



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΜΣ ΚΥΒΕΡΝΟΑΣΦΑΛΕΙΑ
ΚΑΙ ΕΠΙΣΤΗΜΗ ΔΕΔΟΜΕΝΩΝ

MSc CYBERSECURITY
AND DATA SCIENCE
DEPT OF INFORMATICS
UNIVERSITY OF PIRAEUS

Διαχείριση Μεγάλων Δεδομένων

Big Data Management

Lecture 9 – Big data on the Web

Knowledge Graphs & Linked Open Data

George Papastefanatos (gpapas@athenarc.gr)

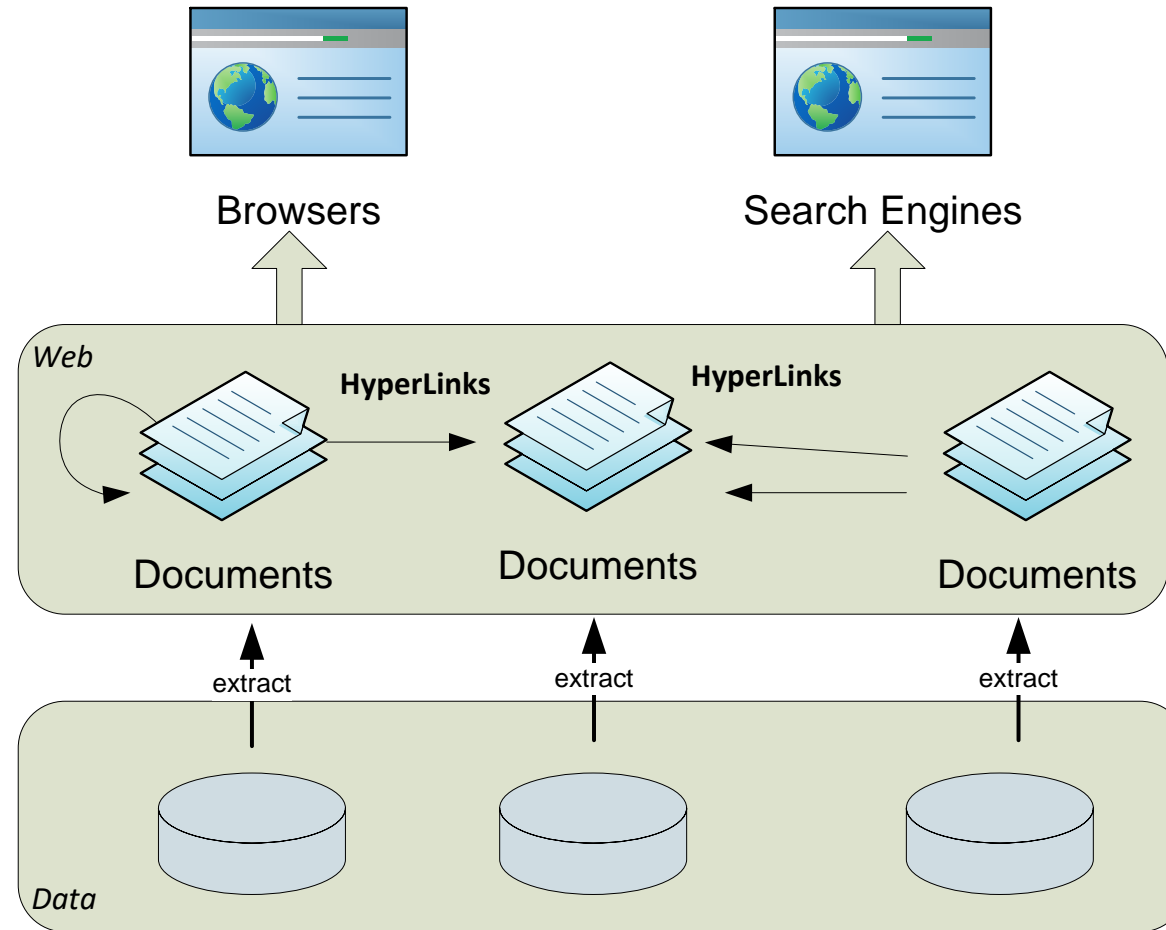
Principal Researcher @ ATHENA Research Center

Περιεχόμενα

- Από το διαδίκτυο εγγράφων στα Διασυνδεδεμένα Δεδομένα και στους Γράφους γνώσης
 - Τι είναι ο Ιστός Δεδομένων
 - Οντολογίες και βασικές Τεχνολογίες.
- Το μοντέλο RDF (Resource Description Framework)
- Η γλώσσα SPARQL
- Ανοιχτά Διασυνδεδεμένα Δεδομένα (Linked Open Data - LOD)
 - Δημοσίευση Πληροφορίας

ΤΟ ΔΙΑΔΙΚΤΥΟ...ΜΕΧΡΙ ΣΗΜΕΡΑ

World Wide Web



Το Διαδίκτυο - Αδυναμίες

- Το μεγαλύτερο μέρος του περιεχομένου του διαδικτύου σήμερα είναι κατάλληλο για χρήση από τον άνθρωπο
- Ακόμη και η δομημένη πληροφορία (π.χ. από μια βάση δεδομένων) χάνεται κατά την παρουσίασή της στο διαδίκτυο
- Η αναζήτηση βασίζεται ακόμα σε λέξεις κλειδιά και κειμενική ανάλυση (π.χ. google)
- Ακόμα και αυτή η φτωχή σε εκφραστικότητα ανάλυση αρκεί και καθιστά το www επιτυχές
- Η πρόσβαση των μηχανών στο περιεχόμενο:
 - είναι χαμηλή
 - παρουσιάζει απουσία σημασιολογίας

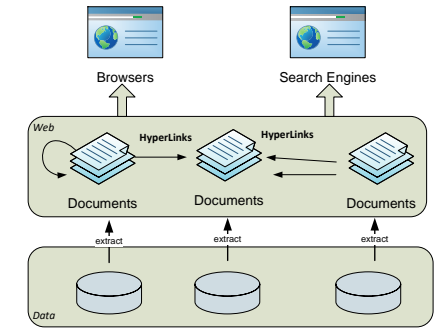
Περιορισμοί & Προβλήματα

- Αναζήτηση πληροφοριών
 - Λέξεις κλειδιά
- Εξαγωγή πληροφοριών
 - Αναγκαία ανθρώπινη παρέμβαση
 - Ομοίως για περιήγηση, ανάκτηση, ερμηνεία
- Συντήρηση πληροφορίας
 - Ασυνέπειες στην ορολογία
 - Μη επικαιροποίηση
- Προβολή πληροφοριών
 - Όψεις

