

Embedding Knowledge in HTML

Some content from a presentations by Ivan Herman of the W3c

Overview

- Why we want to embed structured data in HTML
- RDFa
- Microdata and schema.org
- RDFa lite as an encoding for Microdata
- JSON-LD as an encoding for RDF and Microdata
- Use cases and examples

HTML is Everywhere

- We usually think of HTML as the language of Web pages
- But it's also widely used on/for mobile devices and tablets
 - Readily adapts to different screen sizes/orientations
- And is the basis of many ebook formats
 - E.g. Kindle's formats, mobi, epub, ...
- How can we add knowledge to HTML pages?

Adding Semantic Markup to HTML

- Like to add semi-structured knowledge in HTML
 - Humans see and understand regular HTML content (text, images, videos, audio)
 - Machines see and understand data markup in XML, RDF or some other format
- Possibilities include
 - Add a link to separate document with knowledge
 - Add new HTML tags
 - Embed knowledge as comments, Javascript, etc.
 - Distribute knowledge markup throughout HTML as attributes of existing HTML tags

One page, not two

- Content providers prefer not to generate multiple page:, one for humans (HTML) and another for machines (RDF)
 - RDF serializations are complex
 - Requires separate mechanisms for storage, generation, etc.
 - Introduces redundancy, which can lead to errors if we change one page but not the other
- Simplifies the job of search engines as well

General approach

- Two embedding techniques

- Provide/reuse tag *attributes* to encode the metadata; browsers/apps ignore attributes they don't understand
 - Add a graph for page in JSON-LD

- Several approaches have been developed

- Microformats (~ 2005)
 - RDFa (~ 2007) and JSON-LD (2014)
 - Microdata (aka schema.org) (~ 2012)

- Status 2018 (IMHO)

- *Microformats* still on many pages, but no longer actively used
 - *RDFa* and JSON-LD are preferred encoding of choice
 - Schema.org vocabularies getting large uptake

Microdata approach

- Defined and supported by Google, Bing, Yahoo and Yandex
- Adds new attributes to HTML5 to express metadata
- Works well for simple “single-vocabulary” cases, but not well suited for mixing vocabularies or for complex vocabularies
- No notion of datatypes or namespaces
- Defines a generic mapping to RDF

RDFa approach

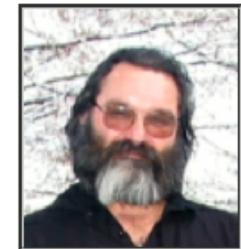
- Adds new (X)HTML/XML attributes
- Has namespaces and URIs at its core
 - So mixing vocabulary is easy, as in RDF
- Complete flexibility for using literals or URI resources
- Is a complete serialization of RDF



Ivan Herman

Who am I?

I graduated as mathematician at the [Eötvös Loránd University of Budapest](#), Hungary, in 1979. After a brief scholarship at the Université Paris VI I joined the Hungarian research institute in computer science ([SZTAKI](#)) where I worked for 6 years (and turned into a computer scientist...). I left Hungary in 1986 and, after a few years in industry in Munich, Germany, I joined the [Centre Mathematics and Computer Sciences \(CWI\)](#) in Amsterdam where I have a tenure position since 1988. I received a PhD degree in Computer Science in 1990 at the [University of Leiden](#), in the Netherlands. I joined the [World Wide Web Consortium \(W3C\)](#) Team as Head of [W3C Offices](#) in January 2001 while maintaining my position at CWI. I served as Head of Offices until June 2006, when I was asked to take the [Semantic Web Activity](#) Lead position, which is now my principal work at W3C.



Before joining W3C I worked in quite different areas (distributed and dataflow programming, language design, system programming), but I spend most of my research years in computer graphics and information visualization. I also participated in various graphics related ISO standardization activities and software developments. My "[professional](#)" [home page](#) contains a list of [my publications](#) (see also [my Mendeley account](#)), [my public presentations](#), and details of the various projects I participated in the past. There is also a [dblp entry for my publications](#) generated automatically (although I am not sure it is complete...). (B.t.w., based on my publications, my Erdős number is ≤ 4 ...)

