

Διασυνδεδεμένα Ανοικτά Δεδομένα (Linked Open Data -LOD)

Μ.Στεφανιδάκης

26-3-2013

Εισαγωγή στα Σημασιολογικά Δεδομένα

Το βασικό μοντέλο οργάνωσης των Σημασιολογικών Δεδομένων

Από τον Σημασιολογικό Ιστό στα Διασυνδεδεμένα Δεδομένα

Χρησιμοποιώντας Διασυνδεδεμένα Ανοικτά Δεδομένα

Ανοικτά Θέματα και Έρευνα

Η αρχή: Το όραμα του Σημασιολογικού Ιστού

Tim Berners-Lee, James Hendler and Ora Lassila, “The Semantic Web”, Scientific American, May 2001, p. 29-37.

...The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users...

...the effectiveness of such software agents will increase exponentially as more machine-readable Web content and automated services become available...

...this structure will open up the knowledge and workings of humankind to meaningful analysis by software agents, providing a new class of tools by which we can live, work and learn together.

Σημασιολογία – “semantics”

- ▶ Τι **σημαίνουν** τα δεδομένα;
 - ▶ Ποια η **έννοιά** τους;
- ▶ **Μετάδοση** έννοιας και ενσωμάτωση στη **γνώση**..
 - ▶ ..έχοντας ως αποτέλεσμα την κατάλληλη **δράση**!
- ▶ Πώς θα γίνει αντιληπτή από τις “**μηχανές**”;
 - ▶ Ός άνθρωποι, χρησιμοποιούμε εύκολα την **έμμεση** σημασιολογία

Έμμεση Σημασιολογία

Παράδειγμα: ονόματα συναρτήσεων ή μεταβλητών

```
quicksort :: (Ord a) => [a] -> [a]
quicksort [] = []
quicksort (x:xs) =
    let smallerOrEqual = [t|t<=xs,t<=x]
        greater = [t|t<=xs,t>x]
    in quicksort smallerOrEqual ++ [x] ++ quicksort greater
```

- ▶ Τι καταλαβαίνετε;
 - ▶ Ανεξάρτητα από το αν γνωρίζετε τη γλώσσα!
 - ▶ Τα σχόλια θα βοηθούσαν περαιτέρω στη μεταφορά σημασιολογίας..
- ▶ Τι μπορεί όμως να καταλάβει η “μηχανή”;

Ιστοσελίδα και μηχανή

Τι “αντιλαμβάνεται” η μηχανή:

```
<!doctype html>
<html>
  <head>
    <meta charset="utf-8">
    <title>????????????????</title>
  </head>

  <body>
    <h1>????????????????</h1>
    ?????????????????
    ?????????????????
  </body>
</html>
```

- ▶ Προσοχή όμως: αν και δεν “αντιλαμβάνεται” το νόημα του περιεχομένου του τίτλου, η μηχανή ξέρει την έννοια “τίτλος”!
 - ▶ π.χ. έχει ιδιαίτερο βάρος στις μηχανές αναζήτησης

Κλασσικά Web Δεδομένα

- ▶ Στο κλασσικό Web, τα δεδομένα είναι “εσωτερική υπόθεση” του παραγωγού τους
 - ▶ Αποθηκευμένα σε βάσεις δεδομένων απροσπέλαστες από τρίτους (“data silos”)
 - ▶ Μετασχηματίζονται από τον ίδιο τον παραγωγό σε ιστοσελίδες
 - ▶ Ή χρησιμοποιούνται από web εφαρμογές που έχει γράψει ο ίδιος ο παραγωγός
 - ▶ Δεν χρειάζεται μετάδοση της σημασιολογίας των δεδομένων (είναι γνωστή εκ των προτέρων!)