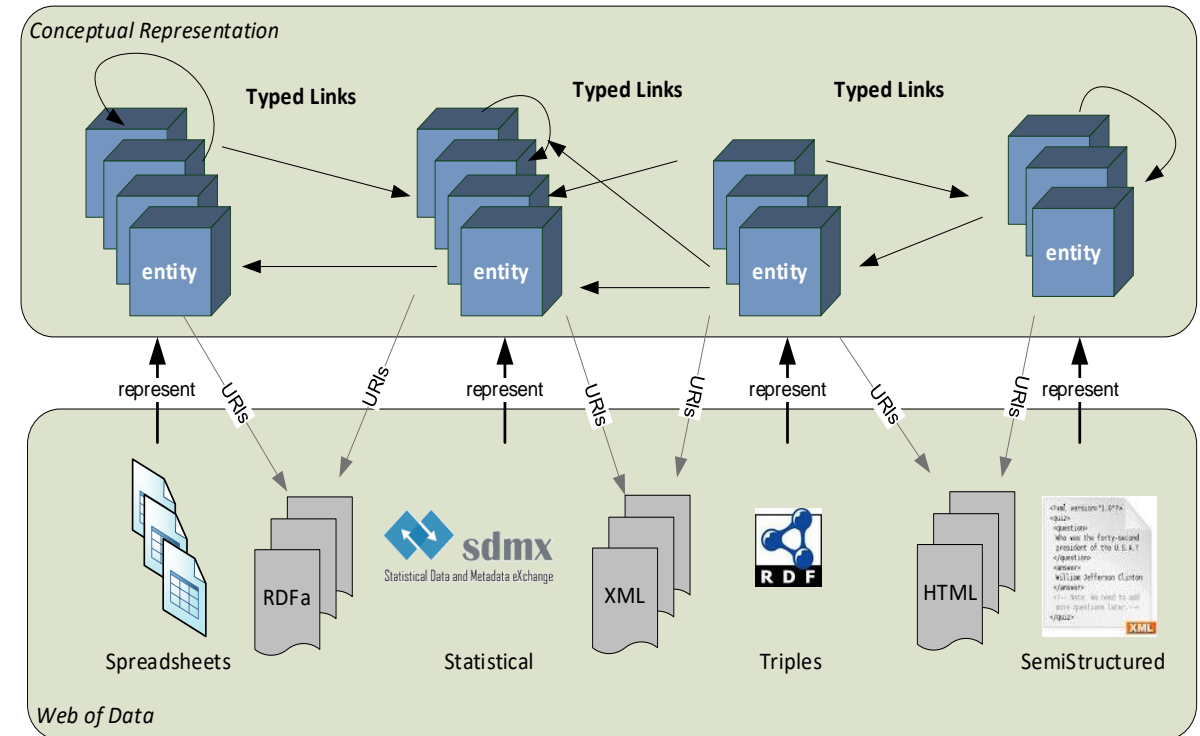
Σημασιολογικός Ιστός

*Από το διαδίκτυο εγγράφων στο **διαδίκτυο γνώσης και δεδομένων***

Web of Documents → Web of Data



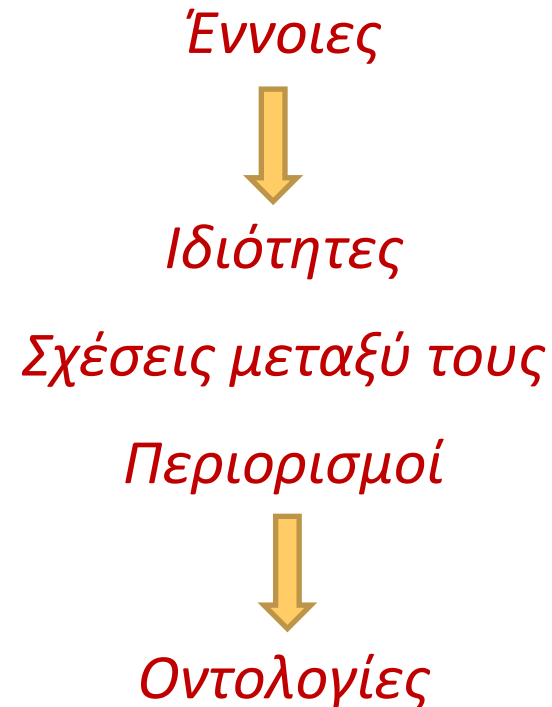
- Σύνδεση δεδομένων από διαφορετικούς τομείς και πηγές
- Βασικά αντικείμενα: Οντότητες και οι περιγραφές τους (και όχι έγγραφα)
- Σημασιολογικές Συνδέσεις (links) μεταξύ οντοτήτων
- Επίπεδο Αναπαράστασης: Από ολόκληρα σύνολα μέχρι ατομικές οντότητες



Adapted from Chris Bizer, Richard Cyganiak, Tom Heath,
available at <http://linkeddata.org/guides-and-tutorials>

Σημασιολογικός Ιστός και Διαχείριση Γνώσης

- Η γνώση οργανώνεται σε **εννοιολογικό** επίπεδο σύμφωνα με το νόημά της

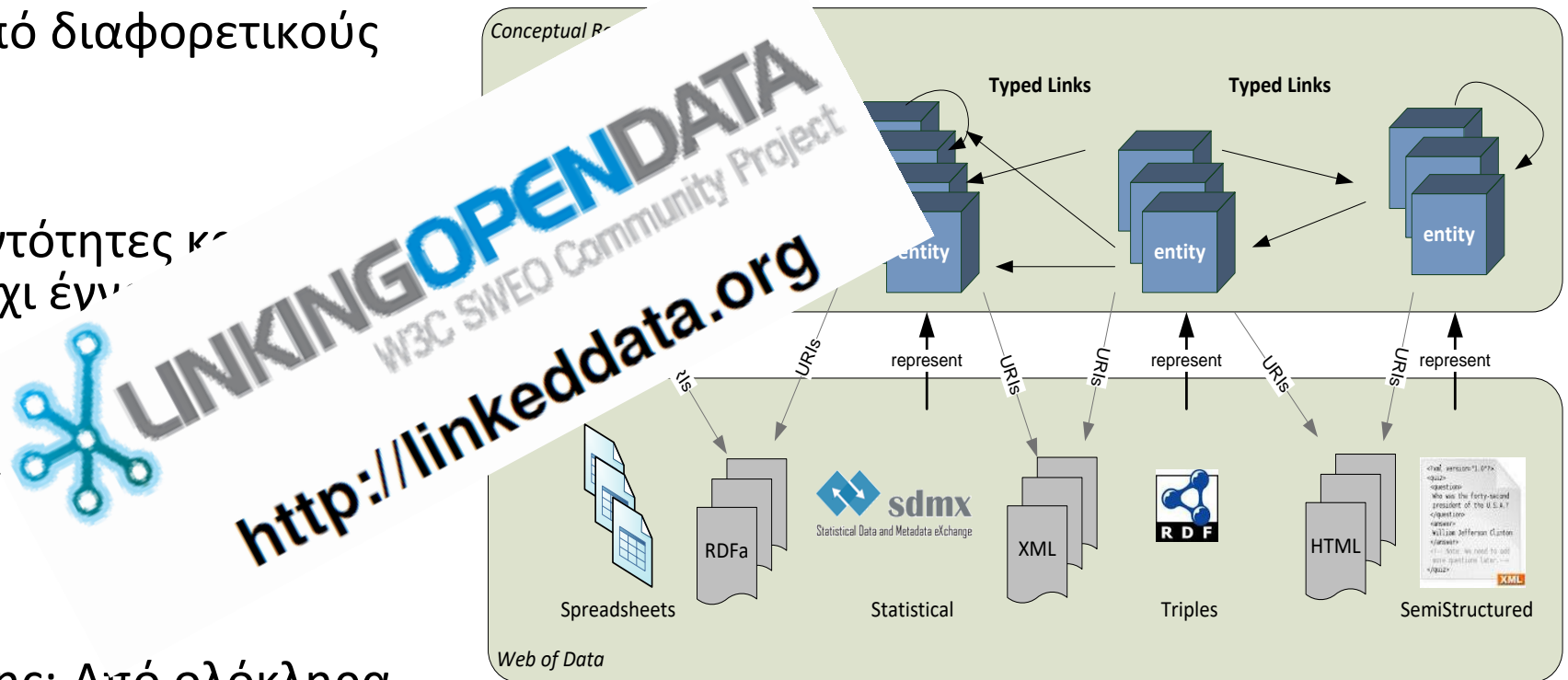


Τι είναι μια οντολογία;

- Τρόπος μοντελοποίησης της γνώσης
- Το σύνολο των εννοιών (π.χ. ορισμοί), οντοτήτων (π.χ. άνθρωπος, αντικείμενο, ψηφιακό τεκμήριο, κτλ) των ιδιοτήτων τους και των σχέσεων μεταξύ τους (περιορισμοί, λογικοί κανόνες, κτλ)
- Κάνουν χρήση προτύπων δεδομένων για την δομική και σημασιολογική μοντελοποίηση όλων των παραπάνω.

