e.g. visually-impaired users, using only voice feedback
 - e.g. motor-impaired users, using only a few buttons

Semantic web

- What is it?

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation"

– May 2001, Scientific American, Tim Berners-Lee

"The Semantic Web is a vision: the idea of having data on the Web defined and linked in a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications."

– W3C 2001



New HTML5 semantic tags

Tag	Description
<article>	Defines an article
<aside>	Defines content aside from the page content
<bdi>	Isolates a part of text that might be formatted in a different direction from other text outside it
<command>	Defines a command button that a user can invoke
<details>	Defines additional details that the user can view or hide
<summary>	Defines a visible heading for a <details> element
<figure>	Specifies self-contained content, like illustrations, diagrams, photos, code listings, etc.
<figcaption>	Defines a caption for a <figure> element
<footer>	Defines a footer for a document or section
<header>	Defines a header for a document or section
<hgroup>	Groups a set of <h1> to <h6> elements when a heading has multiple levels
<mark>	Defines marked/highlighted text
<meter>	Defines a scalar measurement within a known range (a gauge)
<nav>	Defines navigation links
<progress>	Represents the progress of a task
<ruby>	Defines a ruby annotation (for East Asian typography)
<rt>	Defines an explanation/pronunciation of characters (for East Asian typography)
<rp>	Defines what to show in browsers that do not support ruby annotations
<section>	Defines a section in a document
<time>	Defines a date/time
<wbr>	Defines a possible line-break



Figure 1 - General browser layout for fictitious home page

<http://msdn.microsoft.com/en-us/scriptjunkie/gg454786>

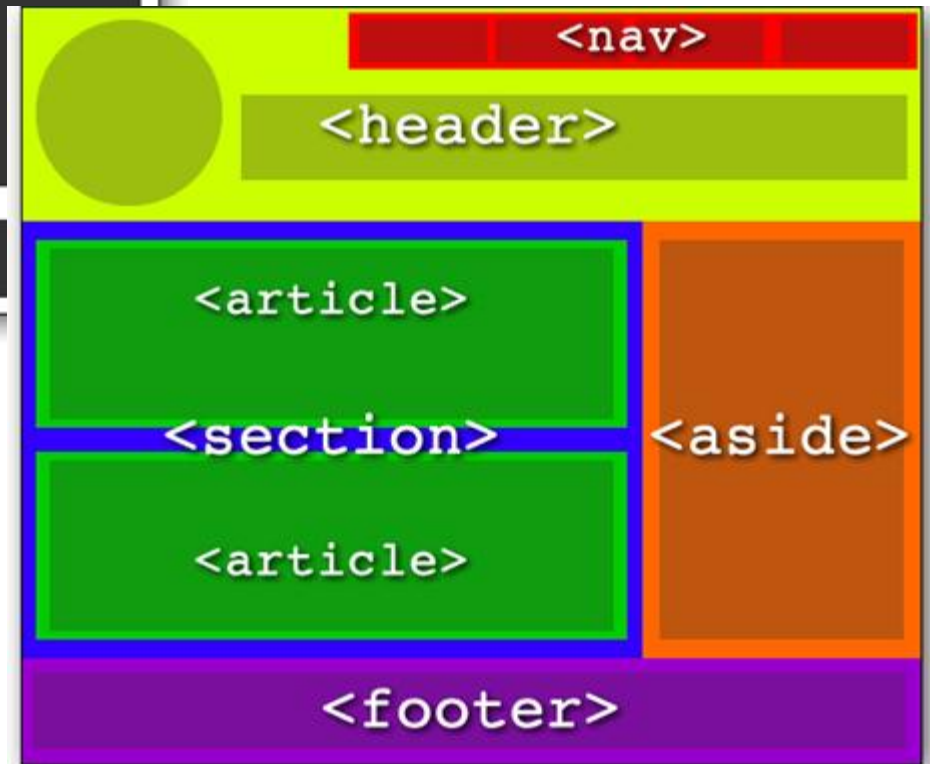


Figure 2 - New HTML5 semantic elements applied to layout for fictitious site

Semantic element support

New semantic elements - Working Draft

HTML5 offers some new elements, primarily for semantic purposes. The elements include: section, article, aside, hgroup, header, footer, nav, figure, figcaption, time, mark.