Κλασσικά Web Δεδομένα (2)

- ▶ Μπορούν όμως τα κλασσικά Web δεδομένα να χρησιμοποιηθούν από τρίτους;
 - ▶ Σε εφαρμογές που δεν γνωρίζει ο παραγωγός των δεδομένων;
 - ▶ Από εφαρμογές που δεν σχεδιάστηκαν ειδικά για τα δεδομένα αυτά;
- ▶ Αυτός είναι ο στόχος των Σημασιολογικών Δεδομένων

Σημασιολογικά Δεδομένα στο Web

- ▶ Πριν από κάθε άλλο, θα πρέπει φυσικά
 - ▶ τα δεδομένα να είναι απευθείας προσπελάσιμα στο Web
 - ▶ σε μια μορφή εύκολα επεξεργάσιμη από τις μηχανές
- ▶ Το κυριότερο όμως είναι να είναι ρητά δηλωμένη και προσπελάσιμη και η σημασιολογία των δεδομένων
 - ▶ Μόνο τότε θα μπορούν να χρησιμοποιηθούν από εφαρμογές που προσαρμόζουν τη συμπεριφορά τους ανάλογα με το τι σημαίνουν τα δεδομένα
 - ▶ εκτελώντας ενέργειες πάνω σε σελτ δεδομένων για τα οποία δεν σχεδιάστηκαν εξ'αρχής

Σημασιολογία και Μεταδεδομένα

- ▶ Μεταδεδομένα:
 - ▶ “δεδομένα για/σχετικά με/που αφορούν τα δεδομένα”
 - ▶ Είναι ακριβώς εκείνα που μεταδίδουν τη σημασιολογία των κανονικών δεδομένων
- ▶ Στις εφαρμογές του Σημασιολογικού Ιστού
 - ▶ Τα μεταδεδομένα μεταδίδονται ρητά μαζί με τα δεδομένα
 - ▶ Έχουν την ίδια αξία και είναι το ίδιο απαραίτητα με τα δεδομένα

Εισαγωγή στα Σημασιολογικά Δεδομένα

Το βασικό μοντέλο οργάνωσης των Σημασιολογικών
Δεδομένων

Από τον Σημασιολογικό Ιστό στα Διασυνδεδεμένα
Δεδομένα

Χρησιμοποιώντας Διασυνδεδεμένα Ανοικτά Δεδομένα

Ανοικτά Θέματα και Έρευνα

Ποιο το κατάλληλο μοντέλο δεδομένων;

- ▶ Τα σημασιολογικά δεδομένα εκτίθενται ως έχουν (raw)
 - ▶ Ποιο είναι το κατάλληλο μοντέλο οργάνωσης;
 - ▶ Και πώς θα εμπεριέχονται και τα μεταδεδομένα;
- ▶ Ας ξεκινήσουμε με ένα απλό μοντέλο: δεδομένα σε μορφή πίνακα
 - ▶ Η σημασιολογία των δεδομένων προκύπτει **έμμεσα από τη θέση τους** (γραμμή-στήλη)

Τύπος	Αριθμός
Λεωφορεία	58.519
ΙΧ	2.062.880
Οδοποιητικά	7.642
Εκχιονιστικά	6
Ποδήλατα	8.540.192

Η έμμεση σημασιολογία του πίνακα

- ▶ Κάθε γραμμή αντιστοιχεί σε μία βασική οντότητα (entity) δεδομένων
 - ▶ Έμμεση πληροφορία
- ▶ Κάθε στήλη αντιστοιχεί σε μια ιδιότητα (attribute)
 - ▶ Ρητή πληροφορία, αν υπάρχει περιγραφή στηλών

Οντότητα _i	→	Τύπος	Αριθμός	↖	Ιδιότητα _j
		Λεωφορεία	58.519		
		ΙΧ	2.062.880		
		Οδοποιητικά	7.642		
		Εκχιονιστικά	6		
		Ποδήλατα	8.540.192		

Προσθήκη ρητού αναγνωριστικού (id)

- ▶ Οικείο σχήμα από τις σχεσιακές βάσεις...
- ▶ Κάθε βασική οντότητα διαθέτει μοναδικό αναγνωριστικό

id	Τύπος	Αριθμός
1	Λεωφορεία	58.519
2	ΙΧ	2.062.880
3	Οδοποιητικά	7.642
4	Εκχιονιστικά	6
5	Ποδήλατα	8.540.192

Έμμεσος Διαχωρισμός Αναγνωριστικών

Γιατί δεν συμπεραίνουμε ότι ο Σμόλικας έχει μέτρια τοξικότητα;

Πίνακας 1: <http://ex.com/mountains>

id	Peak	Height(m)	Mountain range
1	Olympus	2917	Olympus
2	Smolikas	2637	Pindus
3	Kaimaktsalan	2524	Voras

Πίνακας 2: <http://ex.com/plants>

id	Duration	Flower Color	Toxicity
...
2	Perennial	Green	Moderate
...

