From Gore to Obama
New Global Web governance & Linked Data

Michalis Vafopoulous

26/10/2010

WS.01 Web science: lecture 3
Course outline

• Web history, Epistemology, didactics
• Research methodology
• New Web governance & Linked Data
• Project management
• Web Economics and Business
• Web & the Law
• Privacy & trust in the Web
• Web and Psychology
• Web based development
New global Web governance
The Mission of the World Wide Web Foundation

• Advance the Web
• Connect Humanity
• Empower People
W3F: Objectives

- Accelerate number of Web content creators
- Increase availability of life-critical services
- Enable everyone to access and benefit from the Web — regardless of culture, language, literacy, disability, economic status or devices

**1 Web: Free and Open**
- Same content, available to all
- Core technologies are developed openly, and available at no cost
- Communication, creativity, commerce expand on top of free/open Web

**Useful Content**
- Understand the human-system dynamics of the Web
- R&D new technologies — advance the Web
- Support stable evolution

**Usable by All**
Objectives & Activities

Objectives to achieve mission

One Web that is Free and Open

Understand, Improve, Strengthen

Usable by All People

Useful Content and Services

Activities to be funded and coordinated

Community Building

Web Science

Web Standards

Tools

Training

Deployment
Web-based development

We will talk in detail during next lectures for

• Personal **grid workspace** (g-work) [free processing power for every citizen] – *demand side*

• Now we focus on Linked Data- *the supply side*
From Al Gore (internet highways) to Barack Obama (linked data)
The Web of Documents

• Analogy: a global file system
• Designed for: human consumption
• Primary objects: documents
• Links between: documents (or sub-parts of)
• Degree of structure in objects: fairly low
• Semantics of content and links: implicit (humans)

(Tom Heath)
The Web of Documents

• Simple, big and unstructured
• Organized in Silos

But humans are interested in:
• Things, no documents and
• these Things might be in documents or elsewhere
• Humans: Limited capacity to extract meaning...
Limited SEARCH capacity

Search for: Football Players who went to the University of Texas at Austin, played for the Dallas Cowboys as Cornerback

(Juan F. Sequeda)
Google, Bing, yahoo! irrelevant

Headup: football player ut austin dallas cowboy cornerback

Dallas Cowboys wary of containing opening night emotions | News ...
"That's the only advantage in football." Cowboys Injury report: Quarterback Tony Romo ...
Briefly: The names of players in the Cowboys' Ring of Honor will be uncovered during the game. ... Jenkins looks to have won job at right cornerback for Dallas Cowboys ... 3201 Steck Avenue, Austin, Tx 78757 ... ABC News at UT ...
www.kvue.com/.../football/cowboys/.../091909dnspccowbriefs.306d21b.html -
Cached - Similar -

Recruiting Blog | Sports News | News for Dallas, Texas | Dallas ...
Jeffcoat's father, former Dallas Cowboys defensive lineman Jim Jeffcoat, ... that he plans to be at the Texas Tech at Texas game Saturday night in Austin. .... The first day a football player in the Class of 2010 can sign a national .... Lancasier cornerback Tyler Stephenson said he orally committed to Baylor, ... recruitingblog.dallasnews.com/ - Cached - Similar -

Barking Carnival — Blog — Orlando Scandrick Hates Roy Williams and ...
Dallas Cowboys cornerback and possible member of Al Qaeda, ... of the Forever Beloved Non-Longhorn Players fan club at the University of Texas-Austin. ... he jumps up before they get off the field and he is a very good football player ... barkingcarnival.com/.../orlando-scandrick-hates-roy-williams-and-the-dallas-cowboys/ -
Cached - Similar -

Remember that UT-Texas Tech game? You're not the only one
AUSTIN — One more time, Texas players were forced to revisit the play they ... If that's not enough to motivate you, then you're not a Texas football player. ... Dallas Cowboys cornerback Terence Newman (41) is chased by a member of the ... topics.treehugger.com/article/00Tk9Gj792cJm - Cached - Similar -