In my previous life (i.e., before joining W3C...) I was member of the Executive Committee of the [Eurographics Association](#) for 15 years, and I was vice-chair of the Association between 2000 and 2002. I was the co-chair of the [9th World Wide Web Conference](#), in Amsterdam, May 2000; since then, I have also been member of [IW3C2 \(International World Wide Web Conference Committee\)](#), responsible for the World Wide Web Conference series. Since autumn 2007 I am also member of [SWSA \(Semantic Web Science Association\)](#), the committee responsible for the International Semantic Web Conferences (better known as "ISWC") series.

Some personal data

- The Hungarian spelling of my full name is Herman Iván. Ie, my name is Ivan (well, spelled properly: Iván) and my surname is Herman (many in the Netherlands and in Germany mix it up, and use "Herman" as my name... this is aggravated by the fact that, uniquely in Europe, the Hungarian custom is to put surname first).
- Nationalities: French and Hungarian
- Gender: male
- Family: I am married and have a son, David.
- Date and city of birth: 24th February, 1955, [Budapest](#), Hungary
- Email addresses: 'ivan' on my own ivan-herman.net domain, 'ivan' on the w3.org domain, or 'ivan.herman' on the cwi.nl domain
- (Mobile) Phone: +31-641044153
- Skype ID: ivan_herman
- I live in [Amstelveen](#) (see also [geonames](#)), the Netherlands (lat: 52.302063, long: 4.87397). This is a suburb of [Amsterdam](#). The closest airport is Amsterdam Schiphol
- I am the administrator of the Semantic Web Activity Blog at W3C which can either be accessed directly or via [its](#)

some links

- [personal homepage](#)
- more data on me
- [personal blog \(RSS feed\)](#)
- [homepage at W3C](#)
- ["professional" homepage](#)
- ["official" CV](#)
- more about me
- [my photos](#)

"social" links

- [facebook](#)
- [flickr](#)
- [picasa web](#)
- [linkedin](#)
- [IWIW](#)
- [tripit](#)
- [twitter](#)
- [Mendeley](#)
- [Google+](#)
- [freebase](#)



Ivan Herman

Who am I?

I graduated as mathematician at the **Eötvös Loránd University of Budapest**, Hungary, in 1979. After the Université Paris VI I joined the Hungarian research institute in computer science (**SZTAKI**) (and turned into a computer scientist...). I left Hungary in 1986 and, after a few years joined the **Centre Mathematics and Computer Sciences (CWI)** in Amsterdam where I received a PhD degree in Computer Science in 1990 at the **University of Leiden**. Web Consortium (W3C) Team as Head of **W3C Offices** in January 1994 until June 2006, when I was asked to take principal work at W3C.

Before joining W3C I worked in quite different areas (mostly programming), but I spent some time participating in various standardization efforts.

I graduated as mathematician at the Eötvös Loránd University of Budapest, Hungary, in 1979.

I joined the [Team as Head of W3C Offices](#).

** Team as Head of **
<link rel="owl:sameAs" href="http://www.w3.org/2001/07/foaf.rdf#me">
<link rel="foaf:workplaceHomepage" href="http://www.w3.org/2001/07/foaf.rdf#me">
<meta property="schema:jobTitle" content="Team Leader"

<p>Before joining W3C I worked in quite different areas (mostly programming), but I spent some time participating in various standardization efforts.

<p>In my previous life (i.e., before joining W3C...), I graduated as mathematician at the Eötvös Loránd University of Budapest, Hungary, in 1979.

</p>

<h2>

Some personal data

</h2>

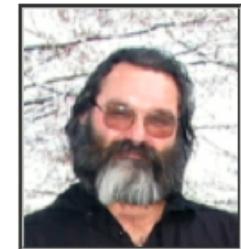
The Hungarian spelling of my full name is Ivan and my surname is Herman. (many in the Netherlands and in Germany mix it up, and use "Herman Ivan" or "Herman Herman")

Ief scholarship at worked for 6 years Munich, Germany, I position since 1988. I joined the World Wide position at CWI. I served as position, which is now my

language design, system information visualization. I also developments. My "professional" home presentations, and details of the various generated automatically (although I am not