Έμμεσος Διαχωρισμός Αναγνωριστικών

Έμμεσα αναγνωρίζουμε ότι:

Πίνακας 1: <http://ex.com/mountains>

id	Peak	Height(m)	Mountain range
http://ex.com/mountains#1	Olympus	2917	Olympus
http://ex.com/mountains#2	Smolikas	2637	Pindus
http://ex.com/mountains#3	Kaimaktsalan	2524	Voras

Πίνακας 2: <http://ex.com/plants>

id	Duration	Flower Color	Toxicity
...
http://ex.com/plants#2	Perennial	Green	Moderate
...

Ρητά Αναγνωριστικά

- ▶ Απαραίτητα για τη διασύνδεση οντοτήτων!

id	Label	Model
...
http://my.com/33	My Crane	http://ex.com/38rb
...

id	Engine	Weight(kg)	Tracks(mm)	Boom(m)
...
http://ex.com/38rb	V8	60000	850	18
...

Το μετα-μοντέλο EAV

- ▶ Η τελευταία μεταμόρφωση: το μοντέλο **Entity-Attribute-Value**
 - ▶ Μορφή τριάδας (triple): (Οντότητα, Ιδιότητα, Τιμή)
 - ▶ Όλα τα μεταδεδομένα δηλώνονται ρητά
 - ▶ Ο “εφιάλτης” του σχεσιακού μοντέλου!!!

Entity	Attribute	Value
http://ex.com/38rb	Engine	V8
http://ex.com/38rb	Weight(kg)	60000
http://ex.com/38rb	Tracks(mm)	850
http://ex.com/38rb	Boom(m)	18

Εισαγωγή στα Σημασιολογικά Δεδομένα

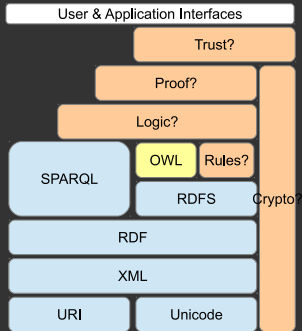
Το βασικό μοντέλο οργάνωσης των Σημασιολογικών
Δεδομένων

Από τον Σημασιολογικό Ιστό στα Διασυνδεδεμένα
Δεδομένα

Χρησιμοποιώντας Διασυνδεδεμένα Ανοικτά Δεδομένα

Ανοικτά Θέματα και Έρευνα

Τα επίπεδα του Σημασιολογικού Ιστού



- ▶ Η αρχική ιδέα για τη δομή του Σημασιολογικού Ιστού
- ▶ Ανάπτυξη πρακτικών προτύπων (κατώτερα επίπεδα)
 - ▶ Για την υλοποίηση των μηχανισμών του “Web of Data”
- ▶ Ακαδημαϊκή έρευνα (ανώτερα επίπεδα)
 - ▶ Συνδυασμός ιδεών Μαθηματικής Λογικής και Βάσεων Γνώσης

Μια (εκ των υστέρων) κριτική ματιά

- ▶ Η πορεία υλοποίησης του Σημασιολογικού Ιστού χωρίστηκε σε δύο κατευθύνσεις:
- ▶ Πρακτικές δράσεις ανάπτυξης τεχνικών οργάνωσης, αποθήκευσης και αναζήτησης των σημασιολογικών δεδομένων
 - ▶ Το αποτέλεσμα ήταν οι **πρότυπες τεχνολογίες** και ένα σύνολο **καλών πρακτικών** για την οργάνωση των σημασιολογικών δεδομένων, με τη μορφή των **Διασυνδεδεμένων Δεδομένων (Linked Data)**
- ▶ Ακαδημαϊκή έρευνα στους τομείς της Μαθηματικής Λογικής και των Βάσεων Γνώσης
 - ▶ Βασικές έννοιες οι οντολογίες και ο αυτόματος συμπερασμός (reasoning)
 - ▶ Μικρό βάρος στη δυνατότητα πρακτικής υλοποίησης
 - ▶ Σήμερα, μικρό μέρος των εννοιών αυτών χρησιμοποιείται και μάλλον με ad-hoc τρόπο

Διασυνδεδεμένα Δεδομένα (Linked Data)

Οι αρχικές “4 Εντολές” (Tim Berners-Lee, 2006):