Web of Data

- Σύνδεση δεδομένων από διαφορετικούς τομείς και πηγές
- Βασικά αντικείμενα: Οντότητες και τύποι συνδέσεων (και όχι έννοιες)
- Σημασιολογικές Συνδέσεις Οντοτήτων
- Επίπεδο Αναπαράστασης: Από ολόκληρα σύνολα μέχρι ατομικές οντότητες



Adapted from Chris Bizer, Richard Cyganiak, Tom Heath,
available at <http://linkeddata.org/guides-and-tutorials>

URI

http://data.europeana.eu/proxy/provider/2048408/item_LJOJLJDD4YFVDHTC4JUFXB7NVYLIQ540



Title	Lesbos, Mytilini. Museum Theofilos. Bauernhochzeit 1933 Helbig, Konrad (Fotograf) and Chatzimichail, Theofilos	FIN View
Description	Chatzimichail, Theofilos, Lesbos, Mytilini. Museum Theofilos. Bauernhochzeit 1933	CAN No ©
People	Creator: Helbig, Konrad (Fotograf) , Chatzimichail, Theofilos	EXP
Classifications	subject: Malerei , Foto , naive Malerei , Fotos , Painting , Photograph	SHA
Time	Date: 1972 , http://semium.org/time/1972 Period: 1972 Place / Time: Theofilos Museum , Mytilini , http://data.europeana.eu/place/base/83762 , http://data.europeana.eu/place/base/87	
Provenance	Identifier: http://www.deutschefotothek.de/obj71435753.html Institution: Deutsche Fotothek Provider: Deutsche Digitale Bibliothek Providing Country: Germany First Published In Europeana: 2017-02-10 Last Updated In Europeana: 2017-07-27	
References And Relations	Dataset: 2048408_Ag_DE_DDB_SLUB-Sub1	
Location	Place / Time: Theofilos Museum , Mytilini , http://data.europeana.eu/place/base/83762 , http://data.europeana.eu/place/base/87 Current Location: http://d-nb.info/gnd/4012995-0	

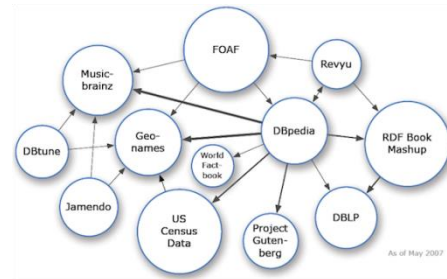
SPARQL for querying

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX edm: <http://www.europeana.eu/schemas/edm/>
PREFIX ore: <http://www.openarchives.org/ore/terms/>
SELECT ?title ?creator ?mediaURL ?date
WHERE {
    ?CHO edm:type "SOUND" ;
        ore:proxyIn ?proxy;
        dc:title ?title ;
        dc:creator ?creator ;
        dc:date ?date .

