

WEB PODATAKA (WEB OF DATA)

Jelena Jovanović

Email: jeljov@gmail.com

Web: <http://jelenajovanovic.net>

Web dokumenata – nedostaci

WWW 2015 FLORENCE, ITALY

24TH INTERNATIONAL
WORLD WIDE WEB CONFERENCE
Florence 18 - 22 May 2015

HOME ORGANIZATION CALLS PARTICIPANTS COLOCATED EVENTS SPONSORS/EXHIBITION

Important Dates

Research track

Abstract submission November 3

Follow us on Facebook

WWW 2015 FLORENCE, ITALY Like

Fokus je na prezentaciji sadržaja na način što pogodniji ljudima, kroz

- * definisanje načina prikaza informacija,
- * postavljanje linkova ka srodnim sadržajima

Semantika sadržaja je dostupna ljudima, ali ne (direktno) i programima...

Web dokumenata – nedostaci

The image shows a screenshot of the WWW 2015 website with several annotations. At the top, the header features the WWW 2015 logo and the text "24TH INTERNATIONAL WORLD WIDE WEB CONFERENCE Florence 18 - 22 May 2015". A callout box on the right points to this header with the text "vreme održavanja konferencije". Below the header is a large image of Florence, Italy, at night. A callout box on the right points to this image with the text "WWW konferencija će se održati 2015. god u Firenci, Italija". Below the image is a navigation menu with links: HOME, ORGANIZATION, CALLS, PARTICIPANTS, COLOCATED EVENTS, SPONSORS/EXHIBITION. A callout box on the right points to this menu with the text "Informacije o slanju radova". Below the navigation menu is a section titled "Important Dates" with a table of submission deadlines. A callout box on the left points to this section. To the right of the "Important Dates" section is a Facebook widget with the text "Follow us on Facebook" and a "Like" button. Below the Facebook widget is a social media post from WWW 2015 dated November 3 at 5:43am, with the text "Please remember that the abstracts deadline is tonight".

vreme održavanja konferencije

WWW konferencija će se održati 2015. god u Firenci, Italija

Informacije o slanju radova

Important Dates	
Research track	
Abstract submission	November 3
Full Paper submission	November 10
Acceptance notification	January 17

Follow us on Facebook

WWW 2015 FLORENCE, ITALY Like

WWW 2015 FLORENCE, ITALY November 3 at 5:43am

Please remember that the abstracts deadline is tonight

Obilje informacija dostupnih ljudima

Web dokumenata – nedostaci

Informacije
raspoložive
programima

The image shows a screenshot of a website with significant text corruption. The text is rendered as a series of nonsensical symbols and characters, likely due to a font or encoding error. The background features a banner for the '24TH INTERNATIONAL WORLD WIDE WEB CONFERENCE' (WWW 2015) held in Florence, Italy, from May 18-22, 2015. The website layout includes a navigation bar with 'HOME' and 'ORCID' links, a main content area with a 'Follow us on Facebook' section, and a footer with an 'Acceptance notification' dated January 17. The overall appearance is that of a broken or corrupted web page.

Web dokumenata – nedostaci

Razumljiv programima, koliko je nama razumljiv:

本店名菜 Hausspezialitäten	
42 德國豬腳 (本道菜限量供應) 先將去皮豬腳用多種香料醃製，烤後淋上黑生啤酒配以新鮮馬鈴薯沙拉上桌	Fr. 27.50
43 烤牛肉 (180 克) 上等牛肉塊，配以獨家醬汁烤後淋上香蒜咖哩醬，配以米飯上桌	Fr. 29.80
41 香煎豬排 (本道菜限量供應) 先以多種香料醃製，之後用油煎，配以新鮮馬鈴薯沙拉上桌	Fr. 24.50
40 脆皮炸梭鱸魚 (220 克，先用特殊醬料醃製然後烤脆) 配以水煮馬鈴薯檸檬片及塔塔醬 (一種用美乃滋做成的甜酸醬)	Fr. 27.50
附餐若選煎馬鈴薯餅或炸薯條 and vegetables	則另 Fr. 3.20
44 市長寶劍 (400 克烤小牛肉串，可供 2 人食用) 以劍狀的鐵叉叉著肉串，然後烤製而成，配香蒜咖哩醬及烤肉醬 配以大份新鮮沙拉，或者煎馬鈴薯餅或炸薯條上桌	Fr. 78.00