1. Χρησιμοποιήστε URIs για να αναγνωρίσετε οντότητες
 - ▶ Εξασφαλίζοντας τη μοναδικότητα των αναγνωριστικών
2. Χρησιμοποιήστε HTTP URIs (δηλαδή URLs)
 - ▶ Για να μπορούν οι άλλοι να προσπελάσουν την πληροφορία σας
3. Στην προσπέλαση, απαντήστε με χρήσιμη πληροφορία σε πρότυπη μορφή (RDF*, SPARQL)
 - ▶ Δώστε δυνατότητα στις μηχανές να ρωτήσουν και να καταλάβουν την απάντηση
4. Διασυνδέστε τα αναγνωριστικά σας URI με άλλα, τρίτων
 - ▶ Έτσι ώστε αυτός που ρωτάει να βρεί και άλλες συμπληρωματικές πηγές πληροφόρησης

Uniform Resource Identifiers (URIs)

- ▶ Υπερσύνολο των διευθύνσεων στο Web (URLs)
 - ▶ Υπάρχουν URIs που δεν αρχίζουν με http://..!
 - ▶ urn:uuid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
 - ▶ urn:nbn:de:eki/ZDB2071735
- ▶ Αλλά **μόνο τα HTTP URIs προσφέρουν μια διεύθυνση αναζήτησης πληροφορίας!**
- ▶ Τα HTTP URIs λύνουν επίσης το πρόβλημα της μοναδικότητας:
 - ▶ Το domain name μέρος της διεύθυνσης έχει ήδη ανατεθεί μονοσήμαντα
 - ▶ στον οργανισμό ή επιχείρηση όπου εργαζόμαστε
 - ▶ στους servers του οποίου θα αναρτήσουμε τη σημασιολογική πληροφορία
- ▶ Προσοχή: το URI **δεν είναι** η οντότητα αλλά **αναφέρεται** στην οντότητα
 - ▶ Η οντότητα μπορεί να είναι οτιδήποτε: άνθρωπος, πράγμα, έννοια, συναίσθημα, κλπ

Αποσαφήνιση: Αναφορές URI (URIrefs)

ο Όλυμπος (η φυσική
οντότητα, όχι η φωτό!)



(σε μορφή
αναγνώσιμη από
τον άνθρωπο και
τη μηχανή)

data
document

πληροφορία για

αναφέρεται σε

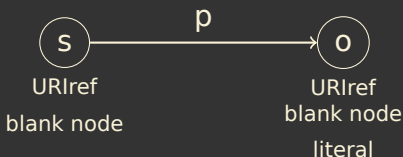
προσπέλαση

http://dbpedia.org/resource/Mount_Olympus

(URIref)

Resource Description Framework (RDF)

- ▶ Ένα πρότυπο για την περιγραφή της οργάνωσης των δεδομένων (data model)
 - ▶ Χρήση **τριάδων** για την περιγραφή **γράφων δεδομένων**
 - ▶ **Subject - Predicate - Object** ή απλά **(s,p,o)**
 - ▶ πολύ κοντά στην απλή φυσική μορφή “υποκείμενο - ρήμα - αντικείμενο”



Resource Description Framework (RDF) (2)

- ▶ **Subject:** Κόμβος του γράφου, είναι
 - ▶ Είτε ένα URI
 - ▶ Είτε **blank node**: κόμβος “χωρίς όνομα”, συμβολίζεται με **_**: και ακολουθεί μια εσωτερική αρίθμηση της βάσης αποθήκευσης
- ▶ **Predicate:** η ετικέτα της ακμής μεταξύ κόμβων, πάντα ένα URI
- ▶ **Object:** Κόμβος του γράφου, είναι
 - ▶ Είτε ένα URI
 - ▶ Είτε blank node
 - ▶ Είτε μια ρητή τιμή (**literal**) με προαιρετικό τύπο δεδομένων ή γλώσσα, π.χ. **"137554788"**, **"Lodovico Giustini"@it** ή **"1685"^^xsd:gYear**
- ▶ Οι κόμβοι διασυνδέονται μεταξύ τους μέσω των predicates, σχηματίζοντας έναν γράφο δεδομένων

Λεξιλόγια Κατηγορημάτων (Vocabularies)

- ▶ Η θεωρητική αντίληψη για την τριάδα (s,p,o):
 - ▶ “οτιδήποτε” – “οτιδήποτε” – “οτιδήποτε”
- ▶ Η πρακτική αντίληψη των Διασυνδεδεμένων Δεδομένων:
 - ▶ “οτιδήποτε” – “μια σταλιά συναίνεση” – “οτιδήποτε”
- ▶ Η συναίνεση εξασφαλίζεται με τη χρήση καθιερωμένων λεξιλογίων στη θέση του κατηγορήματος
 - ▶ Δείτε π.χ. στο <http://lov.okfn.org/dataset/lov/index.html>
 - ▶ Κανείς δεν σας εμποδίζει όμως να εφεύρετε το δικό σας λεξιλόγιο!

