

21. konferenca Dnevi slovenske informatike

Informatika – neizkoriščeni dejavnik razvoja

Semantični splet v spletnih straneh



Ivan Verdonik

15. 04. 2014



Semantični splet v spletnih straneh

Razvoj svetovnega spleta

Svetovni splet je nastal po ideji Tim Berners-Leeja

- statične spletne strani
- dinamične spletne strani
- portali in spletne aplikacije
- socialno usmerjen splet (Splet 2.0)
- semantični splet (Splet 3.0)

Razlogi za semantični splet

- obstoječi splet je v prvi vrsti namenjen ljudem
- na svetovnem spletu je tudi veliko podatkov uporabnih za stroje



Semantični splet v spletnih straneh

Semantični splet

Sloji semantičnega spleta

- URI/IRI
- XML sintaksa
- RDF
- RDFS
- SPARQL
- OWL
- RIF/SWRL
- Še višji sloji (v raziskovalni fazi)



Semantični splet v spletnih straneh

Semantični splet

Uporaba

- Enostavno in zelo zmogljivo iskanje po RDF skladiščih s SPARQL
- Povpraševanja v (prilagojenem) naravnem jeziku
- SPARQL dostop do klasičnih relacijskih podatkovnih baz
- Izpeljava novega znanja preko pravil in Reasonerjev
- Linked Data – povezovanje z drugimi RDF skladišči na spletu
- Vgradnja v Splet 2.0 storitve
- Nadgradnja osnovnih, klasičnih spletnih strani



Semantični splet v spletnih straneh

Semantični splet

Načini vgradnje v spletne strani

- Mikrooblike (Microformats):
Najenostavnejši in (trenutno še) najpogostejši pristop – več kot 65 %
- Mikropodatki (Microdata):
Ne deluje na HTML 4, vgrajeni v HTML 5. Enostaven in zmogljiv pristop
- RDFa (Resource Description Framework in attributes):
Najbolj zmogljiv a nekoliko bolj zahteven pristop



Semantični splet v spletnih straneh

Mikrooblike

Vrste

- hCard: za predstavitev oseb in organizacij. Največ uporabljan.
- hCalendar: Označevanje dogodkov
- XFN (XHTML Friends Network): označevanje odnosov z drugimi ljudmi
- XMDP (XHTML Meta Data Profiles): formaliziranje novih mikrooblik
- adr: označevanje poštних naslovov
- geo: označevanje geolokacij
- Druge: hAtom, hListing, hMedia, hNews

<http://microformats.org> /



Semantični splet v spletnih straneh

Mikrooblike

hCard

Nastal je na osnovi dobro znanega VCARD (Visit Card) zapisa podatkov oseb in organizacij[14]. Tako lahko hCard pretvorimo v VCARD, to pa je oblika, ki jo podpira npr. tudi Microsoft Outlook (stiki). To pomeni, da lahko enega ali več hCard zapisov iz spletne strani preko dodatka (add-on) za brskalnik (npr. Operator za Firefox) direktno uvozimo med stike v Outlooku (lahko preskusite na www.verdonik.com).

Primer mikrooblike hCard za osebo:

```
1.<div align="left" id="hcard-Ivan-Verdonik" class="vcard">
2.
3.<span class="fn n">Ivan Verdonik</span>
4.<div> <a class="url" href="http://www.verdonik.com"> Domaca stran</a></div>
5.<div class="org">UM FG</div>
6.<a class="email" href="mailto:ivanverdonik@hotmail.com">Nabiralnik</a>
7.<div class="geo">
8.<abbr class="latitude" title="46.532724">46&deg;31'57"N</abbr>,
9.<abbr class="longitude" title="15.628513">15&deg;37'42"E</abbr>
10.</div>
11.<div class="adr">
12.<div class="street-address">Pot na okope 7</div>
13.<span class="postal-code">2000</span>
14.<span class="locality">Maribor
15.<div class="country-name">Slovenija</div>
16.</div>
17.<div class="tel">041 865 501</div>
18.</div>
```