Committee of the **Eurographics Association** as the co-chair of the **9th World Wide Web Conference (IW3C2)** (International World Wide Web Conferences). Since autumn 2007 I am also member of the International Semantic Web Conferences

name is Ivan (well, spelled properly: Iván) and my mix it up, and use "Herman" as my name... this is custom is to put surname first).



some links

- [personal homepage](#)
- [more data on me](#)
- [personal blog \(RSS feed\)](#)
- [homepage at W3C](#)
- ["professional" homepage](#)
- ["official" CV](#)
- [more about me](#)
- [my photos](#)

"social" links

- [facebook](#)
- [flickr](#)
- [picasa web](#)
- [linkedin](#)
- [IWIW](#)
- [tripit](#)
- [twitter](#)
- [Mendeley](#)
- [Google+](#)
- [freebase](#)

```

<meta property="foaf:accountName" content="ivan-herman" />
</a>
</li>
<li>
  <a href="https://plus.google.com/u/0/113268051484517627727" typeof="foaf:OnlineAccount">
    <span property="foaf:accountServiceHomepage" href="http://www.mendeley.com/">Google+</span>
    <meta property="foaf:accountName" content="113268051484517627727" />
  </a>
</li>
<li>
  <a about="http://www.ivan-herman.net/foaf#me" rel="owl:sameAs" resource="http://rdf.freebase.com/ns/en.113268051484517627727" type="foaf:Person">
    <span property="foaf:name" href="http://www.ivan-herman.net/foaf#me">Ivan Herman</span>
    <meta property="foaf:title" content="Dr" />
    <h2>Who am I?</h2>
    <p>I graduated as mathematician at the <a rel="foaf:schoolHomepage schema:alumniOf" href="http://www.elte.hu/">ELTE</a> Budapest University of Technology and Economics<br/>I joined the <a rel="schema:worksFor" href="http://www.w3.org" resource="http://www.w3.org/Data#W3C">World Wide Web Consortium (W3C)</a><br/>
      <span property="dc:title">World Wide Web Consortium (W3C)</span>
    </p>
    Team as Head of <a rel="foaf:pastProject" href="http://www.w3.org/Consortium/Offices"><span property="dc:title">W3C</span></a><br/>
    <link rel="owl:sameAs" href="http://www.ivan-herman.net/me" />
    <link rel="owl:sameAs" href="http://www.ivan-herman.net/Ivan_Herman" />
    <link rel="foaf:workplaceHomepage" href="http://www.w3.org"/>
    <meta property="schema:jobTitle" content="Semantic Web Activity Lead" />
</li>
</ul>
</div>

<div id="content" >
  <h1 property="schema:name foaf:name">Ivan Herman</h1>
  <meta property="foaf:title" content="Dr" />
  <h2>Who am I?</h2>
  <p>I graduated as mathematician at the <a rel="foaf:schoolHomepage schema:alumniOf" href="http://www.elte.hu/">ELTE</a> Budapest University of Technology and Economics<br/>I joined the <a rel="schema:worksFor" href="http://www.w3.org" resource="http://www.w3.org/Data#W3C">World Wide Web Consortium (W3C)</a><br/>
    <span property="dc:title">World Wide Web Consortium (W3C)</span>
  </p>
  Team as Head of <a rel="foaf:pastProject" href="http://www.w3.org/Consortium/Offices"><span property="dc:title">W3C</span></a><br/>
  <link rel="owl:sameAs" href="http://www.ivan-herman.net/me" />
  <link rel="owl:sameAs" href="http://www.ivan-herman.net/Ivan_Herman" />
  <link rel="foaf:workplaceHomepage" href="http://www.w3.org"/>
  <meta property="schema:jobTitle" content="Semantic Web Activity Lead" />
</div>

<p>Before joining W3C I worked in quite different areas (distributed and dataflow programming, language design, system design, ...)</p>
<p>In my previous life (i.e., before joining W3C...) I was member of the Executive Committee of the <a rel="foaf:pastProject" href="http://www.w3.org/Consortium/Offices"><span property="dc:title">W3C</span></a><br/>
  <h2>Some personal data</h2>
  <ul>
    <li>The Hungarian spelling of my full name is <span property="foaf:name" lang="hu">Herman Iván</span>. Ie, my name is <span property="foaf:givenname schema:givenName">Ivan</span> (well, spelled properly: <span property="foaf:givenname schema:givenName" lang="hu">Iván</span>) and my surname is <span property="foaf:surname schema:familyName">Herman</span> (many in the Netherlands and in Germany mix it up, and use "Herman" as my name... this is aggravated by the fact that the Dutch spelling "Herman" is identical to the German spelling "Hermann")</li>
  </ul>
</div>