Διάσημα Κατηγορήματα από διάσημα Λεξιλόγια

- ▶ **rdf:type**, τύπος ενός κόμβου RDF (RDF)

```
<http://dbpedia.org/resource/Lodovico_Giustini>  
  <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>  
    <http://dbpedia.org/class/yago/ItalianComposers> .
```

- ▶ **rdfs:label**, περιγραφή ενός κόμβου RDF (RDF Schema)

```
<http://dbpedia.org/resource/Lodovico_Giustini>  
  <http://www.w3.org/2000/01/rdf-schema#label>  
    "Lodovico Giustini"@de .
```

- ▶ **owl:sameAs**, ταυτοσημία κόμβων (OWL),
“ακρογωνιαίος λίθος” των Διασυνδεδεμένων
Δεδομένων

```
<http://dbpedia.org/resource/Lodovico_Giustini>  
  <http://www.w3.org/2002/07/owl#sameAs>  
    <http://rdf.freebase.com/ns/m.08v_23> .
```

Διασυνδεδεμένα Ανοικτά Δεδομένα (Linked Open Data)

Οι νέες “5 Εντολές” (Tim Berners-Lee, 2010):

- ▶ ★☆☆☆☆ Διαθέστε τα δεδομένα σας στο Web, σε οποιαδήποτε μορφή, με ανοικτή άδεια
- ▶ ★★☆☆☆ Διαθέστε τα δεδομένα σας σε δομημένη μορφή, έστω και με κλειστό πρότυπο
- ▶ ★★★☆☆ Χρησιμοποιήστε ανοικτά πρότυπα
- ▶ ★★★★☆ Χρησιμοποιήστε HTTP URIs, για να μπορούν οι άλλοι να αναφερθούν στα δεδομένα σας
- ▶ ★★★★★ Διασυνδέστε τα δεδομένα σας με άλλα, τρίτων

Σχετικά με τα Ανοικτά Δεδομένα

- ▶ Χρησιμοποιούνται πρωταρχικά από τον δημόσιο τομέα και τις κυβερνητικές οργανώσεις
 - ▶ Για να προάγουν τη διαφάνεια, την ηλεκτρονική διακυβέρνηση, τη συμμετοχή των πολιτών
 - ▶ Προσοχή: τα Ανοικτά Δεδομένα δεν είναι πάντοτε Διασυνδεδεμένα...
- ▶ Τι συμβαίνει με τις επιχειρήσεις;
 - ▶ Προς το παρόν υποστηρίζουν τα Ανοικτά Δεδομένα
 - ▶ Κυρίως για τη βελτίωση των αποτελεσμάτων τους στις μηχανές αναζήτησης (SEO)
 - ▶ Αλλά επίσης για την καθιέρωση στον τομέα τους (brand name)