*Usage stats:	Global
Support:	82.36%
Partial support:	3.14%
Total:	85.5%

Show all versions	IE	Firefox	Chrome	Safari	Opera	iOS Safari	Opera Mini	Android Browser	Blackberry Browser
								2.1	
								2.2	
						3.2		2.3	
						4.0-4.1		3.0	
	8.0					4.2-4.3		4.0	
	9.0	18.0	24.0	5.1		5.0-5.1		4.1	
Current	10.0	19.0	25.0	6.0	12.1	6.0	5.0-7.0	4.2	7.0
Near future		20.0	26.0		12.5				10.0
Farther future		21.0	27.0						

Notes [Known issues \(0\)](#) [Resources \(5\)](#) [Feedback](#)

[Edit on GitHub](#)

Partial support refers to missing the default styling. This is easily taken care of by using display:block for all new elements (except time and mark, these should be display:inline anyway).

<http://caniuse.com/#feat=html5semantic>

<section>

- **According to W3C:**
 - "represents a generic section of a document"
 - "a thematic grouping of content, typically with a heading"
- **Not a generic container for styling**
 - That's the job of <div>
 - Avoid if <article>, <aside>, <nav> more appropriate
- **Examples:**
 - Chapters
 - Numbered section of a thesis

<section> example

```
<article>
  <hgroup>
    <h1>Apples</h1>
    <h2>Tasty, delicious fruit!</h2>
  </hgroup>
  <p>The apple is the pomaceous fruit of the apple tree.</p>
  <section>
    <h1>Red Delicious</h1>
    <p>These bright red apples are the most common found in many
    supermarkets.</p>
  </section>
  <section>
    <h1>Granny Smith</h1>
    <p>These juicy, green apples make a great filling for
    apple pies.</p>
  </section>
</article>
```

<article>

- **Self-contained chunk of content**
 - Something you may want to share
 - Independently distributable or reusable (syndication)
- **Examples:**
 - Forum post
 - Magazine/newspaper article
 - Blog entry
 - User-submitted comment

```
<article itemscope itemtype="http://schema.org/BlogPosting">
  <header>
    <h1 itemprop="headline">The Very First Rule of Life</h1>
    <p><time itemprop="datePublished" datetime="2009-10-09">3 days ago</time></p>
    <link itemprop="url" href="?comments=0">
  </header>
  <p>If there's a microphone anywhere near you, assume it's hot and
  sending whatever you're saying to the world. Seriously.</p>
  <p>...</p>
  <section>
    <h1>Comments</h1>
    <article itemprop="comment" itemscope itemtype="http://schema.org/UserComments" id="c1">
      <link itemprop="url" href="#c1">
      <footer>
        <p>Posted by: <span itemprop="creator" itemscope itemtype="http://schema.org/Person">
          <span itemprop="name">George Washington</span>
        </span></p>
        <p><time itemprop="commentTime" datetime="2009-10-10">15 minutes ago</time></p>
      </footer>
      <p>Yeah! Especially when talking about your lobbyist friends!</p>
    </article>
    <article itemprop="comment" itemscope itemtype="http://schema.org/UserComments" id="c2">
      <link itemprop="url" href="#c2">
      <footer>
        <p>Posted by: <span itemprop="creator" itemscope itemtype="http://schema.org/Person">
          <span itemprop="name">George Hammond</span>
        </span></p>
        <p><time itemprop="commentTime" datetime="2009-10-10">5 minutes ago</time></p>
      </footer>
      <p>Hey, you have the same first name as me.</p>
    </article>
  </section>
</article>
```

Headers and footers

- `<header>`, for the tops of:
 - Sections, articles
 - Top of body for main header of your page
 - You can have multiple
- `<footer>`, for the bottoms of:
 - Sections, articles
 - Anywhere you need footer content
 - e.g. Author of document, copyright, privacy policy
 - You can have multiple

Header/footer example

```
<header>  
<h1>Scientist discover way to reduce headaches</h1>  
<b><p>Sleeping with your shoes strongly correlated with  
waking up with a headache</p></b>  
</header>  
<article>  
<p>Blah blah blah blah blah</p>  
<p>Blah blah blah...</p>  
</article>  
<footer>  
Copyright 2012 by Author  
</footer>
```

