

# WEB PODATAKA (WEB OF DATA)

Jelena Jovanović

Email: [jeljov@gmail.com](mailto:jeljov@gmail.com)

Web: <http://jelenajovanovic.net>

# Današnji Web - problemi

**WORLD WIDE WEB 2012 LYON - FRANCE**

16<sup>th</sup> - 20<sup>th</sup> April 2012

Home About Committee Submit Attendees Program Exhibition Sponsors Media Dates

**Keynote speakers announcement :**

**Tim Berners-Lee**

A graduate of Oxford University, Tim Berners-Lee invented the World Wide Web, an internet-based hypermedia initiative for global information sharing while at CERN, the European Particle Physics Laboratory, in 1989. He wrote the first web client and server in 1990. His specifications of URIs, HTTP and HTML were refined as Web technology spread.

looking for something ?

**Platinum Sponsors**

**ONLY LYON** GREATER LYON

**Gold Sponsors**

**YAHOO!**

**Organizers**

With support of

With support of

**Early birds will be closed on Feb. 5th.**

**W4A conference**  
co-located 16th - 17th April.

Lists of accepted **tutorials** and **workshops** are now published with their schedules

**Twitter - about www2012**

charoy about 1 hour ago  
Not a good news @fabien\_gandon:  
#www2012 selection has been very competitive only 12% of papers 108 acceptations out of 885 submissions.

**Calls for submissions: Last update Dec. 15th.**

**Calls for contributions are being published and are regularly updated.**

Please visit the **Submit** top button and its sub-menus, and/or access directly to **the full list of calls for papers**

**Papers camera ready deadline changed from February 28th to February 26th**

**Web Science call for papers updated**

**Contribute to www2012 organisation:**

- Become one of our (foreign) correspondents!
- Become one of our local volunteer!

**Join (or become part of) our**

With support of

**idci**

Omogućeno je:  
definisane načina  
prikaza  
informacija,  
postavljanje linkova  
ka srodnim  
sadržajima

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

# Današnji Web - problemi

The screenshot shows the homepage of the World Wide Web 2012 Lyon - France conference website. At the top left is the logo featuring a stylized lion's head and the text 'WORLD WIDE WEB 2012 LYON - FRANCE'. A pink banner across the top reads '16<sup>th</sup> - 20<sup>th</sup> April 2012'. Below the logo is a navigation menu with links: Home, About, Committee, Submit, Attendees, Program, Exhibition, Sponsors, Media, Dates. The main content area is divided into several sections:

- Keynote speakers announcement:** A yellow box containing a photo of Tim Berners-Lee and a text block describing his role in inventing the World Wide Web. An arrow points from the word 'keynote' to this section.
- Early birds will be closed on Feb. 5th.** A pink box.
- W4A conference** co-located 16th - 17th April. A blue box.
- Lists of accepted tutorials and workshops** are now published with their schedules. A yellow box.
- Twitter - about www2012** A pink box with a tweet from 'charoy' about the selection process.
- Contribute to www2012 organisation:** A blue box with two bullet points: 'Become one of our (foreign) correspondents!' and 'Become one of our local volunteer!'.
- Join (or become part of) our** A yellow box at the bottom.
- Calls for contributions are being published and are regularly updated.** A pink box with text about submitting papers and a link to 'the full list of calls for papers'. An arrow points from this section to the text 'informacije o radovima' on the right.
- Platinum Sponsors:** ONLY LYON GREATER LYON.
- Gold Sponsors:** YAHOO!.
- Organizers:** A logo for the organizing institution.
- With support of:** The French Republic logo (Liberté • Égalité • Fraternité) and the Ministry of Higher Education and Research.
- With support of:** idci logo.

datum konferencije

WWW konferencija se ove godine održava u Lionu, Francuska

Informacije dostupne ljudima

informacije o radovima

# Današnji Web - problemi

16<sup>th</sup> - 20<sup>th</sup> April 2012

The image shows a screenshot of the World Wide Web 2012 website, which is heavily obscured by a dense, chaotic overlay of various symbols, characters, and icons. The background text is mostly illegible due to this overlay. Visible elements include the event logo at the top left, navigation links such as 'Home', 'Attendees', 'Program', 'Exhibition', 'Sponsors', 'Media', and 'Dates', and several content sections with titles like 'Keynote speakers announcement', 'W4A conference', and 'Calls for contributions'. A search bar is also visible on the right side of the page.