```

Yielding this RDF

```
<http://www.ivan-herman.net/foaf#me>
    schema:alumniOf      <http://www.elte.hu> ;
    foaf:schoolHomePage <http://www.elte.hu> ;
    schema:worksFor     <http://www.w3.org/W3C#data>
;
...
<http://www.elte.hu>
    dc:title "Eötvös Loránd University of Budapest" .
...
<http://www.w3.org/W3C#data>
    dc:title "World Wide Web Consortium (W3C)"
```

The Telegraph

Monday 09 April 2012

[HOME](#) [NEWS](#) [SPORT](#) [FINANCE](#) [COMMENT](#) [BLOGS](#) [CULTURE](#)
[TRAVEL](#) [LIFESTYLE](#) [FASHION](#) [TECH](#)
[Dating](#) [Offers](#) [Jobs](#)
[Film](#) [Music](#) [Art](#) [Books](#) [TV and Radio](#) [Theatre](#) [Hay Festival](#) [Dance](#) [Opera](#) [Photography](#) [Comedy](#) [Video](#) [In the Know](#)
[Oscars](#) [Film Reviews](#) [Cinema Trailers](#) [Coming Soon](#) [Talking Movies](#) [Interviews](#) [DVDs](#) [Film Life](#) [Film Video](#)
[HOME](#) » [CULTURE](#) » [FILM](#) » [FILM REVIEWS](#)

Oscars 2012: The Artist, review

The Artist, an utterly beguiling silent, black-and-white celebration of early Hollywood won Best Picture at the Oscars 2012.



To attend an Info Session or Masterclass Or download a brochure

[Click here](#)

The Business School
for the World®

Global Executive MBA


[TELEGRAPH TICKETS »](#)


```
<li class="list-item">
<li><a href="#">
```

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

```
<!-- googleon: all -->
```

```
<div id="tmglBody" >
<div class="access"><a name="a1" href="#">
```

```
<div class="twoThirdsThird2 gutter">
<div class="twoThirds gutt">
<div class="story">
```

```
<h1 itemprop="name">Oscars 2012: The Artist, review</h1>
<h2 itemprop="description">
```

The Artist, an utterly beguiling silent, black-and-white celebration of early Hollywood won Best Picture at the Oscars 2012

```
</h2>
```

```
<div class="rating" itemprop="reviewRating">
<meta itemprop="worstRating" content="0.0" />
<meta itemprop="bestRating" content="5.0" />
<span itemprop="ratingValue" content="4.5" />

```

```
</div>
<div class="artIntro">
```

```
<div id="storyEmbed" style="border: 1px solid #ccc; padding: 5px; margin-bottom: 10px;">
<div class="slideshow ssIntro">
<div class="nextPrevLayer">
<div class="prevLayer" style="float: left; width: 40px; height: 40px; background: url(/template/ver1-0/i/icon_left.gif) no-repeat; border-radius: 50%;">

```

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

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

```
<div class="oneHalf gutter">
<div class="story">
<div class="cl"> </div>
<!-- remove the whitespace added by escenic before end of -->
```