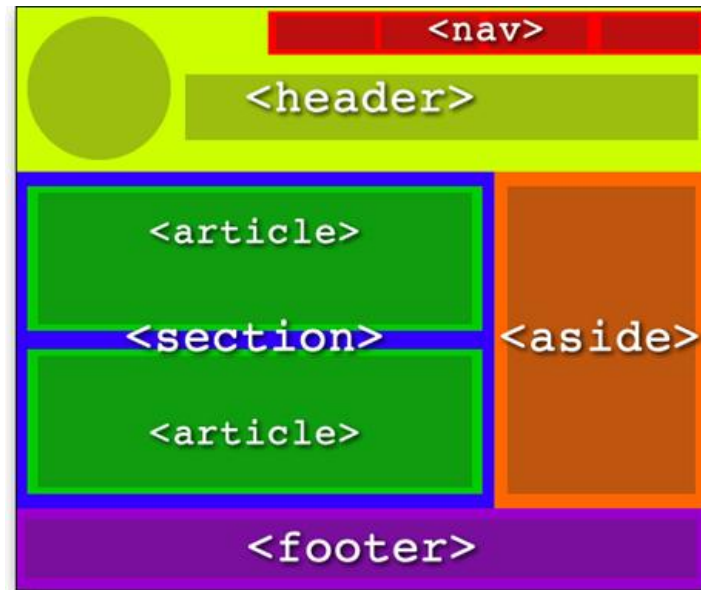
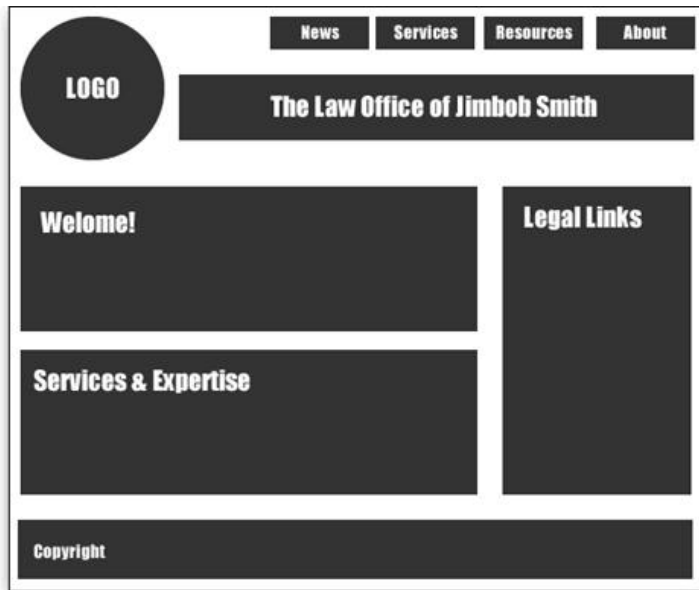
<hgroup>

- Represents the headings of a section
 - Group set of h1-h6 elements when heading has multiple levels
 - e.g. subheadings, alternative titles, taglines

```
<hgroup>  
  <h1>The reality dysfunction</h1>  
  <h2>Space is not the only void</h2>  
</hgroup>  
<hgroup>  
  <h1>Dr. Strangelove</h1>  
  <h2>Or: How I Learned to Stop Worrying and Love the Bomb</h2>  
</hgroup>
```

<nav>

- Navigation and links
 - Used for groups of links, not a single link
 - e.g. Links to all articles in a forum thread
 - Not needed for links in <header>, <footer>