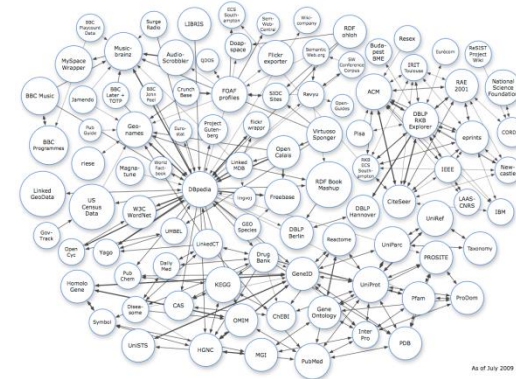
    ?proxy edm:isShownBy ?mediaURL .
}
LIMIT 100
```


Εξέλιξη των Linked Open Data

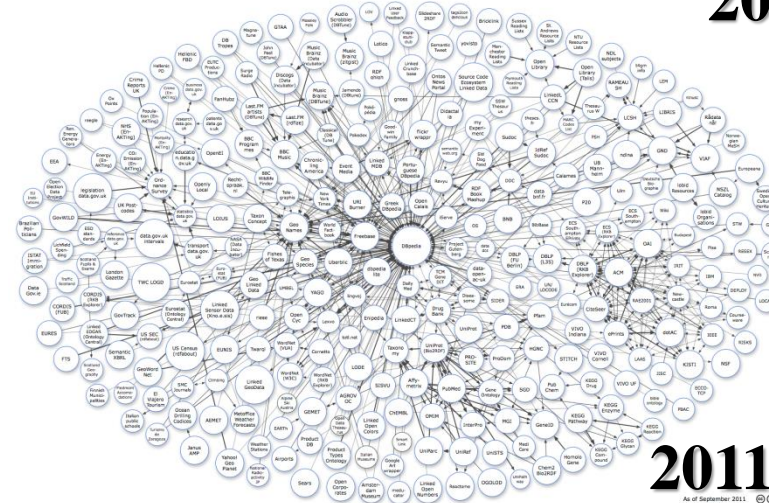
- Τα τελευταία 10 χρόνια → Τεράστια αύξηση στη δημοσίευση ανοιχτών δεδομένων στον Ιστό
 - Κυβερνητικά
 - Ακαδημαϊκά
 - Πολιτισμικά
 - Επιστήμες Ζωής
 - ...



2007



2009

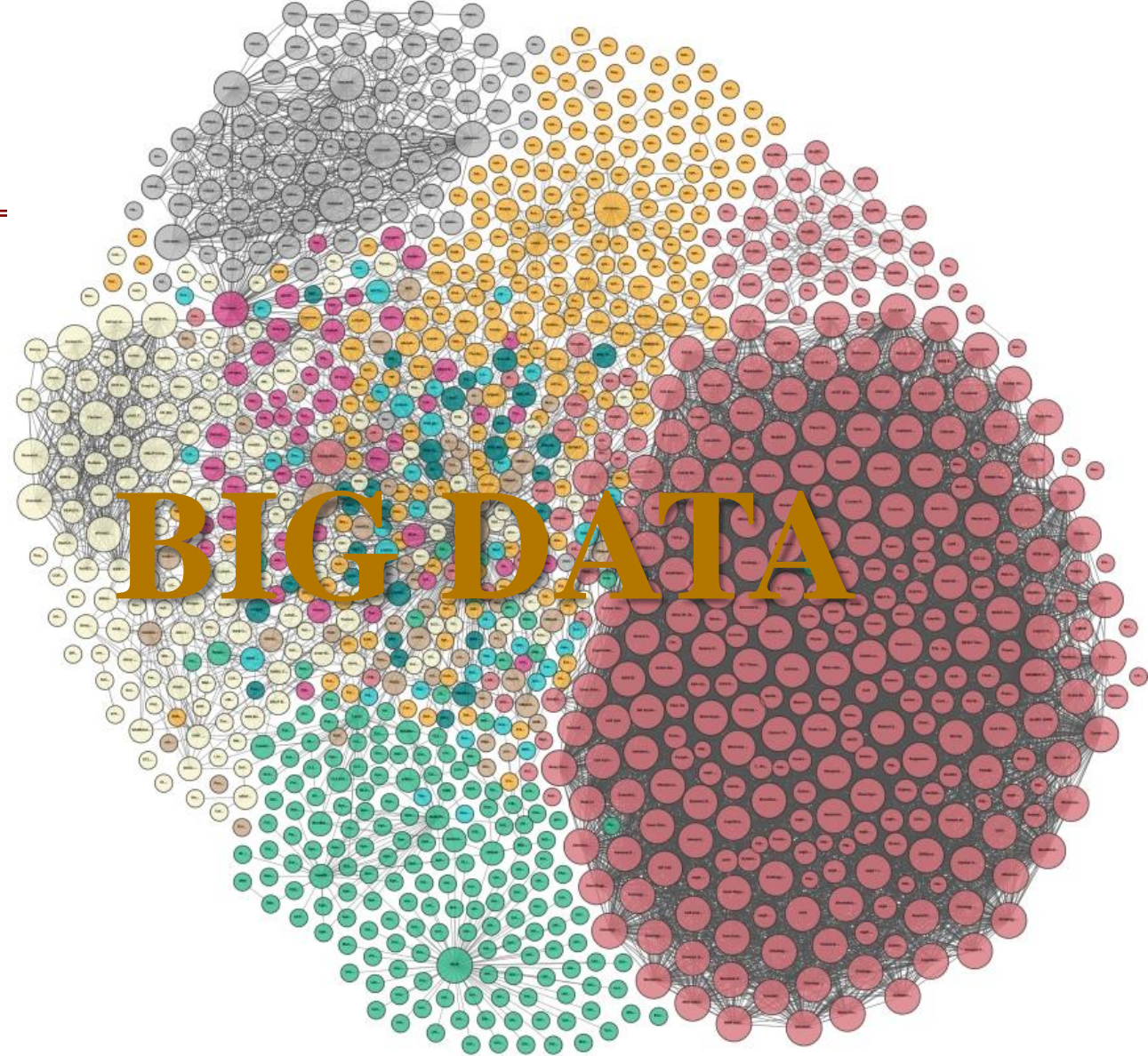


2011

2023

Legend

- Cross Domain
- Geography
- Government
- Life Sciences
- Linguistics
- Media
- Publications
- Social Networking
- User Generated
- Incoming Links
- Outgoing Links



Linking Open Data cloud diagram 2017, by Andrejs Abele, John P. McCrae, Paul Buitelaar, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>

- **Όγκος**
 - Π.χ. Europeana-LOD: ~2.4M entities
- **Πολλές Διασυνδέσεις**
 - ~117M triples
- **Δυναμικό περιβάλλον**
 - Updated >200 data providers

<https://old.datahub.io/dataset/europeana-lod-v1>

RDF MODEL

Τι είναι η RDF;



- Η Resource Description Framework (RDF) είναι **ένα μοντέλο δεδομένων** για την αναπαράσταση πληροφορίας (**δεδομένα & μεταδεδομένα**) για **πόρους** στο Web.
- RDF μπορεί επίσης να χρησιμοποιηθεί για την αναπαράσταση πληροφορίας για **οντότητες** (και όχι μόνο έγγραφα) , οι οποίες μπορούν να αναγνωριστούν στο Web, ακόμα και αν δεν μπορούν άμεσα να ανακτηθούν (π.χ. ένα πρόσωπο, μια έννοια, ένα βιβλίο, κτλ)
- RDF κυρίως στοχεύει στην αναπαράσταση πόρων στο διαδίκτυο που χρειάζονται να **προσπελαστούν από εφαρμογές** παρά να παρουσιαστούν σε ανθρώπους.

Πριν την RDF ...

- RDF βασίζεται σε ιδέες και τεχνολογίες αναπαράστασης γνώσης, τεχνητής νοημοσύνης και διαχείρισης δεδομένων:
 - Semantic networks
 - Frames
 - Conceptual graphs
 - Logic-based knowledge representation
 - Relational databases
 - XML

Τα βασικά της RDF

- RDF βασίζεται στην ιδέα να αναγνωρίζουμε πόρους στον Ιστό χρησιμοποιώντας μοναδικούς **Web identifiers** και να περιγράφουμε τους πόρους αυτούς με απλές ιδιότητες (**properties**) και τιμές ιδιοτήτων (**property values**).
- Οι identifiers είναι
 - **Uniform Resource Identifiers (URIs)**
 - **URI references (URIrefs)**.
- **Ορισμός:** Ένας πόρος είναι οτιδήποτε μπορεί να αναγνωριστεί από ένα URI.

Παράδειγμα ενός πόρου

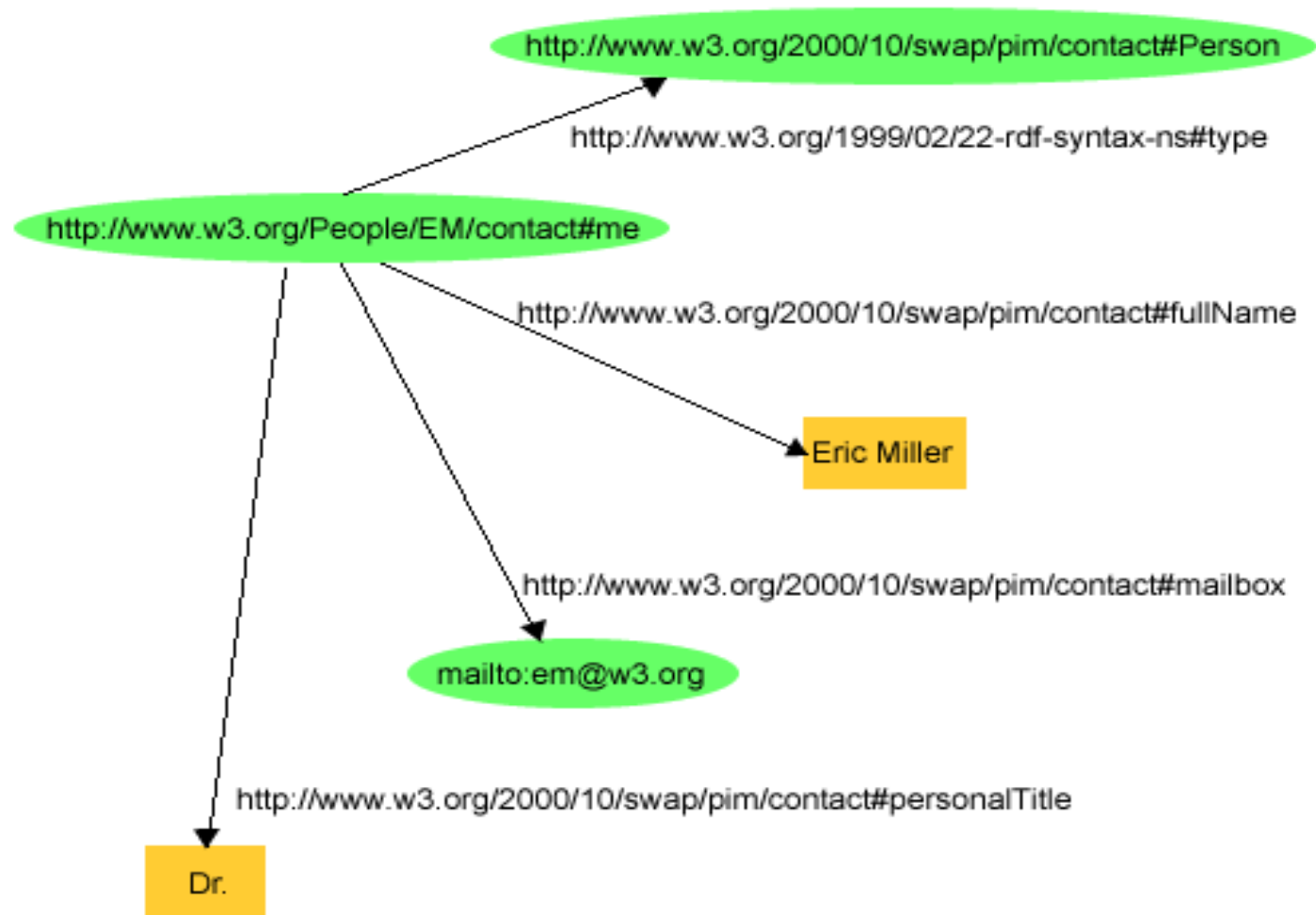
Ένα πρόσωπο μπορεί να αναγνωριστεί από το URI

`http://www.w3.org/People/EM/contact#me,`

του οποίου

- Το όνομα είναι Eric Miller,
- Το email είναι em@w3.org,
- Ο τίτλος του είναι “Dr.”

Παράδειγμα (cont'd)



Τα Βασικά της RDF (cont'd)

- Η RDF βασίζεται στην ιδέα ότι
 - κάθε πόρος έχει **ιδιότητες** που χαρακτηρίζονται από **τιμές** και ότι
 - κάθε πόρος μπορεί να περιγραφεί με τη μορφή **δηλώσεων** (statements) που προσδιορίζουν τις ιδιότητές του και τις τιμές του.
- Ορολογία
 - Το τμήμα που προσδιορίζει την οντότητα μια τέτοιας δήλωσης ονομάζεται υποκείμενο (**subject**).
 - Το τμήμα που προσδιορίζει την ιδιότητα, τη σχέση μιας δήλωσης ονομάζεται κατηγορημα (**predicate**).
 - Το τμήμα που προσδιορίζει την τιμή της ιδιότητας μιας δήλωσης ονομάζεται αντικείμενο (**object**).

Παράδειγμα

`http://www.example.org/index.html` has a creator whose value is "John Smith"

- **Subject** = URL `http://www.example.org/index.html`
- **Predicate** = "creator"
- **Object** = "John Smith"

Βασικές έννοιες της RDF

- Βασικές έννοιες της RDF
 - Αρχικά: resources, properties, values, statements, triples
 - URIs and URIrefs
 - RDF graphs
 - Literals
 - Shorthand notation: QNames
 - URIs as vocabularies
 - RDFS
 - Other data modeling concepts: structured values and blank nodes
- Serialization of RDF graphs: XML/RDF and Turtle

Uniform Resource Identifiers

Το Web παρέχει μια γενική μορφή **μοναδικού (unique)** και **σταθερού (persistent)** αναγνωριστικού, **Uniform Resource Identifier (URI)**, για να αναγνωρίζουμε (ονοματίζουμε) πόρους.

URIs VS URL

- Uniform Resource Locators χρησιμοποιούνται παραδοσιακά για να προσδιορίζουν μια δικτυακή τοποθεσία - διεύθυνση στον ιστό (π.χ. ένα αρχείο σε ένα απομακρυσμένο υπολογιστή). Π.χ.

<https://gunet2.cs.unipi.gr/modules/document/document.php?course=CDS110&openDir=/61603867SwGR> προσδιορίζει το φάκελο με τις διαλέξεις του μαθήματος

- URIs δεν προσδιορίζουν **απαραίτητα** αρχεία που βρίσκονται σε μια δικτυακή τοποθεσία. Π.χ. διαφορετικά **URI schemes (URI forms)** έχουν αναπτυχθεί και χρησιμοποιούνται για διάφορους λόγους. Π.χ.
 - `http:` (Hypertext Transfer Protocol, for Web pages)
 - `mailto:` (email addresses), e.g., `mailto:em@w3.org`
 - `ftp:` (File Transfer Protocol)
 - `urn:` (Uniform Resource Names, → persistent location-independent resource identifiers), π.χ. `urn:isbn:0-520-02356-0` (για ένα βιβλίο)
- Κανένας άνθρωπος ή οργανισμός δεν αποδίδει κεντρικά URIs ή ελέγχει πως θα χρησιμοποιηθούν. Π.χ. ενώ κάποια URI schemes, όπως το URL's `http:`, εξαρτώνται από κεντρικά συστήματα όπως το DNS, άλλα schemes, είναι τελείως **αποκεντριοποιημένα**.

URIs στην RDF

- Η RDF χρησιμοποιεί URI για να αναγνωρίζει και να ονοματίζει **subjects, predicates, objects** σε κάθε δήλωση.
- Τα RDF URI μπορούν να περιέχουν **Unicode** characters, επιτρέποντας URIs σε διαφορετικές γλώσσες → RDF χρησιμοποιεί **Internationalized Resource Identifiers (IRIs)** και IRIs.

Δείτε επίσης στο <https://www.w3.org/TR/cooluris/> για το πώς φτιάχνουμε URIs & IRIs

RDF Triples

- Κάθε δήλωση RDF μπορεί να γραφτεί με τη μορφή μιας τριπλέτας (**triple**)

Subject	Predicate	object .
http://www.example.org/index.html	Creator	"John Smith"

- **Παραδειγμα:**