Εισαγωγή στα Σημασιολογικά Δεδομένα

Το βασικό μοντέλο οργάνωσης των Σημασιολογικών
Δεδομένων

Από τον Σημασιολογικό Ιστό στα Διασυνδεδεμένα
Δεδομένα

Χρησιμοποιώντας Διασυνδεδεμένα Ανοικτά Δεδομένα

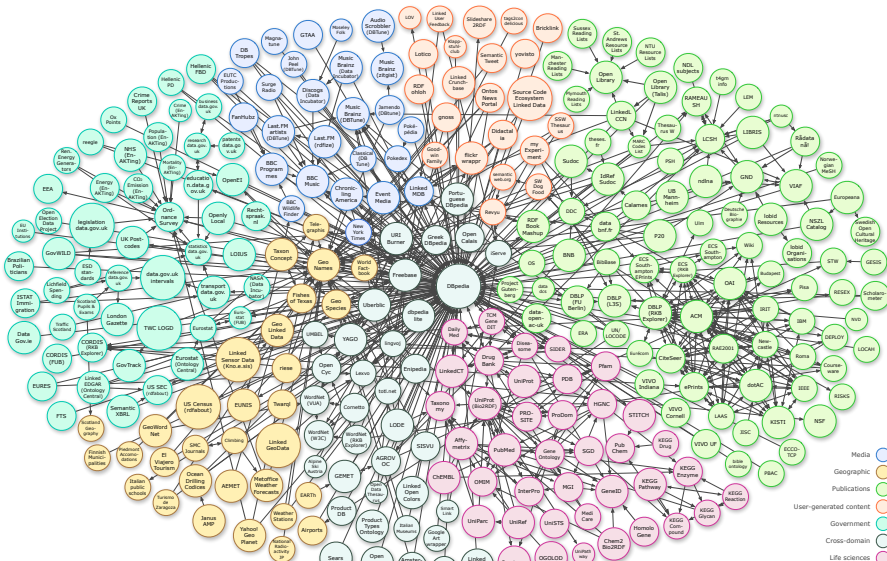
Ανοικτά Θέματα και Έρευνα

Κατανάλωση Διασυνδεδεμένων Ανοικτών Δεδομένων

- ▶ Πώς θα ενημερωθούμε για τα σετ δεδομένων που υπάρχουν;
 - ▶ Υπάρχουν ειδικά sites για τον σκοπό αυτόν
 - ▶ Όπως και πασίγνωστα σετ δεδομένων, όπως π.χ. η DBpedia (το περιεχόμενο της Wikipedia επεξεργασμένο σε μορφή RDF)
- ▶ Πώς θα εξάγουμε την επιθυμητή πληροφορία;
 - ▶ Διαδοχική Προσπέλαση URIs (crawling)
 - ▶ Ερωτήσεις σε SPARQL endpoints
 - ▶ Κατέβασμα και επεξεργασία data dumps (offline επεξεργασία μόνο!)

LOD cloud

Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch. <http://lod-cloud.net/>



The Data Hub

http://datahub.io/

The screenshot shows a web browser window with the URL `http://datahub.io/dataset?q=greece`. The page header includes the Data Hub logo, the tagline "The easy way to get, use and share data", and links for "Search", "Groups", "About", "Login", and "Register".

The main content area is titled "Search - the Data Hub" and features a search bar with the text "greece". Below the search bar, it indicates "53 datasets found".

The search results are listed as follows:

- Regions of Greece**: This file includes the geographical limits of 13 Regions (Eastern Macedonia and Thrace, Central Macedonia, Western Macedonia, Epirus, Thessaly, Ionian Islands, Western Greece, Central...). **41 views**. Formats: SHP, GML, KML, MAP. Buttons: OPEN, DATA.
- Greece - National Statistical Service of Greece**: About National statistics for Greece. Data Topics include: Census Of Population 1991 Censuses 2000-2001 Demography Labour Market Indices... **163 views**. Buttons: OPEN, DATA.
- Greek public wifi**: This is an effort to tag Greek open data. **69 views**. Buttons: OPEN, DATA.
- Diavgeia**: text/csv text/html application/rdf+xml text/plain example/rdf+xml meta/void application/vnd.ms-excel application/vnd.ms-excel **103 views**. Buttons: OPEN, DATA.
- The Clarity program introduces for the first time in Greece the obligation to publish all the decisions on the Internet, with the exception of decisions that contain sensitive personal...** **103 views**. Buttons: OPEN, DATA.
- Hellenic Fire Brigade**: application/rdf+xml example/rdf+xml application/x-ntriples api/sparql meta/sitemap meta/void **601 views**. Buttons: OPEN, DATA.
- The Hellenic Fire Brigade project encompasses efforts to extract valuable information from Greek Open Data originating from the Ministry of Public Order & Citizen Protection and in...** **601 views**. Buttons: OPEN, DATA.