Search - enhanced by Google

Monday 09 April 2012

Dating Offers Jobs

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE

TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

HOME NEWS WORLD CHANNEL COMMENT BLOGS CULTURE TRAVEL LIFESTYLE FASHION TECH

To attend an Info Session or Masterclass Or download a brochure

Click here

INSEAD

The Business School
for the World®

Global Executive MBA



TELEGRAPH TICKETS *



```

        <li class="first"><a href="/">Home</a><span>&raquo;</span></li>
        <li><a href="http://www.telegraph.co.uk/culture/">Culture</a><span>&raquo;</span></li>
            <li><a href="http://www.telegraph.co.uk/culture/film/">Film</a><span>&raquo;</span></li>
            <li class="styleSix"><a href="http://www.telegraph.co.uk/culture/film/filmreviews/">Film F
        <div class="cl"></div>
    </div>

<!-- googleon: all -->
<div id="tmglBody" >
    <div class="access"><a name="article"></a></div>

    <div class="twoThirdsThird2 gutterUnder">
        <div class="twoThirds gutter" itemscope itemtype="http://schema.org/Review">
            <div class="storyHead">

                <h1 itemprop="name">Oscars 2012: The Artist, review</h1>
                <h2 itemprop="description">
The Artist, an utterly beguiling silent, black-and-white celebration of early
Hollywood won Best Picture at the Oscars 2012.

            </h2>
                <div class="rating" itemprop="reviewRating" itemscope itemtype="http://schema.org/Rating">
                    <meta itemprop="worstRating" content = "0.5">
                    <meta itemprop="bestRating" content = "5">
                    <span itemprop="ratingValue" class="hidden">5</span>
                    
                </div>
                <div class="artIntro">
                    <div id="storyEmbSlide">
                        <div class="slideshow ssIntro">
                            <div class="nextPrevLayer">
                                <div class="ssImg">
                                    
                                        <div class="ingCaptionCredit">
                                            <span class="caption">Bérénice Bejo as ris
                                    </div>
                                </div>
                            </div>
                        </div>
                    </div>
                </div>
            </div>
            <div class="oneHalf gutter">
                <div class="story">
                    <div class="cl"></div>
                    <!-- remove the whitespace added by escenic before end of </a> tag -->

```

Yielding this RDF

```
[ rdf:type schema:Review ;
  schema:name "Oscars 2012: The Artist, review" ;
  schema:description "The Artist, an utterly
beguiling..." ;
  schema:ratingValue "5" ;
...
]
```

Rich Snippets



- Search engines add text under results to preview what's on page and why it's relevant
- Text often extracted from structured data embedded on the page
- Of great interest to SEO, see this for more information

The screenshot displays two search results with rich snippets:

Little Water Cantina - Eastlake - Seattle, WA
www.yelp.com › Restaurants › Mexican
★★★★★ 90 reviews - Price range: \$\$
90 Reviews of Little Water Cantina. "Three things are on my list when I eat out: great food, atmosphere, and

[Vegetarian Vegan Pizza No Cheese\) Recipe - Food.com - 248865](#)
www.food.com/recipe/vegetarian-vegan-pizza-no-c...
★★★★★ 2 reviews - 1 hr 32 mins - 242.9 cal
Aug 26, 2007 – This is from my dad, who developed some **vegan recipes** that don't have any cheese, and you

Leonard Cohen – Free listening, videos, concerts, stats, & pictures at...
www.last.fm/music/Leonard+Cohen
Watch videos & listen to **Leonard Cohen**: Suzanne, Hallelujah & more, plus 132 pictures. **Leonard Cohen**, (born September 21, 1934 in Montréal, Québec, ...

Track	Duration
Suzanne	3:48
The Darkness	4:29
Going Home	3:51
Hallelujah	6:12

Delicious LocalData TR 2012 My Mercurial Private Mailing lists Social SW Python RDFA it! Bookmarklets To... Web Data Inspector

Everything

Images

Maps

Videos

News

Shopping

More

Amsterdam

Change location

Any time

Past hour

Past 24 hours

Past week

Past month

Past year

Custom range...

More search tools

The Artist showtimes for Amsterdam

[Pathe Tuschinski](#) - Reguliersbreestraat 26-34, Amsterdam - [Map](#)
11:50 - 14:05 - 19:10

[Filmtheater "De Uitkijk"](#) - Prinsengracht 452, Amsterdam - [Map](#)
12:15 - 19:00 - 21:15

[Filmtheater Rialto](#) - Ceintuurbaan 338, Amsterdam - [Map](#)
12:45

[+ Show more theaters](#)

The Artist (2011) - IMDb

www.imdb.com/title/tt1655442/

Silent **movie** star George Valentin bemoans the coming era of talking ... Still of Jean Dujardin and Missi Pyle in **The Artist** Still of Bérénice Bejo in **The Artist** Reem ...