Informacije  
dostupne  
programima

# Današnji Web - problemi

Razumljiv mašinama, 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>

# Današnji Web - problemi

Ideja je mašinama omogućiti razumevanje sadržaja



Slika preuzeta sa:

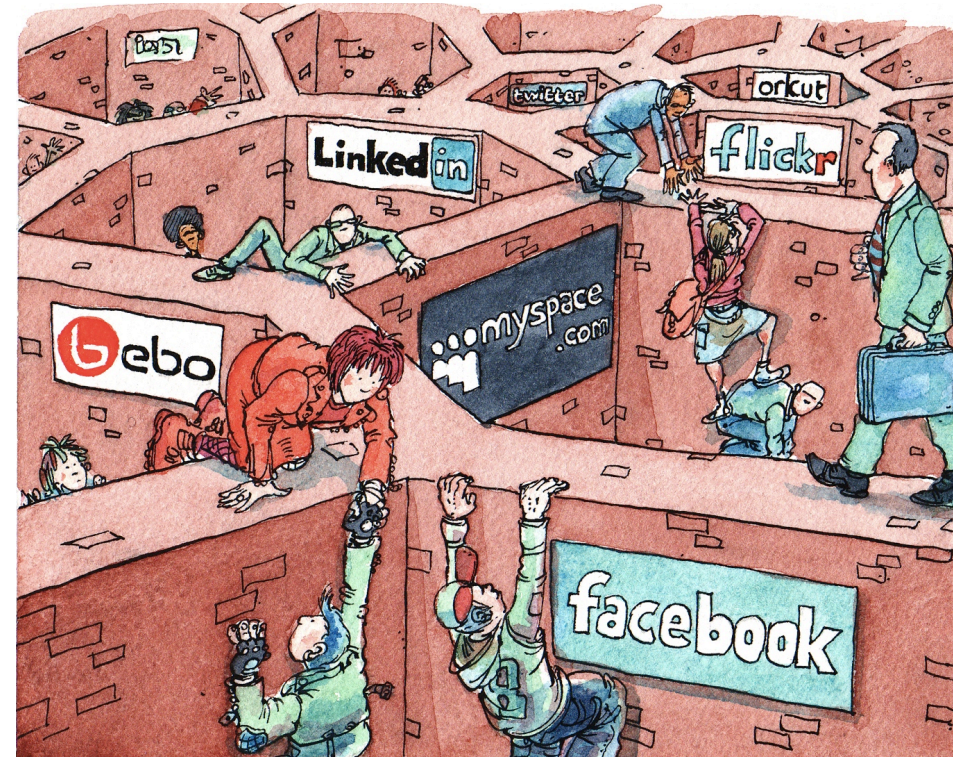
<http://chaxiubao.typepad.com/photos/uncategorized/pb060002.JPG>

# Današnji Web - 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**

# Današnji Web - problemi

- Fragmentacija podataka na Web-u:
  - Dostupni samo u okviru Web sajta/aplikacije
  - Nizak nivo ili potpuno odsustvo povezivanja podata iz razlicitih izvora (aplikacija/sajtova)
  - “Wall gardens”, “Data silos”



Slika preuzeta sa:

<http://www.economist.com/node/10880936>



# Današnji Web - problemi

- Današnji Web je mesto gde su...

- ...računari zaduženi za *prezentaciju* sadržaja (jednostavno)

- ...ljudi zaduženi za *integrisanje* i *interpretiranje* sadržaja (teško)

Zašto ne bi računari preuzeli  
da rade teži deo posla?

# Današnji Web - izazovi

## ▪ Izazov 1: Integracija podataka

- Realizacija upita koji zahtevaju integrisanje podataka iz različitih izvora
  - Npr.: prikaži mi sve blog postove mojih poslovnih kontakata koji su relevantni za studiju na kojoj trenutno radim

# Današnji Web - izazovi

## ▪ Izazov 2: Razvoj naprednih servisa

– Npr. servisa za preporuku ho(s)tela

- Prikazi 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
  - Iz bilo kog izvora, filtrirani proizvoljnim 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

# Web podataka

- Web podataka (Web of Data)  
je vizija Web-a kao jedne  
*gigantske globalne baze 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

# Web podataka – primer pretrage

- Koji evropski gradovi imaju najvišu koncentraciju 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

	<b>Web dokumenata</b>	<b>Web podataka</b>
Dizajniran za	Ljude	Ljude koje opslužuju mašine
Primarni objekti	Dokumenti (i njihovi delovi)	Resursi i opisi resursa
Linkovi između	Dokumenata	Resursa
Stepen strukturiranosti (opisa) stvari	Prilično nizak	Visok
Semantika sadržaja i linkova	Implicitna	Eksplicitna
Analogija	Globalni fajl sistem	Globalna baza podataka