The right sidebar contains two sections:

- Tags**:
 - lod (7)
 - country-greece (7)
 - sidinfo (6)
 - provenance-metadata (5)
 - license-metadata (5)
 - greek (5)
 - format-skos (5)
 - format-rtf (5)
 - published-by-producer (4)
 - no-proprietary-vocab (4)
- Resource formats**:
 - SHP (12)
 - KML (12)
 - GML (12)
 - XLS (8)
 - application/rdf+xml (7)
 - ODS (7)
 - example/rdf+xml (6)
 - meta/void (5)
 - api/sparql (5)
 - text/csv (4)

The bottom section is titled **Groups** and lists:

- Greece (22)
- Linking Open Data Cloud (3)
- Economics Datasets (2)
- The Greek public spending (1)
- OpenSpending (1)
- Open Archaeology (1)
- Linguistic Resources (party)

SPARQL

- ▶ Πρότυπη γλώσσα για ερωτήσεις σε σετ δεδομένων σε μορφή RDF
 - ▶ Οι ερωτήσεις έχουν τη μορφή τριάδων που συγκροτούν ένα βασικό σχέδιο γράφου
 - ▶ Στις τριάδες αυτές κάποια μέρη (s,p ή o) έχουν συγκεκριμένες τιμές
 - ▶ Ενώ άλλα μέρη αποτελούνται από αδέσμευτες μεταβλητές
- ▶ Ο γράφος της ερώτησης ταιριάζει πάνω στον RDF γράφο των δεδομένων
 - ▶ Προκαλώντας την ανάθεση (πιθανώς πολλαπλών) τιμών στις μεταβλητές
 - ▶ Οι τιμές αυτές επιστρέφονται ως απαντήσεις
- ▶ Η SPARQL απαντάται στην έκδοση 1.0 και στη νεώτερη 1.1, με πολλά πρόσθετα χαρακτηριστικά

Παράδειγμα SPARQL (1)

<http://dbtune.org/classical/snorql>

- Ταυτόσημοι με τον συνθέτη Lodovico Giustini

```
SELECT ?o WHERE{  
<http://dbtune.org/classical/resource/composer/giustini\_lodovico>  
  owl:sameAs ?o .  
}
```

(δες)

- Πότε γεννήθηκε ο Lodovico Giustini;

```
SELECT ?x WHERE{  
<http://dbtune.org/classical/resource/composer/giustini\_lodovico>  
  <http://dbtune.org/classical/resource/vocab/birth> ?b .  
  ?b <http://purl.org/dc/elements/1.1/date> ?x .  
}
```

(δες)

Παράδειγμα SPARQL (2)

<http://dbtune.org/classical/snorql>

- Ποιοι γεννήθηκαν το 1685;

```
SELECT ?x WHERE{
  ?x <http://dbtune.org/classical/resource/vocab/birth> ?b .
  ?b <http://purl.org/dc/elements/1.1/date> "1685"^^xsd:gYear .
}
```

(δες)

- Ποιος συνθέτης δεν επηρεάστηκε από άλλους;

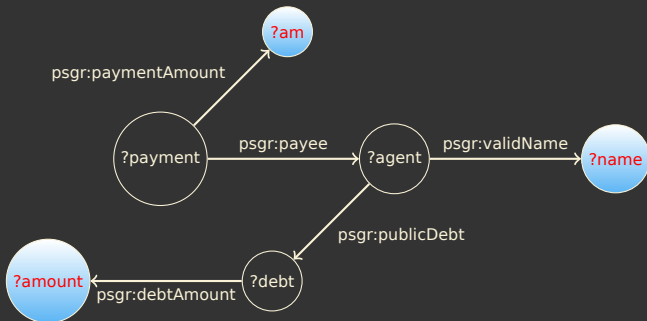
```
SELECT ?o WHERE{
  ?o a classical:Composer.
  OPTIONAL {
    ?s sim:method simmethods:cmn-influence ;
        sim:object ?o;
        sim:subject ?i .
  }
  FILTER (!bound(?i))
}
LIMIT 10
```

(δες)

Παράδειγμα SPARQL 1.1

<http://publicspending.medialab.ntua.gr/endpoint>

```
select ?name (?amount) as ?totalDebt
      (sum(xsd:decimal(?am)) as ?totalPayments)
from <http://publicspending.medialab.ntua.gr/Decisions>
where{
  ?payment psgr:payee ?agent ; psgr:paymentAmount ?am .
  ?agent psgr:publicDebt ?debt ; psgr:validName ?name .
  ?debt psgr:debtAmount ?amount .
}ORDER BY DESC(?totalPayments)
```