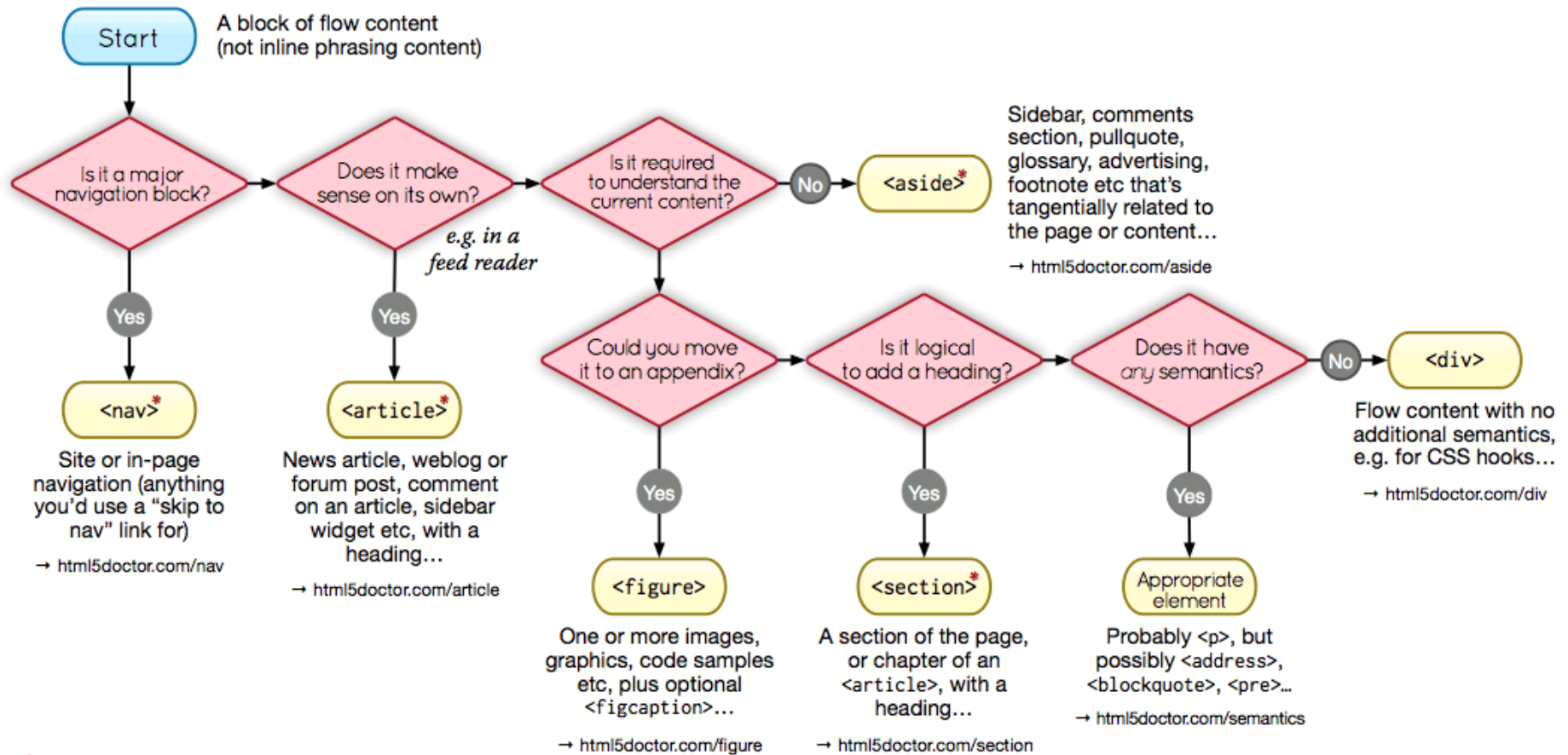

```

<header>
  <h1>Wake up sheeple!</h1>
  <p><a href="news.html">News</a> -
    <a href="blog.html">Blog</a> -
    <a href="forums.html">Forums</a></p>
  <p>Last Modified: <span itemprop="dateModified">2009-04-01</span></p>
  <nav>
    <h1>Navigation</h1>
    <ul>
      <li><a href="articles.html">Index of all articles</a></li>
      <li><a href="today.html">Things sheeple need to wake up for today</a></li>
      <li><a href="successes.html">Sheeple we have managed to wake</a></li>
    </ul>
  </nav>
</header>
<div>
  <article itemprop="blogPosts" itemscope itemtype="http://schema.org/BlogPosting">
    <header>
      <h1 itemprop="headline">My Day at the Beach</h1>
    </header>
    <div itemprop="articleBody">
      <p>Today I went to the beach and had a lot of fun.</p>
    </div>
    <footer>
      <p>Posted <time itemprop="datePublished" datetime="2009-10-10">Thursday</time>.</p>
    </footer>
  </article>
</div>
<footer>
  <p>Copyright 2010 The Example Company</p>
  <p><a href="about.html">About</a> -
    <a href="policy.html">Privacy Policy</a> -
    <a href="contact.html">Contact Us</a></p>
</footer>

```

Some other elements

- `<aside>`
 - Chunks of content outside main flow of text
 - Sidebar, quote, after-though
- `<time>`
 - Machine readable time/date
- `<abbr>`
 - Expansion of an abbreviation
- `<mark>`
 - Mark words, for highlighting or editing



* Sectioning content element

These four elements (and their headings) are used by HTML5's outlining algorithm to make the document's outline
→ html5doctor.com/outline

Even more meaning

- How does W3C know what I actually need?
 - I want to markup:
 - Names: companies, first names, last names, pet names
 - Sarcastic comments in forum posts
 - Info about the CDs in my library
 - Extend HTML in some way so I can add my own tags or attributes