# Web podataka

Preporuka:

Pogledati video **What is Linked Data?** raspoloživ na:

[http://www.youtube.com/watch?v=4x\\_xzT5eF5Q](http://www.youtube.com/watch?v=4x_xzT5eF5Q)

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

# Web podataka – Otvoreni podaci

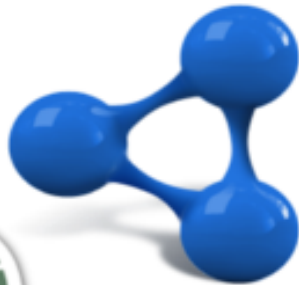
## ■ Open Data Inicijativa

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

**london.gov.uk**

**DATA.GOV**

open  
dataTO



Proyecto  
**aporta**  
Realizador de la Información  
del Sector Público



**europaana**



**open data.**

Open Knowledge Foundation

# Web podataka – Otvoreni podaci

- 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

PHOTO CREDIT HERE

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

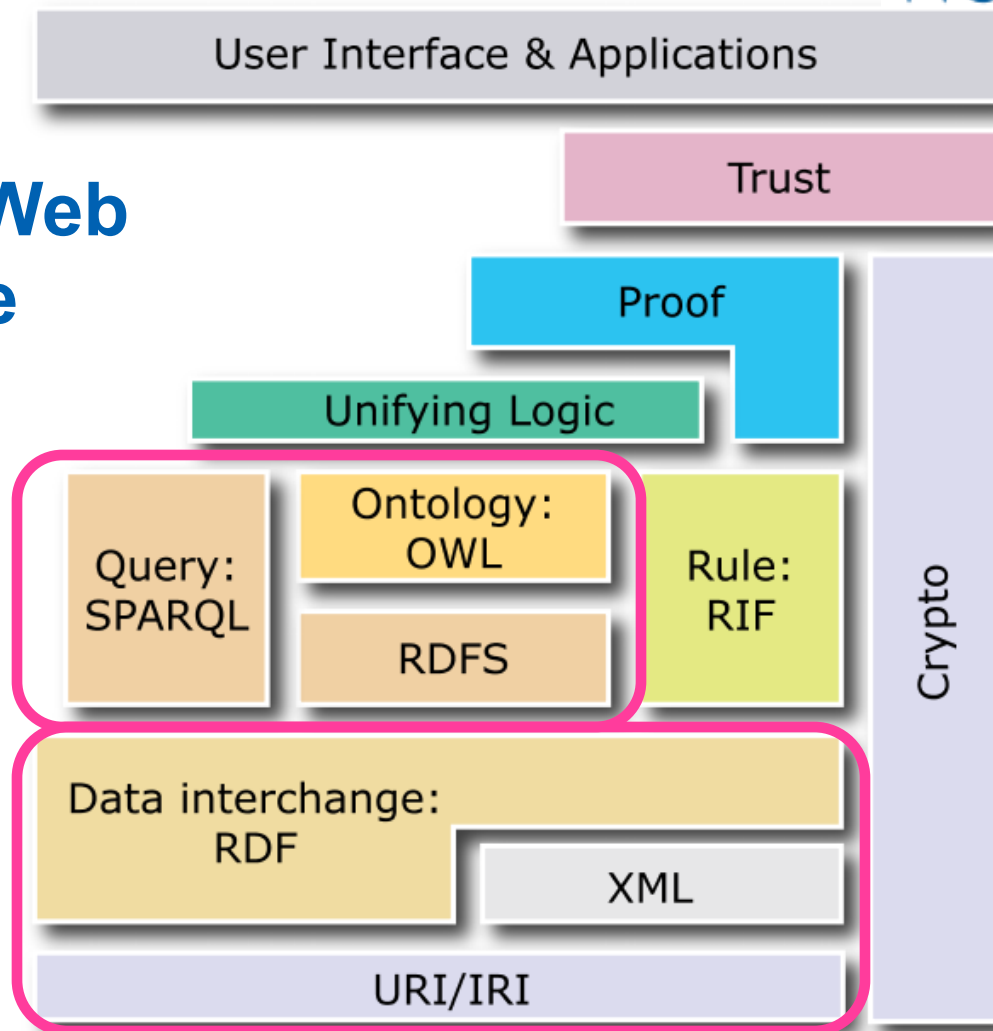
# The Semantic Web in 6 minutes

# Tehnologije Semantičkog Web-a

- Tehnologije i otvoreni standardi



## Semantic Web Layer Cake



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

<http://goo.gl/cqdp3l>