DallasCowboys.com | Team | Bios | Players
Acams, Flozell, Anderson, Deon, Austin, Miles, Ball, Alan .... Became just the third rookie cornerback in club history to start a season opener against Arizona (9/7) .... 2005-2007 Brain Awareness Week, hosted by UT Dallas. ... Dallas Cowboys Draft Day Fan Appreciation. NFL
Wikipedia through LD: relevant

Tom Landry
From Wikipedia, the free encyclopedia

Thomas Wade "Tom" Landry (September 11, 1924 – February 12, 2000) was an American football player and coach. He is legendary for his successes as the coach of the Dallas Cowboys. He is ranked as one of the greatest and most innovative coaches in NFL history. He created many new formations and methods. He invented the now popular 4-3 defense, and the "flex defense" system made famous by the "Doomsday Defense" squads he created during his tenure with the Dallas Cowboys.

Landry won 2 Super Bowl titles (VI, XII), 5 NFC titles, 13 Divisional titles, and compiled a 270-178-6 record, the 3rd most wins of all time for an NFL coach. His 20 career playoff victories are the most of any coach in NFL history. He was named the NFL Coach of the Year in 1966 and the NFC Coach of the Year in 1975.

His most impressive professional accomplishment is his record for coaching the Dallas Cowboys to 20 consecutive winning seasons (1966-1985), an NFL record that remains unbroken and unchallenged. It remains one of the longest winning streaks in all of professional sports history.

Contents [hide]
1 Early life, World War II
2 NFL playing career
3 NFL coaching career
   3.1 The Great Innovator
4 Beyond the NFL
5 Tom Landry in popular culture
6 Quotations
7 References
8 External links

Early life, World War II

Born to Ray (an auto mechanic and volunteer fireman) and Ruth Landry in Mission, Texas, Landry was the second of four children (Robert, Tommy, Ruthie and Jack). After playing quarterback (primary passer and runner, and also punter) for Mission High School (including leading his team to a 12-0 record his senior season),[1] he attended the University of Texas in Austin as an industrial engineering major; but interrupted his education after a semester to serve in the United States Army Air Forces during World War II. Landry earned his wings and a commission as a Second Lieutenant at Lubbock Army Air Field, and was assigned to the 483d Bombardment Group at RAF Debden, England, as a B-17 Flying Fortress bomber co-pilot in the 860h Bomb Squadron. From November 1944 to April 1945, he completed a combat tour of 30 missions, and survived a crash landing in Belgium after his bomber ran out of fuel.
The Web of Data

• Analogy: a global filesystem ----> **global database**
• Designed for: human consumption --> **machines first, humans later**
• Primary objects: documents --> **things** (or descriptions of things)
• Links between: documents --> **things**
• Degree of structure in objects: fairly low ---> **high**
• Semantics of content and links: implicit --> **explicit**

(Tom Heath)
The Modigliani Test

• Show me all the locations of all the original paintings of Modigliani
• Daniel Koller (@dakoller) showed that you can find this with a SPARQL query on DBpedia

Thanks Richard MacManus - ReadWriteWeb
SPARQL Explorer for http://dbpedia.org/sparql

SPARQL:
```sparql
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdfs: <http://www.w3.org/2001/XMLSchema#>
PREFIX dbpedia: <http://dbpedia.org/resource/>
PREFIX dbpedia2: <http://dbpedia.org/property/>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX do: <http://purl.org/dc/elements/1.1/>
PREFIX do: <http://dbpedia.org/resource/>
PREFIX do: <http://dbpedia.org/property/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>

SELECT * WHERE { 
  ?s foaf:depiction ?img . 
  ?s dbpedia2:museum ?m . 
  ?m rdfs:label ?ml . 
  FILTER (lang(?sl) = "en") . 
  FILTER (lang(?ml) = "en") 
}
```

Results: Browse Go! Reset

SPARQL results:

<table>
<thead>
<tr>
<th>sl</th>
<th>ml</th>
<th>img</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Ship of Fools (painting)&quot;@en</td>
<td>&quot;Louvre&quot;@en</td>
<td><a href="http://upload.wikimedia.org/wikipedia/commons/a/2/Jheronimus_Bosch_011.jpg">http://upload.wikimedia.org/wikipedia/commons/a/2/Jheronimus_Bosch_011.jpg</a></td>
<td>Paris</td>
</tr>
<tr>
<td>&quot;The Last Judgment (Bosch triptych fragment)&quot;@en</td>
<td>&quot;Alte Pinakothek&quot;@en</td>
<td><a href="http://upload.wikimedia.org/wikipedia/commons/d/d9/TheLastJudgementBoschTriptychFragment.jpg">http://upload.wikimedia.org/wikipedia/commons/d/d9/TheLastJudgementBoschTriptychFragment.jpg</a></td>
<td>Munich</td>
</tr>
</tbody>
</table>
Results of the Modigliani Test

- Atanas Kiryakov from Ontotext
- Used LDSR – Linked Data Semantic Repository
  - Dbpedia
  - Freebase
  - Geonames
  - UMBEL
  - Wordnet

Published April 26, 2010:
### SPARQL Query

Results for `PREFIX fb: <http://rdf.free...>` (8)

<table>
<thead>
<tr>
<th>painting_1</th>
<th>owner_1</th>
<th>city_fb_con</th>
<th>city_db_loc</th>
<th>city_db_clt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jeanne Hébuterne@en</td>
<td>Barnes Foundation@en</td>
<td>Philadelphia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reclining Nude@en</td>
<td>Museum of Modern Art</td>
<td>New York City</td>
<td>Manhattan</td>
<td></td>
</tr>
<tr>
<td>Portrait of Diego Rivera@en</td>
<td>The São Paulo Museum of Art@en</td>
<td></td>
<td>São Paulo</td>
<td></td>
</tr>
<tr>
<td>Portrait of a Woman@en</td>
<td>School of the Art Institute of Chicago@en</td>
<td></td>
<td></td>
<td>Chicago</td>
</tr>
<tr>
<td>Head@en</td>
<td>Museum of Modern Art</td>
<td>New York City</td>
<td>Manhattan</td>
<td></td>
</tr>
<tr>
<td>Woman with a Necklace@en</td>
<td>School of the Art Institute of Chicago@en</td>
<td></td>
<td></td>
<td>Chicago</td>
</tr>
<tr>
<td>Madam Pompadour@en</td>
<td>School of the Art Institute of Chicago@en</td>
<td></td>
<td></td>
<td>Chicago</td>
</tr>
<tr>
<td>Anna Zborowska@en</td>
<td>Museum of Modern Art</td>
<td>New York City</td>
<td>Manhattan</td>
<td></td>
</tr>
</tbody>
</table>
The Web of Data: why?

- encourages reuse
- reduces redundancy
- maximises its (real and potential) interconnectedness
- enables network effects to add value to data
The Web of Data: how?

– current state on the Web
  • Relational Databases
  • APIs
  • XML
  • CSV
  • XLS

(see EXHIBIT)

Computers can’t consume data because:
  • Different formats & models
  • Not inter-connected
The Web of Data: how?

– we need to create a standard way of publishing Data on the Web (like HTML for docs)

This is the Resource Description Framework (RDF)

(a simple example here from Juan F. Sequeda), more next semester!)
Resource Description Framework (RDF)

• A data model
  – A way to model data
  – Inspired form Relational databases and Logic
• RDF is a triple data model
• Labeled Graph (semantic networks)
• Subject, Predicate, Object

<Isidoro> <was born in> <Chios>
<Chios> <is part of> <Greece>
Example: Document on the Web

Programming the Semantic Web

Build Flexible Applications with Graph Data
By Toby Segaran, Colin Evans, Jamie Taylor
Publisher: O'Reilly Media
Released: July 2009
Pages: 304

Write a Review

Description

With this book, the promise of the semantic web -- in which machines can find, share, and combine data on the Web -- is not just a technical pos
Web demonstrates several ways to implement semantic web applications, u: technologies recently introduced. Each chapter walks you through a single p:solve real problems.