```
<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/creator>  
<http://www.example.org/staffid/85740> .
```

```
<http://www.example.org/index.html> <http://www.example.org/terms/creation-date> "August  
16, 1999" .
```

```
<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/language> "en" .
```

- **Σημείωση:** Σε αυτή τη μορφή όλα τα URIs γράφονται ολόκληρα μέσα σε αγκύλες <>.

RDF Τριπλέτες → Γράφοι Γνώσης

- Γραφικά, η RDF μοντελοποιεί δηλώσεις με τη μορφή κόμβων και ακμών σε ένα **γράφο**.
- Μια τριπλέτα RDF αναπαρίσταται με τη μορφή γράφου ως:
 - Ένας κόμβος για το `subject`
 - Μια ακμή για το `predicate`, που ξεκινάει από το `subject` και καταλήγει στο `object`.
 - Ένας κόμβος για το `object`
- Ένα κόμβος προσδιορίζεται (έχει μια ετικέτα) από ένα URI ή μπορεί να είναι λεκτικό (επίσης μπορεί να είναι ένας κενός κόμβος).
- Μια ακμή μπορεί να είναι μόνο URI.

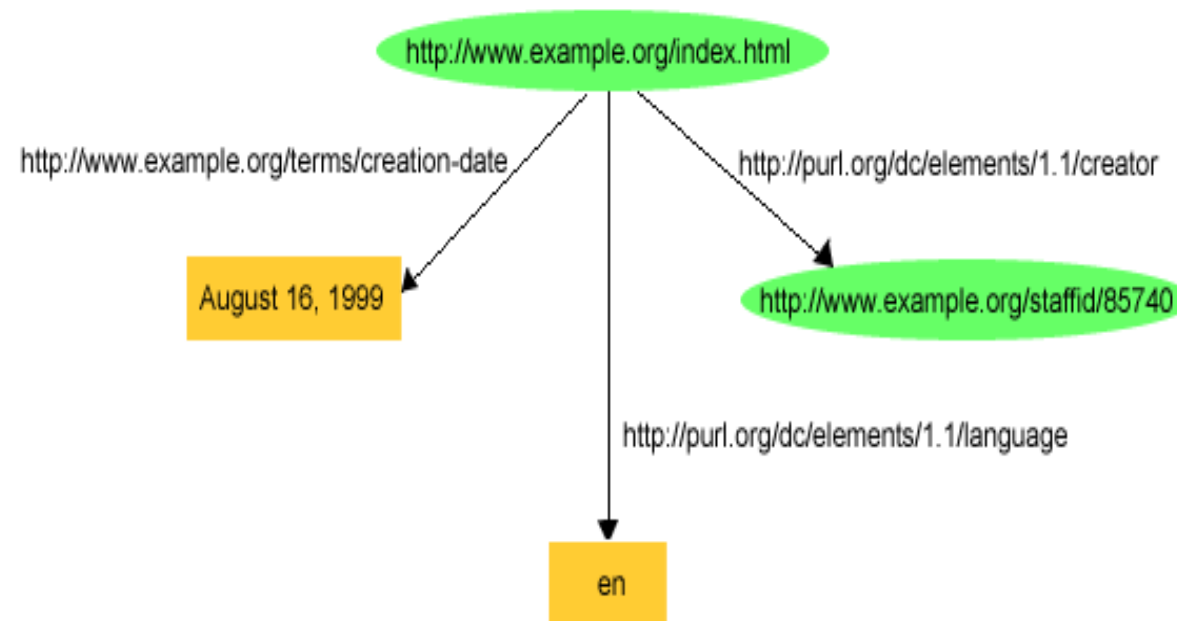
...ο γράφος του προηγούμενου παραδείγματος

3 RDF triples

`<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/creator> <http://www.example.org/staffid/85740> .`

`<http://www.example.org/index.html> <http://www.example.org/terms/creation-date> "August 16, 1999" .`

`<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/language> "en" .`



RDF και άλλα μοντέλα δεδομένων

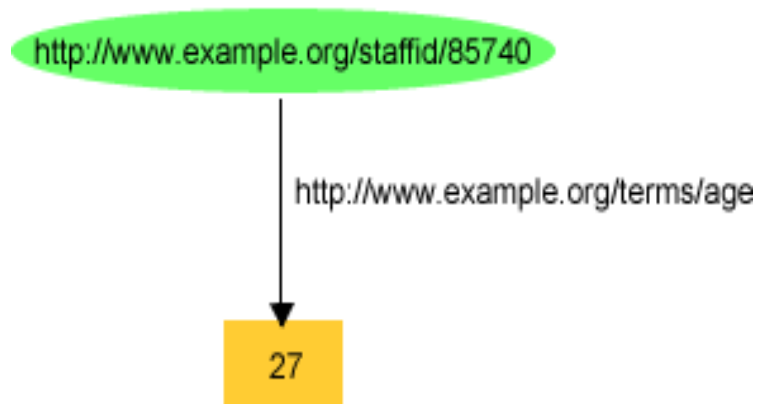
- Σε σχέση με το **σχεσιακό μοντέλο**, ένα RDF statement μπορεί να θεωρηθεί μια **tuple (εγγραφή) σε μια σχέση (πίνακα)**
 - Π.χ. ένας πίνακας «*Triples or Graph*» με 3 κολώνες *Subject*, *Predicate* και *Object*.

Triples or Graph

Subject	Predicate	Object
http://www.example.org/index.html	creation-date	August 16, 1999
http://www.example.org/index.html	language	English

Λεκτικά - Literals

- Υπάρχουν 2 είδη λεκτικών
 - Απλά (untyped).
 - Ακολουθούν κάποιο τύπο δεδομένων (typed).



- Το 27 τι είναι; Αριθμός ή αλφαριθμητικό;

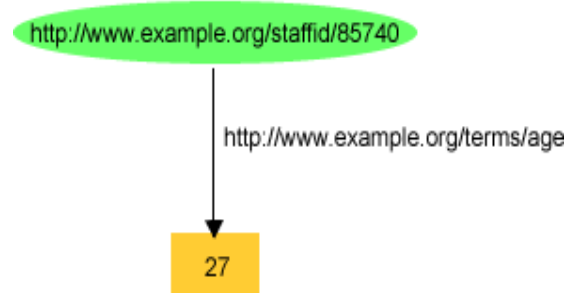
Απλά Λεκτικά

- Έχουν μια **αλφαριθμητική μορφή** (η τιμή τους) και προαιρετικά μια ετικέτα – σήμανση για τη **γλώσσα**
- Π.χ.:
 - "27",
 - "Hello world"@en
 - "Πειραιάς"@el

Λεκτικά βάσει τύπου δεδομένων

- Ένα **RDF typed literal** σχηματίζεται με το να αναθέτουμε ένα τύπο δεδομένων στο λεκτικό →

Συνοδεύουμε την τιμή (^^) με ένα URI που προσδιορίζει **το τύπο δεδομένων**



- Π.χ.:

```
<http://www.example.org/staffid/85740> <http://www.example.org/terms/age> "27"^^http://www.w3.org/2001/XMLSchema#integer
```

Triple Notation – Συντομεύσεις

- Η πλήρης αναπαράσταση RDF σε triple notation μπορεί να καταλήξει σε μακριές δηλώσεις λόγω των URIs.

`<http://www.example.org/staffid/85740> <http://www.example.org/terms/age> "27"^^http://www.w3.org/2001/XMLSchema#integer`



- **Συντόμευση:** Μπορούμε να χρησιμοποιήσουμε ένα **qualified name** (or **QName**) χωρίς αγκύλες <> για να αναφερθούμε σε ένα full URI reference.
- Ένα **QName** αποτελείται από ένα πρόθεμα (**prefix**) το οποίο έχει αντιστοιχηθεί σε ένα **namespace URI**, ακολουθούμενο από ένα **colon (:)**, και μετά το **τοπικό όνομα του πόρου**. Το πλήρες URIref σχηματίζεται από το QName με τον να αντικαταστήσουμε το πρόθεμα με το URI και να το ενώσουμε με το τοπικό όνομα του πόρου.
- Οι έννοιες **χώροι ονομάτων (namespace)** χρησιμοποιούνται στην RDF από την XML.

QNames (cont'd)

- **Παράδειγμα:**

- QName prefix `foo` συμβολίζει το namespace URI <http://example.org/somewhere/>,
- Τότε το QName `foo:bar` → <http://example.org/somewhere/bar>.

- **Γνωστά Namespaces:**

prefix **rdf:**, namespace URI: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

prefix **rdfs:**, namespace URI: <http://www.w3.org/2000/01/rdf-schema#>

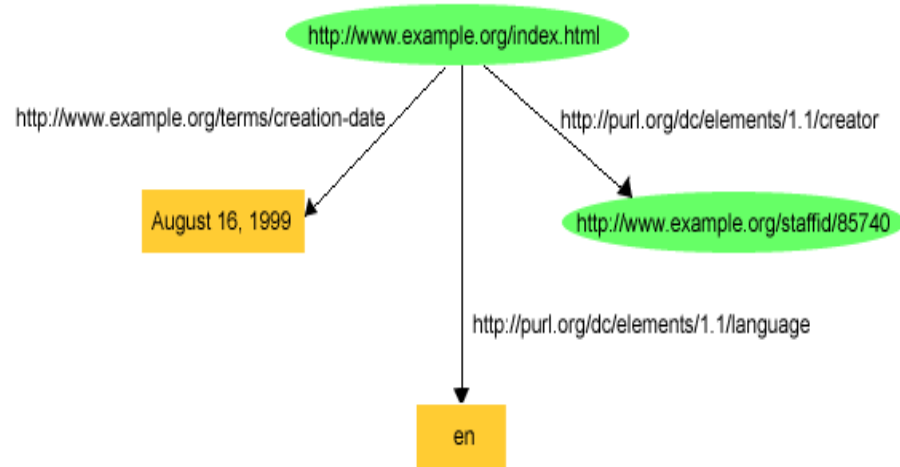
prefix **dc:**, namespace URI: <http://purl.org/dc/elements/1.1/>

prefix **owl:**, namespace URI: <http://www.w3.org/2002/07/owl#>

prefix **ex:**, namespace URI: <http://www.example.org/> (or <http://www.example.com/>)

prefix **xsd:**, namespace URI: <http://www.w3.org/2001/XMLSchema#>

Παράδειγμα



- Ο γράφος σε triple notation :

`ex:index.html dc:creator exstaff:85740 .`

`ex:index.html exterms:creation-date "August 16, 1999" .`

`ex:index.html dc:language "en" .`

URI as Vocabulary

- Αφού η RDF χρησιμοποιεί URIs αντί για λέξεις για να ονοματίζει αντικείμενα σε δηλώσεις, τότε μπορούμε να χρησιμοποιήσουμε τα URIs για να ορίσουμε ένα **λεξιλόγιο (δλδ μεταδεδομένα)** στην RDF.
- Τα URIs σε RDF vocabularies οργανώνονται ως ένα **σύνολο** από **QNames με ένα κοινό prefix**:
 - Ένα κοινό namespace επιλέγεται για όλους τους όρους ενός λεξιλογίου, συνήθως ένα URI το οποίο καθορίζει το λεξιλόγιο.
 - Τα URIs που περιέχονται στο λεξιλόγιο σχηματίζονται με την προσθήκη μεμονωμένων τοπικών ονομάτων στο τέλος του κοινού URI

URI as Vocabulary (cont'd)

- Η RDF χρησιμοποιεί την ίδια προσέγγιση για να ορίσει το λεξιλόγιο – τους όρους που έχουν ιδιαίτερη σημασία στην RDF:
 - Όλα τα URIs όρων στο λεξιλόγιο της RDF ξεκινάνε με:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>,
→prefix `rdf:` .
 - Η RDF Schema Language ορίζει ένα σύνολο όρων με URIs που αρχίζουν από
<http://www.w3.org/2000/01/rdf-schema#>,
→prefix `rdfs:` .
- Το prefix χρησιμοποιείται πάντα σε συνδυασμό με ένα σύνολο όρων. Για αυτό συνηθίζεται το QName prefix να χρησιμοποιείται σαν **το όνομα του λεξιλογίου**. Π.χ. συνήθως λέμε το "**`rdfs: vocabulary`**".

URIs as Vocabulary (cont'd)

- **Σύμβαση:** Οι οργανισμοί συνήθως χρησιμοποιούν ένα namespace URI σαν το prefix των όρων του λεξιλογίου που χρησιμοποιούν.
- **Παράδειγμα:** Το prefix `dc :` χρησιμοποιείται για το <http://purl.org/dc/elements/1.1> και αναφέρεται στο **Dublin Core vocabulary**.
 - Η πρόσβαση σε αυτόν τον χώρο ονομάτων από ένα πρόγραμμα περιήγησης στο Web θα ανακτήσει πρόσθετες πληροφορίες σχετικά με το λεξιλόγιο Dublin Core
 - **Υπενθύμιση:** αυτή είναι μόνο μια χρήσιμη σύμβαση. Το RDF δεν υποθέτει ότι ένα namespace URI προσδιορίζει έναν πόρο Web που μπορεί να ανακτηθεί

URIs as Vocabulary (cont'd)

- Η χρήση URIs σαν subjects, predicates, objects στις RDF δηλώσεις υποστηρίζει την **ανάπτυξη και επαναχρησιμοποίηση** κοινών λεξιλογίων στο Web.
- Οι άνθρωποι μπορούν να **ανακαλύψουν και να αρχίσουν να χρησιμοποιούν** λεξιλόγια τα οποία ήδη χρησιμοποιούνται από άλλους για να περιγράψουν πόρους. Με αυτόν τον τρόπο επιτυγχάνεται μια **κοινή αντίληψη εννοιών** στον Ιστό.

- Ας πάρουμε την τριπλέτα

`ex:index.html dc:creator exstaff:85740 .`

Το predicate `dc:creator`, δηλ το URI στο οποίο αντιστοιχεί, είναι μια **μη διφορούμενη** αναφορά στην ιδιότητα "creator" που περιγράφεται στο Dublin Core metadata

Ο συγγραφέας της τριπλέτας δηλώνει ότι: Η σχέση μεταξύ της σελίδας `index.html` και του δημιουργού της είναι μια σαφής έννοια που ταυτοποιείται και περιγράφεται από:
<http://purl.org/dc/elements/1.1/creator>.

Κάποιος που βρίσκει τι σημαίνει `dc:creator` (π.χ. βλέποντας τον ορισμό στο αντίστοιχο URL) θα γνωρίζει τι σημαίνει αυτή η σχέση. Επιπρόσθετα, τα **προγράμματα** που κατασκευάζονται μπορούν να αντιμετωπίζουν την σχέση αυτή σύμφωνα με την έννοιά της όταν επεξεργάζονται και ανακτούν τριπλέτες με την σχέση `dc:creator`.

Άλλο Παράδειγμα



- Friend of a Friend (FOAF) vocabulary @ <http://xmlns.com/foaf/spec/>.
- Το FOAF vocabulary έχει δημιουργήσει ένα δίκτυο από machine-readable σελίδες (σε RDF) που περιγράφουν **πρόσωπα**, τις σχέσεις μεταξύ τους και γενικότερα τα πράγματα που δημιουργούν και κάνουν.

ex:gpapas foaf:name "George Papastefanatos"

<http://web.imsi.athenarc.gr/~gpapas/index.rdf>