↳ [Full cast and crew](#) - [The Artist Trailer \(Official ...](#) - Bérénice Bejo - Jean Dujardin

The Artist (film) - Wikipedia, the free encyclopedia

[en.wikipedia.org/wiki/The_Artist_\(film\)](http://en.wikipedia.org/wiki/The_Artist_(film))

The Artist is a 2011 French romantic comedy drama in the style of a black-and-white silent film written and directed by Michel Hazanavicius, starring Jean ...

↳ [Jean Dujardin](#) - [Bérénice Bejo](#) - [Uggie](#) - [Diegesis](#)

The Artist Trailer 2011 HD - YouTube

www.youtube.com/watch?v=O8K9AZcSQJE

25 Aug 2011 - 3 min - Uploaded by TrailersApplecom

I love how George Clooney, and Brad Pitt, lost the Best actor catagory to this **film**. It just shows that there is ...

[More videos for the artist movie »](#)

Oscars 2012: The Artist, review - Telegraph

www.telegraph.co.uk/Culture/Film/Film_Reviews

★★★★★ Review by Robbie Collin

27 Feb 2012 – **The Artist**, the final **film** to be released in 2011 and also the most heart-swellingly joyful one, is a silent **movie**, screened in black and white and ...

[The Artist is the perfect **film** about Hollywood | Hadley Freeman](#)

RDFa and Microdata: similarities

- RDFa and Microdata are modern options
- Both have similar approaches
 - Structured data encoded in *HTML attributes only* – no new elements
 - Define some special *attributes*
 - e.g., **itemscope** for microdata, **resource** for RDFa
 - Reuse *some* HTML core attributes (e.g., **href**)
 - Use textual content of HTML source, if needed
- RDF data can be extracted from both

RDFa and microdata: differences

- Microdata *optimized* for simpler use cases:
 - One vocabulary at a time
 - Tree shaped data
 - Few datatypes
- RDFa provides full serialization of RDF in XML or HTML
 - Price is extra complexity over Microdata
- RDFa 1.1 Lite is a simplified authoring profile of RDFa, very similar to microdata

Amount of structured data on Web?

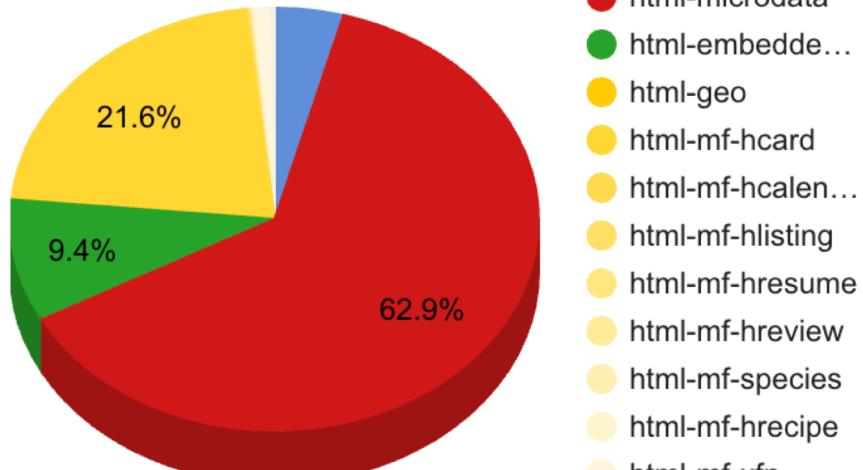


- Web Data Commons project uses Common Crawl data to estimate amount of structured data on Web
- Looks for Microdata, RDFa, other formats (e.g., hCalendar, hCard) in URLs parseable as HTML
- Nov. 2017 crawl analyzed 3.2B HTML pages:
 - 1.2B pages (30%) with structured data in 5.6M domains of 34M (17%)
 - 39B triples about 9B typed entities
- Data can be downloaded

Amount of structured data on Web?