Approaches

- **Microdata**

- WHATWG HTML specification
- Web Hypertext Technology Working Group
 - Development of HTML / APIs, formed by Apple, Mozilla, Opera
 - Response to "W3C's direction with XHTML, lack of interest in HTML and apparent disregard for the needs of real-world authors"
- Now a W3C working draft

- **Microformats**

- Grassroots effort, not a standards body
- 34 published formats

- **RDFa** (Resource Description Framework with attributes)

- W3C recommendation
- Set of attribute extensions to XHTML, HTML4, HTML5

Google rich snippets

- What custom semantic tags should I add?
 - One answer: those supported by Google
- Google rich snippets
 - "designed to give users a sense of what's on the page and why it's relevant to their query"
 - Microdata (recommended)
 - Microformats
 - RDFa
 - Content types:
 - Reviews, people, products, businesses and organizations, recipes, events, music

<http://www.google.com/webmasters/tools/richsnippets>

Microdata details

- **Microdata vocabularies**
 - Meaning for an item
 - Design your own custom one, or link to one
 - <http://data-vocabulary.org>
- **Attributes**
 - itemscope
 - Creates the item, descendants of this element has the information
 - itemtype
 - URL to the vocabulary that describes the item
 - itemprop
 - Value of a particular property of the item

Microdata person example

```
<section itemscope itemtype="http://data-vocabulary.org/Person">
  <h2 itemprop="name">Keith Vertanen</h2>
  
  <p>I am an assistant professor at
  <span itemprop="affiliation">Montana Tech</span>.</p>
  <address itemprop="address" itemscope
    itemtype="http://data-vocabulary.org/Address">
    <b>Address:</b>
    <p itemprop="locality">Butte, Montana</p>
  </address>
</section>
```

<http://keithv.com/websci/keith.html>

Property	Description
name (fn)	Name.
nickname	Nickname.
photo	An image link.
title	The person's title (for example, Financial Manager).
role	The person's role (for example, Accountant).
url	Link to a web page, such as the person's home page.
affiliation (org)	The name of an organization with which the person is associated (for example, an employer). If <code>fn</code> and <code>org</code> have the exact same value, Google will interpret the information as referring to a business or organization, not a person.
friend	Identifies a social relationship between the person described and another person.
contact	Identifies a social relationship between the person described and another person.
acquaintance	Identifies a social relationship between the person described and another person.
address (adr)	The location of the person. Can have the subproperties <code>street-address</code> , <code>locality</code> , <code>region</code> , <code>postal-code</code> , and <code>country-name</code> .

Microdata product example

```
<div itemscope itemtype="http://data-vocabulary.org/Product">
  <span itemprop="brand">ACME</span> <span itemprop="name">Executive Anvil</span>
  <br />
  <span itemprop="description">Sleeker than ACME's Classic Anvil, the Executive Anvil is perfect
for the business traveler looking for something to drop from a height.
</span><br />Category:
<span itemprop="category" content="Hardware > Tools > Anvils">Anvils</span><br />
Product #: <span itemprop="identifier" content="mpn:925872">925872</span><br />
<span itemprop="review" itemscope
  itemtype="http://data-vocabulary.org/Review-aggregate">
  <span itemprop="rating">4.4</span> stars, based on
  <span itemprop="count">89</span> reviews <br />
  $<span itemprop="price">119.99</span>
</span>
</div>
```

Property	Description
brand	The brand of the product-for example, ACME.
category	The product category-for example, "Books-Fiction", "Heavy Objects", or "Cars".
description	Product description.
name (fn)	Product name.
image	URL of product photo.
review	A nested review-aggregate of the product (for example, the average rating). If there are multiple reviews of the product, mark up aggregated review data (for example, the average rating from all users) using Review-aggregate rather than individual reviews.
identifier	The product identifier. Google recommends including brand and at least one identifier for each product. Recognized types include: <code>asin</code> , <code>isbn</code> , <code>mpn</code> , <code>sku</code> , <code>upc</code> .
offerdetails	An offer to sell the product. Includes a nested Offer or Offer-aggregate.

Microdata

Q: Why are Google, Bing, Yandex and Yahoo! collaborating? Aren't you competitors?

Currently, there are many standards and schemas for marking up different types of information on web pages. As a result, it is difficult for webmasters to decide on the most relevant and supported markup standards to use.

Creating a schema supported by all the major search engines makes it easier for webmasters to add markup, which makes it easier for search engines to create rich search features for users.