```
<rdf:RDF>
  <foaf:PersonalProfileDocument rdf:about="">
    <foaf:maker rdf:resource="#me"/>
    <foaf:primaryTopic rdf:resource="#me"/>
  </foaf:PersonalProfileDocument>
  <foaf:Person rdf:about="http://web.imis.athena-innovation.gr/~gpapas/index.rdf">
    <foaf:name>George Papastefanatos</foaf:name>
    <foaf:title>Dr</foaf:title>
    <foaf:givenname>George</foaf:givenname>
    <foaf:family_name>Papastefanatos</foaf:family_name>
    <foaf:nick>gpapas</foaf:nick>
    <foaf:mbox_sha1sum>164e68227742040496903441f7a20287acc7a0c1</foaf:mbox_sha1sum>
    <foaf:homepage rdf:resource="http://web.imis.athena-innovation.gr/~gpapas/">
    <foaf:phone rdf:resource="tel:++302106875402"/>
  </foaf:Person>
  <foaf:knows>
    <foaf:Person>
      <foaf:name>Yannis</foaf:name>
      <foaf:mbox_sha1sum>c578343d83ab26e1b5779203f775579a95522eb1</foaf:mbox_sha1sum>
      <rdfs:seeAlso rdf:resource="http://web.imis.athena-innovation.gr/~yannis/index.rdf"/>
    </foaf:Person>
  </foaf:knows>
  <foaf:knows>
    <foaf:Person>
      <foaf:name>Theodore</foaf:name>
      <foaf:mbox_sha1sum>41adce671a5cf30801ad1ec5c724ae5cee54deae</foaf:mbox_sha1sum>
      <rdfs:seeAlso rdf:resource="http://web.imis.athena-innovation.gr/~dalamag/people/people.rdf#1"/>
    </foaf:Person>
  </foaf:knows>
</foaf:Person>
</rdf:RDF>
```

URI as Vocabulary (cont'd)

- Η RDF δίνει έννοια στους όρους που ορίζονται στα σχετικά RDF vocabularies `rdf:` and `rdfs:`.
 - `rdf:type`
 - `rdfs:seeAlso`

<https://www.w3.org/TR/rdf-syntax-grammar/>
- Άλλοι οργανισμοί έχουν ορίσει άλλους σημαντικούς όρους λεξιλογίων, π.χ.,
 - `dc:`
 - `owl:`
 - `foaf:`
 - `Geonames:`
 - `SKOS:`
 - ...

RDF Schema

- Το RDFS (RDF Schema) ή RDF Vocabulary Schema αποτελεί σημασιολογική επέκταση του RDF
- Έχει σχεδιαστεί για να περιγράφει πόρους και/ή τις μεταξύ τους σχέσεις, χρησιμοποιώντας ένα σύνολο δεσμευμένων λέξεων που ονομάζεται RDFS λεξιλόγιο
- Συγκεκριμένα παρέχει δομές για τις περιγραφές:
 - τύπων αντικειμένων (κλάσεων-classes),
 - τύπους ιεραρχιών (υπο-κλάσεις),
 - ιδιότητες που αντιπροσωπεύουν χαρακτηριστικά αντικειμένων (properties)
 - ιεραρχίες ιδιοκτησίας (subproperties).

Τεχνικά Χαρακτηριστικά του RDFS

- Η RDFS παρέχει την έννοια της κλάσης (class) για τον ορισμό των κατηγοριών που θα χρησιμοποιηθούν για την ταξινόμηση των πόρων (resources)
- Η σχέση μεταξύ ενός στιγμιότυπου και της κλάσης του ορίζεται μέσω της ιδιότητας *τύπος (type)*.
- Μπορούμε να δημιουργήσουμε ιεραρχίες κλάσεων και υποκλάσεων, όπως και ιεραρχίες ιδιοτήτων και υπο-ιδιοτήτων.
- Περιορισμοί στους τύπους των αντικειμένων ή/και των υποκειμένων συγκεκριμένων τριάδων μπορούν να οριστούν μέσω του πεδίου ορισμού τους (domain) και του συνόλου τιμών τους (range).

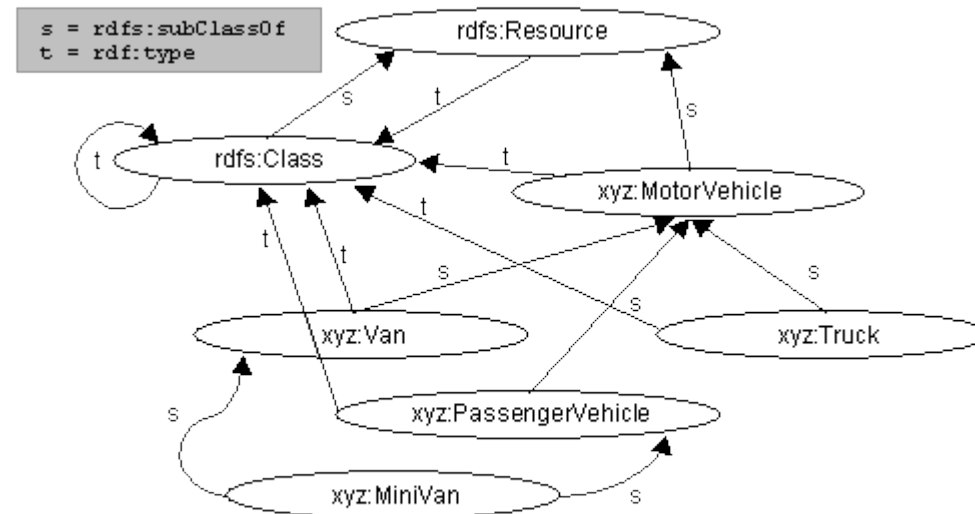
Για παράδειγμα, αντικείμενα και υποκείμενα των τριάδων, στα οποία συμμετέχει το URI <http://www.example.org/friendOf> ως ιδιότητα, θα πρέπει να είναι στιγμιότυπα (instances) της κλάσης <http://www.example.org/Person>.

Δομές Μοντελοποίησης του RDFS

Δομή	Σύνταξη τριάδας	Περιγραφή
Class (μια κλάση)	C rdf:type rdfs:Class .	C (ένας πόρος) είναι μια κλάση RDF
Property (μια κλάση)	P rdf:type rdf:Property .	P (ένας πόρος) είναι μια ιδιότητα RDF
type (μια ιδιότητα)	I rdf:type C .	I (ένας πόρος) είναι ένα στιγμιότυπο της C (μια κλάση)
subClassOf (μια ιδιότητα)	C1 rdfs:subClassOf C2 .	C1 (μια κλάση) είναι υποκλάση της C2 (μια κλάση)
subPropertyOf (μια ιδιότητα)	P1 rdfs:subPropertyOf P2 .	P1 (μια ιδιότητα) είναι υπο-ιδιότητα της P2 (μια ιδιότητα)
domain (μια ιδιότητα)	P rdfs:domain C .	Το πεδίο ορισμού P (μιας ιδιότητας) είναι C (μια κλάση)
range (μια ιδιότητα)	P rdfs:range C .	Το σύνολο τιμών P (μιας ιδιότητας) είναι C (μια κλάση)

Παράδειγμα RDFS

- Ορίζουμε την κλάση 'MotorVehicle', και τις υποκλάσεις της 'PassengerVehicle', 'Truck' και 'Van'.
- Ορίζουμε την κλάση 'Minivan' ως υπο-κλάση των 'Van' και 'PassengerVehicle'.



Παράδειγμα RDFS

- Βάσει της σύνταξης RDF/XML και αξιοποιώντας το RDFS
 - Ο ορισμός της κλάσης 'MotorVehicle':