Triples



Top Domains by Extracted Triples

1. [blogspot.com](#) (2,326,407,936 triples)
2. [wordpress.com](#) (836,982,484 triples)
3. [theclothdiaperwhisperer.com](#) (258,310,473 triples)
4. [skyrock.com](#) (232,476,462 triples)
5. [moosejaw.com](#) (147,690,597 triples)
6. [canalblog.com](#) (110,509,475 triples)
7. [hotels.com](#) (109,331,631 triples)
8. [peternyssen.com](#) (106,756,630 triples)
9. [untitled-magazine.com](#) (106,468,802 triples)
10. [justia.com](#) (104,903,804 triples)
11. [leadferret.com](#) (92,088,394 triples)
12. [drom.ru](#) (80,352,891 triples)
13. [tutete.com](#) (74,136,046 triples)
14. [prom.ua](#) (73,993,196 triples)
15. [epicsports.com](#) (72,744,802 triples)
16. [repairpal.com](#) (68,768,136 triples)
17. [callersmart.com](#) (68,117,826 triples)
18. [lyst.com](#) (67,716,741 triples)
19. [kidsroom.de](#) (66,900,221 triples)
20. [More](#)

M Inbo Dona Intro Link Depa Face UMB TPM Talkir pralj Cx Appli UMB view Web Tim

https://www.moosejaw.com/moosejaw/shop/...

Moosejaw

FREE SHIPPING on Orders \$35+

What are you looking for?

Brands Jackets Clothing Footwear Hike & Camp Climb Snow More Activities Sale

Free Almost Gift Card up to \$100 - Use Code: **FREEMONEY** | **Free 2-Day Shipping on Jackets** over \$99 [DETAILS](#)

Home | Womens Jackets | Womens Canada Goose Jackets

VIDEO

360°





Canada Goose Women's Shelburne Parka

★★★★★ 4.8

\$925.00

Select a Size
[Size Guide](#)

Select a Color

Qty: This Item SHIPS FREE
When will I get it?

Add To Cart

Call 877-MOOSEJAW | [Gift Cards](#) | [Store Locator](#)

OUTDOOR KNOWLEDGE | MADNESS

Sign-up for Our Email

Microdata RDFa POSH

Statement Collection #1

Entity https://www.moosejaw.com/moosejaw/shop/product_Canada-Goose-Women-s-Shelburne-Parka_10254734_10208_10000001_-1

Attributes

- [rdf:type](#)
- [schema:Product](#)
- [See Statement Collection #2](#)
- [See Statement Collection #4](#)
- [See Statement Collection #6](#)
- [See Statement Collection #8](#)
- [See Statement Collection #10](#)
- [See Statement Collection #12](#)
- [See Statement Collection #14](#)
- [See Statement Collection #16](#)
- [See Statement Collection #18](#)
- [See Statement Collection #20](#)
- [See Statement Collection #22](#)
- [See Statement Collection #24](#)
- [See Statement Collection #26](#)
- [See Statement Collection #28](#)
- [See Statement Collection #30](#)
- [See Statement Collection #32](#)
- [See Statement Collection #34](#)
- [See Statement Collection #36](#)
- [See Statement Collection #38](#)
- [Name](#)
- [schema:brand](#)
- [schema:image](#)

Canada Goose Women's Shelburne Parka

Canada Goose

//s7d1.scene7.com/is/image/MoosejawMB/10254734x1018148_zm?
\$product1000\$

Search  |  | 

Activities & Camp Climb Snow More Activities Sale

NEY | Free 2-Day Shipping on Jackets over \$99 DETAILS

Canada Goose Women's Shelburne Parka

★★★★★ 4.8

\$925.00

Select a Size 

Select a Color 

Qty: This Item SHIPS FREE
When will I get it?

Add To Cart

Copyright © 2015-2017 OpenLink Software

ver: 2.15.3 OpenLink Structured Data Sniffer

Call 877-MOOSEJAW | Gift Cards | Store Locator

OUTDOOR KNOWLEDGE | MADNESS

Sign-up for Our Email

Sign-up for Our Email  

Conclusions

- The amount of structured data on the web is growing steadily
- Microdata shows the strongest growth
- RDFa also common