

Jak dostat RDF na web



ISKM89 Organizace dat - sémantický web | podzim 2023
Zuzana Nevěřilová | Centrum zpracování přirozeného jazyka

Strukturovaná data na webu

- mikroformáty (specifické pro doménu)
- RDFa
- HTML5 microdata (schema.org)
- HTML metadata via OpenGraph (by Meta/Facebook)
- JSON-LD
- Dublin Core
- Twitter Cards

Mikroformáty

Od roku 2005 se přidávají do standardního HTML, konkrétně do atributu `class`.

- Jednoduché koncepty
- Lidé mají přednost před stroji

vCard, also known as VCF (Virtual Contact File)

```
BEGIN:VCARD
VERSION:4.0
FN:Simon Perreault
N:Perreault;Simon;;;ing. jr,M.Sc.
BDAY:--0203
GENDER:M
EMAIL;TYPE=work:simon.perreault@viagenie.ca
END:VCARD
```

```
<div class="vcard">
  <div class="fn n">
    <span class="given-name">Simon</span>
    <span class="family-name">Perreault</span>
    <span class="honorofic-suffix">ing. jr,M.Sc.</span>
  </div>
  <div>B-Day:
    <span class="bday">February 03</span>
  </div>
  <div><span class="type">work</span> <span>email</span>
    <span class="email">simon.perreault@viagenie.ca</span>
  </div>
</div>
```

<https://microformats.org/wiki/hcard>

Mikrodata a mikroformáty

Výhodné pro hotové mikroformáty.

Může být složité najít podporu pro nový mikroformát v komunitě.

RDFa

RDFa 1.1 a RDFa Lite 1.1 (2015)

Nové atributy v HTML:

- vocab
- typeof
- property
- resource
- prefix

```
<p>  
  Google Inc.<br>  
  P.O. Box 1234<br>  
  Mountain View, CA<br>  
  94043<br>  
  United States<br>  
</p>
```

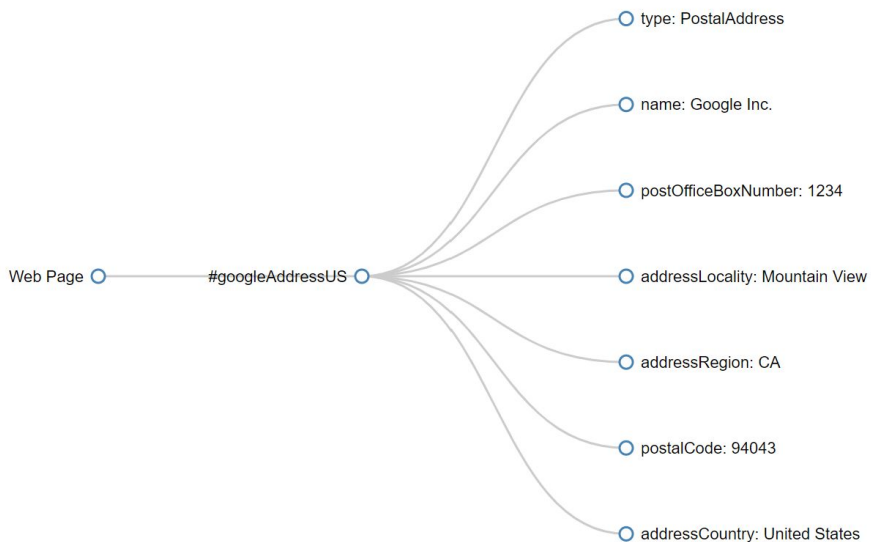
```
<p vocab="http://schema.org/"  
  typeof="PostalAddress"  
  resource="#googleAddressUS"><br>  
  <span property="name">Google Inc.</span><br>  
  P.O. Box  
  <span property="postOfficeBoxNumber">1234</span><br>  
  <span property="addressLocality">Mountain View</span>,  
  <br>  
  <span property="addressRegion">CA</span><br>  
  <span property="postalCode">94043</span><br>  
  <span property="addressCountry">United States</span><br>  
</p>
```

<https://schema.org/PostalAddress>

<https://rdfa.info/play/>

RDFa

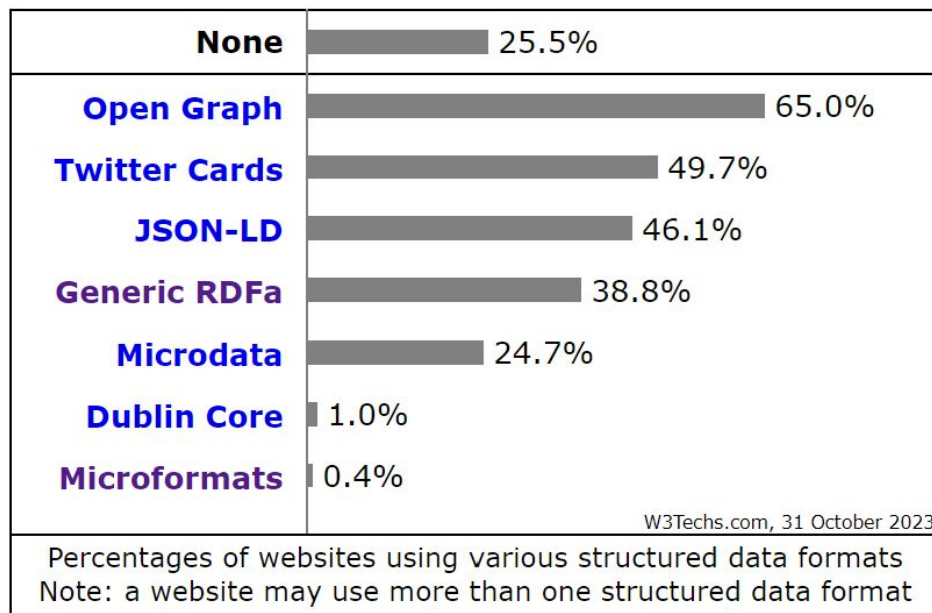
RDFa Tools



```
<p vocab="http://schema.org/"
  typeof="PostalAddress"
  resource="#googleAddressUS"><br>
<span property="name">Google Inc.</span>
<br>
P.O. Box
<span property="postOfficeBoxNumber">1234</span>
<br>
<span property="addressLocality">Mountain
View</span>,
<br>
<span property="addressRegion">CA</span>
<br>
<span property="postalCode">94043</span>
<br>
<span property="addressCountry">United
States</span>
<br>
</p>
```

<https://rdfa.info/play/>

Jaké technologie se používají?



https://w3techs.com/technologies/overview/structured_data

Jaké technologie se používají?

OpenGraph - Facebook

(<https://ogp.me/>)

TwitterCards - X

(<https://developer.twitter.com/>)

JSON-LD - Google Search

```
<html>
  <head>
    <title>Party Coffee Cake</title>
    <script type="application/ld+json">
      {
        "@context": "https://schema.org/",
        "@type": "Recipe",
        "name": "Party Coffee Cake",
        "author": {
          "@type": "Person",
          "name": "Mary Stone"
        },
        "datePublished": "2018-03-10",
        "description": "This coffee cake is awesome and perfect for parties.",
        "prepTime": "PT20M"
      }
    </script>
  </head>
  <body>
    <h2>Party coffee cake recipe</h2>
    ...
```