Full Description
Databases back up documents

THINGS have PROPERTIES:
A Book as a Title, an author, ...

<table>
<thead>
<tr>
<th>Isbn</th>
<th>Title</th>
<th>Author</th>
<th>PublisherID</th>
<th>ReleasedData</th>
</tr>
</thead>
<tbody>
<tr>
<td>978-0-596-15381-6</td>
<td><em>Programming the Semantic Web</em></td>
<td>Toby Segaran</td>
<td>1</td>
<td>July 2009</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

This is a THING:
A book title “Programming the Semantic Web” by Toby Segaran, ...

<table>
<thead>
<tr>
<th>PublisherID</th>
<th>PublisherName</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>O’Reilly Media</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Data representation in RDF

<table>
<thead>
<tr>
<th>Isbn</th>
<th>Title</th>
<th>Author</th>
<th>PublisherID</th>
<th>ReleasedData</th>
</tr>
</thead>
<tbody>
<tr>
<td>978-0-596-15381-6</td>
<td><em>Programming the Semantic Web</em></td>
<td>Toby Segaran</td>
<td>1</td>
<td>July 2009</td>
</tr>
</tbody>
</table>

PublisherID | PublisherName  | 1 | O’Reilly Media |
Everything on the web is identified by a URI!
link the data to other data

http://.../isbn978

- title: Programming the Semantic Web
- author: Toby Segaran
- isbn: 978-0-596-15381-6
- publisher: O’Reilly

http://.../publisher1

- name: publisher1
consider the data from Revyu.com
start to link data
Juan Sequeda publishes data too

http://juansequeda.com/id

livesIn

name

http://dbpedia.org/Austin

Juan Sequeda
Let’s link more data

- ISBN: 978
- Reviewer: Juan Sequeda
- Review: Awesome Book
- Website: www.juansequeda.com/id
- Location: Austin

The diagram shows the relationships between these entities.
And more

- Name: Juan Sequeda
- Has Reviewer: Toby Segaran
- Publisher: O'Reilly
- Lives In: "http://dbpedia.org/Austin"
- Previous book title: Programming the Semantic Web
Linked data = internet + http + RDF
Linked Data Principles

1. Use URIs as names for things
2. Use URIs so that people can look up (dereference) those names.
3. When someone looks up a URI, provide useful information.
4. Include links to other URIs so that they can discover more things.
Web as a database

Linked Data makes the web exploitable as ONE GIANT HUGE GLOBAL DATABASE!

Is there any query language like sql? SPARQL... (next semester)
History of LD

- Linked Data Design Issues by TimBL July 2006
- Linked Open Data Project WWW2007
- First LOD Cloud May 2007
- 1st Linked Data on the Web Workshop WWW2008
- 1st Triplification Challenge 2008
- How to Publish Linked Data Tutorial ISWC2008
- BBC publishes Linked Data 2008
- 2nd Linked Data on the Web Workshop WWW2009
- NY Times announcement SemTech2009 - ISWC09
- 1st Linked Data-a-thon ISWC2009
- 1st How to Consume Linked Data Tutorial ISWC2009
- Data.gov.uk publishes Linked Data 2010
- 2nd How to Consume Linked Data Tutorial WWW2010
- 1st International Workshop on Consuming Linked Data COLD2010
- El.dbpedia.org
Oct 2007
Feb 2008
Sept 2008

As of September 2008
July 2009
What is a Linked Data application/service?

Software system that makes use of data on the Web from multiple datasets and that benefits from links between the datasets.
Characteristics of Linked Data Applications

• **Consume** data that is published on the web following the Linked Data principles: an application should be able to request, retrieve and process the accessed data.

• **Discover** further information by **following the links** between different data sources: the fourth principle enables this.