Slika preuzeta sa:
<http://goo.gl/YRs7Yq>

Web dokumenata – nedostaci

Osnovna ideja: omogućiti programima razumevanje Web sadržaja time što će se ti sadržaji opisati mašinski razumljivim jezikom



Slika preuzeta sa:

<http://chaxiubao.typepad.com/photos/uncategorized/pb060002.JPG>

Web dokumenata - karakteristike

- Dizajniran za: **direktno korišćenje od strane ljudi**
- Primarni objekti: **dokumenti i multi-medija**
- Stepen strukturiranosti objekata: **prilično nizak**
- Linkovi: **između dokumenata (ili njihovih delova)**
- Semantika sadržaja i linkova: **implicitna**
- Analogija: **globalni fajl sistem**

Web dokumenata – izazovi

- Izazov #1: Integracija podataka iz različitih izvora
 - Pružanje informacija kroz integraciju (i obradu) podataka iz različitih, najčešće heterogenih izvora (aplikacija, sajtova)
 - Heterogenost se odnosi na različite oblike strukturiranja i organizacije podataka, i/ili oblike u kojima se podaci čine dostupnim
 - Npr.: prikaži mi sve Web sadržaje (blog postove, poruke sa diskusionih foruma, online tutorijale i sl.) koji bi mogli biti relevantni za problem koji rešavam; selektuj samo sadržaje čiji su autori stručni za dati problem

Današnji Web - izazovi

▪ Izazov #2: Razvoj naprednih servisa

– Npr. servis za preporuku ho(s)tela

- Preporuči mi ho(s)tele u Lisabonu koji su bili visoko rangirani od strane ljudi koji imaju isti “stil putovanja” kao ja

[pri tome, uzmi u obzir podatke iz što većeg broja izvora na Web-u]

Današnji Web - izazovi

- Izazov #3: Kreiranje adaptivnih RSS feeds
 - selekcija izvora na osnovu opisa karakteristika izvora (tema, pouzdanost, aktuelnost sadržaja i sl), a *ne* isključivo direktnim definisanjem URL-a (RSS feed-a) izvora
 - filtriranje proizvoljnim, ad-hoc definisanim skupom kriterijuma
 - Npr., prikaži mi sve blog postove čiji su autori moji omiljeni blogeri, i to samo one postove koji su imali više od 5 komentara
- Izazov #4: Ad-hoc, dinamičko kreiranje mash-ups
 - Npr. [FreebaseMaps](#)

Web podataka

- Glavne karakteristike:
 - Podaci (na Web-u) su **strukturirani** i **interlinkovani**
 - **Semantika** podataka i linkova je **eksplicitno data**
 - Omogućeno je izvršavanje **složenih upita** nad više izvora
 - Vizija Web-a kao jedne **gigantske globalne baze podataka**

Web podataka – primer pretrage

- Koji evropski gradovi imaju dobru kolekciju radova Caravaggio-a?
 - ...i pri tome, do njih postoji direktan let iz mog grada?
 - ...i to avio kompanijom koja je ocenjena kao dobra ili odlična?
 - ...i to uzimajući u obzir moje ocene i/ili ocene mojih prijatelja?