Q: Why microdata? Why not RDFa or microformats?

Focusing on microdata was a pragmatic decision. Supporting multiple syntaxes makes documentation for webmasters more complex and introduces more overhead in terms of defining new formats. Microformats are concise and easy to understand, but they don't offer an open extensibility mechanism and the reuse of the class tag can cause conflicts with website CSS. RDFa is extensible and very expressive, but the substantial complexity of the language has contributed to slower adoption. Microdata is the most recent well-known standard, created along with HTML5. It strikes a balance between extensibility and simplicity, and is most suitable for building the schema.org.

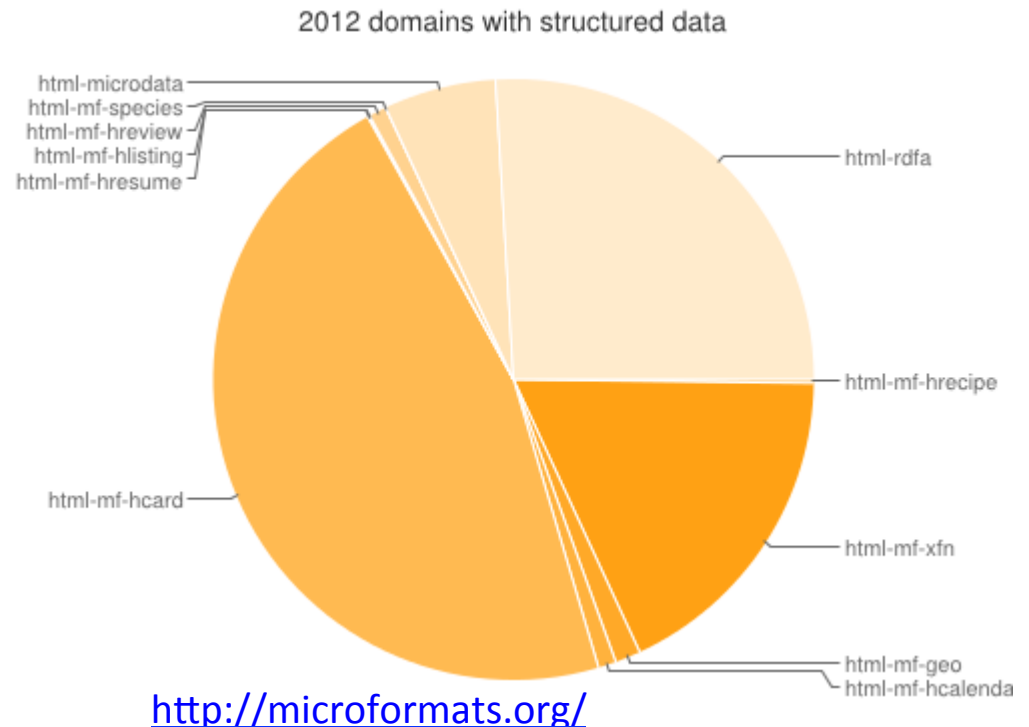
<http://schema.org/docs/faq.html#0>

Microformat example

```
<div class="vcard">
  
  <strong class="fn">Bob Smith</strong>
  <span class="title">Senior editor</span> at <span class="org">ACME Reviews</span>
  <span class="adr">
    <span class="street-address">200 Main St</span>
    <span class="locality">Desertville</span>, <span class="region">AZ</span>
    <span class="postal-code">12345</span>
  </span>
</div>
```

- **Microformat**

- Uses class tag
- "Humans first, machines second"



RDFa example

```
<html>
<body>
<div xmlns:v="http://rdf.data-vocabulary.org/#" typeof="v:Person">
My name is <span property="v:name">Bob Smith</span>,
but people call me <span property="v:nickname">Smithy</span>.
Here is my homepage:
<a href=http://www.example.com rel="v:url">www.example.com</a>.
I live in Albuquerque, NM and work as an
<span property="v:title">engineer</span>
at
<span property="v:affiliation">ACME Corp</span>.
</div>
</body>
</html>
```

- RDFa
 - Implemented by Best Buy
 - 30% increase in organic search traffic
 - 15% in Click-through rate (CTR)

Summary

- Semantic markup
 - Using tags/attributes to describe meaning
 - Separates presentation from semantics
 - Makes computer processing easier
 - Very hard to parse meaning from arbitrary page: AI complete
 - Multiple standards
 - Microdata
 - Microformats
 - RDFa