Semantični splet v spletnih straneh

Mikropodatki

itemscope
itemtype
itemprop

Primer mikropodatkov:

```
1.<!DOCTYPE html><html lang="en">
2.<head><meta charset=utf-8><meta name="author" content="Ivan Verdonik">
3.<meta name="description" content="primer mikropodatkov">
4.<meta name="keywords" content="microdata, mikropodatki">
5.<title>Primer mikropodatkov</title></head><body>
6.<h2 align="center">Oglejmo si mikropodatke</h2>
7.<div itemscope itemtype="http://schema.org/Person">
8.<div itemprop="name">Ivan Verdonik</div>
9.<span itemprop="givenName">Ivan</span>
10.<span itemprop="familyName">Verdonik</span> <br>sem
11.<span itemprop="jobTitle">raziskovalec</span><br>
12.<a href="http://verdonik.com" itemprop="affiliation">Janko</a>
13.<div>Moja Slika:</div>
14.<p>Moj naslov: </p> Ivan Verdonik
15.<div itemprop="address" itemscope itemtype="http://schema.org/PostalAddress">
16.  <div itemprop="streetAddress">Pot na okope 7</div>
17.  <span itemprop="postalCode">2000</span>
18.  <span itemprop="addressLocality">Maribor</span></div>
19.</div> </body></html>
```



Semantični splet v spletnih straneh

RDFa

vocab typeof property

Primer HTML 5 dokumenta z RDFa 1.1:

```
1.<!DOCTYPE html>
2.<html> <head><title>HTML5 RDFa 1.1 primer</title>
3.</head><body>
4.<div vocab="http://schema.org/" typeof="Person">
5.<div property="name">Ivan Verdonik</div>
6.<span property="givenName">Ivan</span>
7.<span property="familyName">Verdonik</span> <br>sem
8.<span property="jobTitle">raziskovalec</span><br>
9.<a href="http://verdonik.com" property="affiliation">Janko</a>
10.<div>Moja Slika:</div>
11.<p>Moj naslov: </p> Ivan Verdonik
12.<div property="address" typeof="PostalAddress">
13.    <div property="streetAddress">Pot na okope 7</div>
14.    <span property="postalCode">2000</span>
15.    <span property="addressLocality">Maribor</span></div></div>
16.</body></html>
```



Semantični splet v spletnih straneh

GRDDL

Zajem semantičnih podatkov iz spletnih strani v RDF skladišče

Primer z GRDDL in mikroobliko XFN:

```
1.<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
2.<head profile="http://gmpg.org/xfn/11">
3.<title>Ivan prijatelji</title>
4.<link rel="transformation" href="http://www.w3.org/2003/12/rdf-in-xhtml-xslts/grokXFN.xsl"/>
5.</head><body><h3>Test</h3>
6.<a href="http://lektoriranje.net" rel="sibling">Darinka</a>
7.<a href="http://daniela.um.si" rel="friend met colleague">Daniela</a>
8.<a href="http://verdonik.com" rel="me"> Ivan</a>
9.</body></html>
```

```
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/transformation> <http://www.w3.org/2003/12/rdf-in-xhtml-xslts/grokXFN.xsl> .
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/sibling> <http://lektoriranje.net> .
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/friend> <http://daniela.um.si> .
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/met> <http://daniela.um.si> .
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/colleague> <http://daniela.um.si> .
<http://www.verdonik.com/grddl/xfngrddltest.html> <http://www.verdonik.com/grddl/me> <http://verdonik.com> .
```



Semantični splet v spletnih straneh

Zaključek

Med pristopi vpeljave semantike v spletne strani so mikrooblike še vedno najbolj razširjene, saj imajo same več kot polovični delež. Mikropodatki imajo sicer zaenkrat le 5% delež, a se hitro širijo, posebej ker so vgrajeni v HTML 5. RDFa ima približno 25% delež, je pa med vsemi najbolj razširljiv in zmogljiv. Vsi trije najpomembnejši spletni iskalniki (Google, Yahoo, Bing) znajo razbrati in višje razvrščajo strani, ki vsebujejo semantične podatke. Poleg tega za najbolj razširjene spletne brskalnike obstajajo dodatki (Add on), ki omogočajo ekstrakcijo semantičnih podatkov (zaenkrat predvsem za mikrooblike).



Hvala za vašo pozornost !

Vprašanja?

Pripombe?

Predlogi?