Slika preuzeta sa:

<http://ceo.ca/caravaggio-and-the-myth-of-the-boring-painter/>

Web podataka

- Dizajniran za: ljude koje opslužuju mašine
- Analogija: globalna baza podataka
- Primarni objekti: stvari (*things*) i opisi stvari
- Linkovi: između stvari
- Stepen strukturiranosti (opisa) stvari: visok
- Semantika sadržaja i linkova: eksplicitna

Web dokumenata vs. Web podataka

	Web dokumenata	Web podataka
Dizajniran za	Ljude	Ljude koje opslužuju programi
Primarni objekti	Dokumenti (uključujući multimediju)	Resursi* i opisi resursa
Linkovi između	Dokumenata	Resursa
Stepen strukturiranosti	Prilično nizak	Visok
Semantika sadržaja i linkova	Implicitna	Eksplicitna
Analogija	Globalni fajl sistem	Globalna baza podataka

* Npr., Beograd je resurs, a njegov opis je [DBpedia opis Beograda](#)

Web podataka

Preporuka:

Pogledati video **What is Linked Data?**

raspoloživ na: http://www.youtube.com/watch?v=4x_xzT5eF5Q

Zašto? Zato što će pomoći da još bolje razumete koncept Web-a podataka i linkovanja podataka na Web-u

Web podataka - izazovi

- Ograničen javni pristup podacima, a time i integracija podataka iz različitih izvora
 - Pun pristup podacima najčešće moguć samo u okviru jedne Web aplikacije/sajta i to uz eksplicitno odobrenje i/ili plaćanje
 - Tehnološka rešenja su raspoloživa, ali (poslovni/politički) interesi stvaraju prepreke
 - Podaci postaju predmet sve intenzivnije trgovine (data markets)

Web otvorenih/dostupnih podataka

■ Open Data Inicijativa

- Inicijativa svetskih razmera da se podaci učine javno dostupnim na Web-u,
- Inicijativa prisutna u najrazlicitijim domenima
 - Politika, ekonomija, kultura, umetnost, obrazovanje,...



Web otvorenih/dostupnih podataka

- Open Data inicijativa

- U New York-u je 2012. donet zakon (“Local Law 11 of 2012”) koji zahteva da

- “all internal city data that can be made public has to be made public”*

- takodje, u zakonu stoji da

- “The city’s open data has to be published in machine-readable format”*



Semantički Web

- Naredni korak u evoluciji Web-a podataka
- ‘Inteligentni’ Web
 - Rezonovanje nad podacima integrisanim iz različitih izvora
 - Sistemi za preporuku na nivou čitavog Web-a
 - Inteligentni agenti vrše pretragu i preporuku sadržaja
 - ...

Scientific American, May 2001:



THE SEMANTIC WEB

A new form of Web content
that is meaningful to computers
will unleash a revolution of new abilities

by
TIM BERNERS-LEE,
JAMES HENDLER and
ORA LASSILA

<http://www.sciam.com/article.cfm?id=the-semantic-web>

The Semantic Web in 6 minutes

Otvorene tehnologije i standardi

Web-a podataka i Semantičkog Web-a



User Interface & Applications

Trust

Proof

Unifying Logic

Query:
SPARQL

Ontology:
OWL

Rule:
RIF

RDFS

Crypto

Data interchange:
RDF

JSON-LD,
Turtle, N3, XML

URI/IRI

**Semantic Web
Technology
stack**

Neki interesantni tekstovi na temu Web-a podataka

- Semantics in the Real World: Where to Begin?
 - a set of four articles that
 - starts with the characteristics a use case should have to be a good target for semantic technology application,
 - continues to provide surveys of semantic technology applications on the Web and in the enterprise
- Examples of benefits offered by Semantic Web technologies
 - Why Most Job Sites Miss the Boat (blog post at LinkedIn)
 - Startup Pitch: Sépage sees semantic as the answer
 - Sports are the Semantic Focus In Britain at The BBC...
 - Interview: Oracle on Data on the Web
 - Poderopedia Uses Public Data to Take on Powerful Interests in Chile

(Anonimni) upitnik za vaše kritike,
komentare, predloge:

<http://goo.gl/cqdp3l>