Εισαγωγή στα Σημασιολογικά Δεδομένα

Το βασικό μοντέλο οργάνωσης των Σημασιολογικών Δεδομένων

Από τον Σημασιολογικό Ιστό στα Διασυνδεδεμένα Δεδομένα

Χρησιμοποιώντας Διασυνδεδεμένα Ανοικτά Δεδομένα

Ανοικτά Θέματα και Έρευνα

Δεδομένα από πολλαπλές πηγές

- ▶ Όταν συνδυάζονται Διασυνδεδεμένα Ανοικτά Δεδομένα από πολλές πηγές
 - ▶ Πόσο **εμπιστευόμαστε** την συνολική απάντηση;
 - ▶ Πώς εκφράζεται η **αβεβαιότητα** της απάντησης;
 - ▶ Πώς συνυπολογίζεται η **ποιότητα** των δεδομένων;
 - ▶ Τι πρέπει να γίνει σε περίπτωση **αντικρουόμενων** δεδομένων;
 - ▶ Τι αξία έχουν τα δεδομένα που **παλαιώνουν**;
 - ▶ Και ποιός θα τα κρίνει όλα αυτά; Ο τελικός χρήστης;
- ▶ Απαιτούνται μηχανισμοί για τη μεταφορά των συμφραζομένων των ερωτήσεων και τη βαθμονόμηση των διαφορετικών απαντήσεων που προκύπτουν!

Το χάσμα μεταξύ χρήστη και LOD

- ▶ Τα Διασυνδεδεμένα Ανοικτά Δεδομένα είναι ένας συνεκτικός χώρος
 - ▶ Αρκεί να ξέρουμε τι να ρωτήσουμε, για να λάβουμε τις σωστές απαντήσεις
 - ▶ (κάτι που ονομάζεται και “crisp reasoning”)
- ▶ Όμως...
 - ▶ Πώς θα ανακαλύψουμε τα URIs (s,p ή o) για τα οποία ενδιαφερόμαστε;
 - ▶ Ξεκινώντας από περιγραφές π.χ. σε φυσική γλώσσα;
 - ▶ Και πώς θα χειριστούμε το ranking που αναπόφευκτα προκύπτει;
- ▶ Είναι πλέον επιτακτική η ανάγκη της εφαρμογής μεθόδων μηχανικής μάθησης σε συνδυασμό με τις υπάρχουσες τεχνικές των Διασυνδεδεμένων Ανοικτών Δεδομένων!

Αύξηση του βαθμού διασύνδεσης

- ▶ Ο **βαθμός διασύνδεσης** προσδίδει πρόσθετη αξία στα Διασυνδεδεμένα Ανοικτά Δεδομένα
 - ▶ Η διασύνδεση σήμερα είναι **μονόδρομη**: δεν υπάρχει τρόπος να μάθετε ότι κάποιος αναφέρεται σε ένα URI σας!
- ▶ Η ειδοποίηση αναφοράς από τρίτους ισοδυναμεί με την εγκαθίδρυση ανάστροφης διασύνδεσης (**backlink**)
 - ▶ Η διασύνδεση γίνεται πλέον **αμφίδρομη**!
- ▶ Για την ενημέρωση των backlinks απαιτείται ένα πλαίσιο εφαρμογής που επιτρέπει την αυτοματοποιημένη αποστολή και λήψη ειδοποιήσεων κατά τη δημιουργία αναφορών σε ξένα URIs
 - ▶ μια λύση παρουσιάζεται στο <http://swrg.ionio.gr>

Νέες Εφαρμογές

- ▶ Τα Διασυνδεδεμένα Ανοικτά Δεδομένα επιτρέπουν την ανάπτυξη **έξυπνων εφαρμογών**
 - ▶ Δυναμικά προσαρμοζόμενες στα δεδομένα
 - ▶ Σε μορφή mashups, ιδίως σε σχέση με τα γεωγραφικά ανοικτά δεδομένα
 - ▶ Με ενδιαφέρουσα οπτικοποίηση (visualization)
- ▶ Όμως: το **“killer LOD application”** δεν έχει ακόμα κατασκευαστεί...
 - ▶ Τα Διασυνδεδεμένα Ανοικτά Δεδομένα είναι διαθέσιμα
 - ▶ Τα εργαλεία ανάπτυξης είναι διαθέσιμα
- ▶ **Μπορείτε να είστε εσείς οι κατασκευαστές!**
 - ▶ Το μόνο που χρειάζεται είναι φαντασία και δημιουργικότητα