• **Combine** the consumed linked data with data from sources (not necessarily Linked Data)

• **Expose the combined data** back to the web following the Linked Data principles

• **Offer value** to end-users
the 5 stars of open linked data

★ make your stuff available on the Web (whatever format)
★★ make it available as structured data (e.g. excel instead of image scan of a table)
★★★★ non-proprietary format (e.g. csv instead of excel)
★★★★★ use URLs to identify things, so that people can point at your stuff
★★★★★★ link your data to other people’s data to provide context

http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/
Examples

- Data.gov.uk
- Swisstrain.ch
- LOC
- Greece – not in Linked Data!
- Fuelprices.gr
- E-prices.gr
- Et.diavgeia.gov.gr
More Examples

• http://data-gov.tw.rpi.edu/wiki
• http://dbrec.net/
• http://fanhu.bz/
• http://data.nytimes.com/schools/schools.html
• http://sig.ma
• http://visinav.deri.org/semtech2010/
Our project

Greek Linked Data Cloud

– Ideas for Linked Data Services
Describe and rationalize the idea of a proposed Service (public or commercial) for the Greek LD cloud.

Deadlines

• 28-1-2011 presentation (class: 5 min)
## Specifications

<table>
<thead>
<tr>
<th>WP</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP0</td>
<td>Description and Objectives</td>
</tr>
<tr>
<td>WP1</td>
<td>Review of Current state in Greece and other countries (UK, US, AUS etc)</td>
</tr>
<tr>
<td>WP2</td>
<td>Requirements analysis</td>
</tr>
<tr>
<td>WP3</td>
<td>Datasets</td>
</tr>
<tr>
<td>WP4</td>
<td>Create RDF, LD and Links</td>
</tr>
<tr>
<td>WP5</td>
<td>Architectural design for proposed ontologies and Services</td>
</tr>
<tr>
<td>WP6</td>
<td>Implementation for proposed ontologies</td>
</tr>
<tr>
<td>WP7</td>
<td>Linked Data Services (web &amp; mobile web)</td>
</tr>
<tr>
<td>WP8</td>
<td>Evaluation of Linked Data Services (quantitative &amp; qualitative)</td>
</tr>
</tbody>
</table>
Example

WP0 Description & Objectives
Create the Greek Linked Open Commerce and link to local and international resources in order to provide efficient, structured, reusable and free information to consumers.
Review of Current state in Greece and other countries

Theory


• A Framework for Investigating Mobile Web Success in the Context of E-commerce etc

Applications

http://linkedopencommerce.com/
### Review table based on important criteria

<table>
<thead>
<tr>
<th>STUDIES</th>
<th>DEFINITIONS</th>
<th>DATA</th>
<th>METHODOLOGY</th>
<th>RESULTS</th>
</tr>
</thead>
</table>
| 1       | Jansen & Pooch (2002) [12] | **Session**: the entire sequence of queries entered by a searcher  
**Query**: string of zero or more characters entered into the Web IR system  
**Term**: string of characters separated by some delimiter (space, colon or a period) | Fireball Search Engine:  
1-31 July 1998  
16,252,902 queries  
Excite Search Engine:  
10 March 1997  
54,573 queries  
Alta Vista Search Engine:  
2 August-13 September 1998  
993,208,159 queries | Transaction log Analysis  
Statistical analysis | -Query length: 2 terms  
-Two queries per session  
-No complex query syntax  
-View of no more than 10 documents from the results list  
-Almost no use of Boolean operators, |
| 2       | Wolfram, Spink | 200,000 Excite users  
1,000,000 queries | Query Log Analysis  
Compare and contrast | -Move towards greater simplicity, shorter queries and shorter sessions with little modification of terms in subsequent queries |
Datasets

• Upload at el.ckan.net
• Real and valid hypotheses (use star system)
  • frontistiria prices ★
  • idiotika ekpedeutiria ★
  • e-prices.GR ★
• antikimenikes axies (no star)
• dimotika teli (no star)
• Cost of ISP, mobile phones (no star)
• Cost of travelling (no star)
Interesting Research issues

• search: distributed queries
• Interlinking Algorithms
• LD economics and business
• Institutional engineering for LD provision
• Provenance and Trust
• Evaluation