```
<rdf:Description ID="MotorVehicle">  
  <rdf:type resource="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Class"/>  
  <rdfs:subClassOf rdf:resource="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Resource"/>  
</rdf:Description>
```

Παράδειγμα RDFS

- Βάσει της σύνταξης RDF/XML και αξιοποιώντας το RDFS
 - Ο ορισμός της υπο-κλάσης 'PassengerVehicle':

```
<rdf:Description ID="PassengerVehicle">  
  <rdf:type resource="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Class"/>  
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>  
</rdf:Description>
```

Παράδειγμα RDFS

- Βάσει της σύνταξης RDF/XML και αξιοποιώντας το RDFS
 - Ο ορισμός της υπο-κλάσης 'Mini-Van':

```
<rdf:Description ID="MiniVan">  
  <rdf:type resource="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#Class"/>  
  <rdfs:subClassOf rdf:resource="#Van"/>  
  <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>  
</rdf:Description>
```

Machine Readable Formats για RDF

Η RDF μπορεί να γραφτεί με διάφορες μορφές:

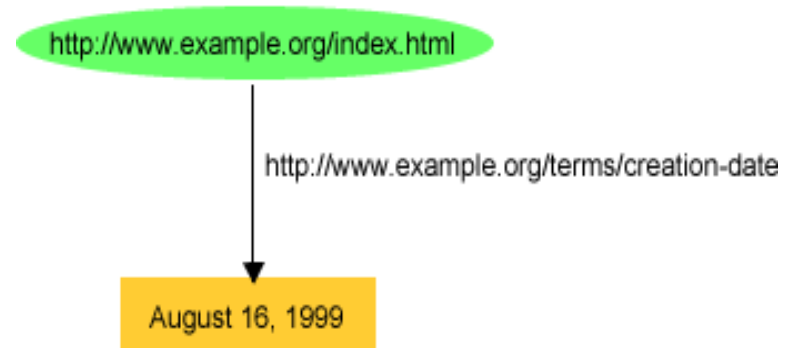
- XML/RDF
- Turtle (Terse RDF Triple Language)
- N3
- ...

XML Syntax for RDF: RDF/XML

- Το μοντέλο της RDF είναι ένας γράφος. Η XML ακολουθεί ιεραρχίες (δέντρο).
- Η RDF παρέχει μια XML syntax για συγγραφή και ανταλλαγή RDF γράφων → **RDF/XML**.
- RDF/XML είναι η κανονική σύνταξη για να γράφουμε RDF.

Παράδειγμα

<http://www.example.org/index.html> έχει creation-date με τιμή August 16, 1999.



Παράδειγμα σε XML/RDF

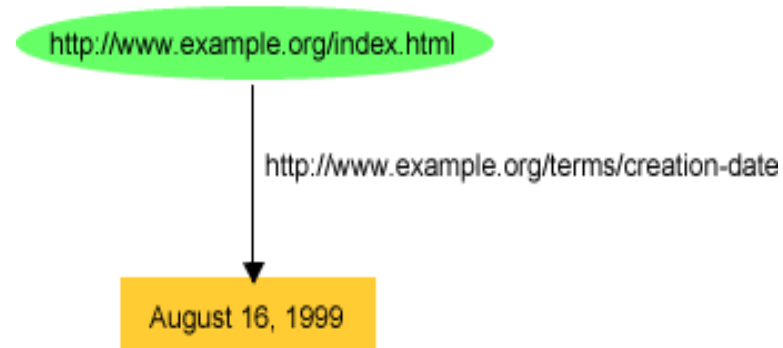
```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:exterms="http://www.example.org/terms/">
  <rdf:Description rdf:about="http://www.example.org/index.html">
    <exterms:creation-date>
      August 16, 1999
    </exterms:creation-date>
  </rdf:Description>
</rdf:RDF>
```

Turtle

- **Η Turtle** είναι μια εναλλακτική μορφή σύνταξης RDF.
- **Η Turtle** βασίζεται στην **triple notation**.

Παράδειγμα Ι

<http://www.example.org/index.html> έχει creation-date με τιμή August 16, 1999.



Παράδειγμα I σε Turtle

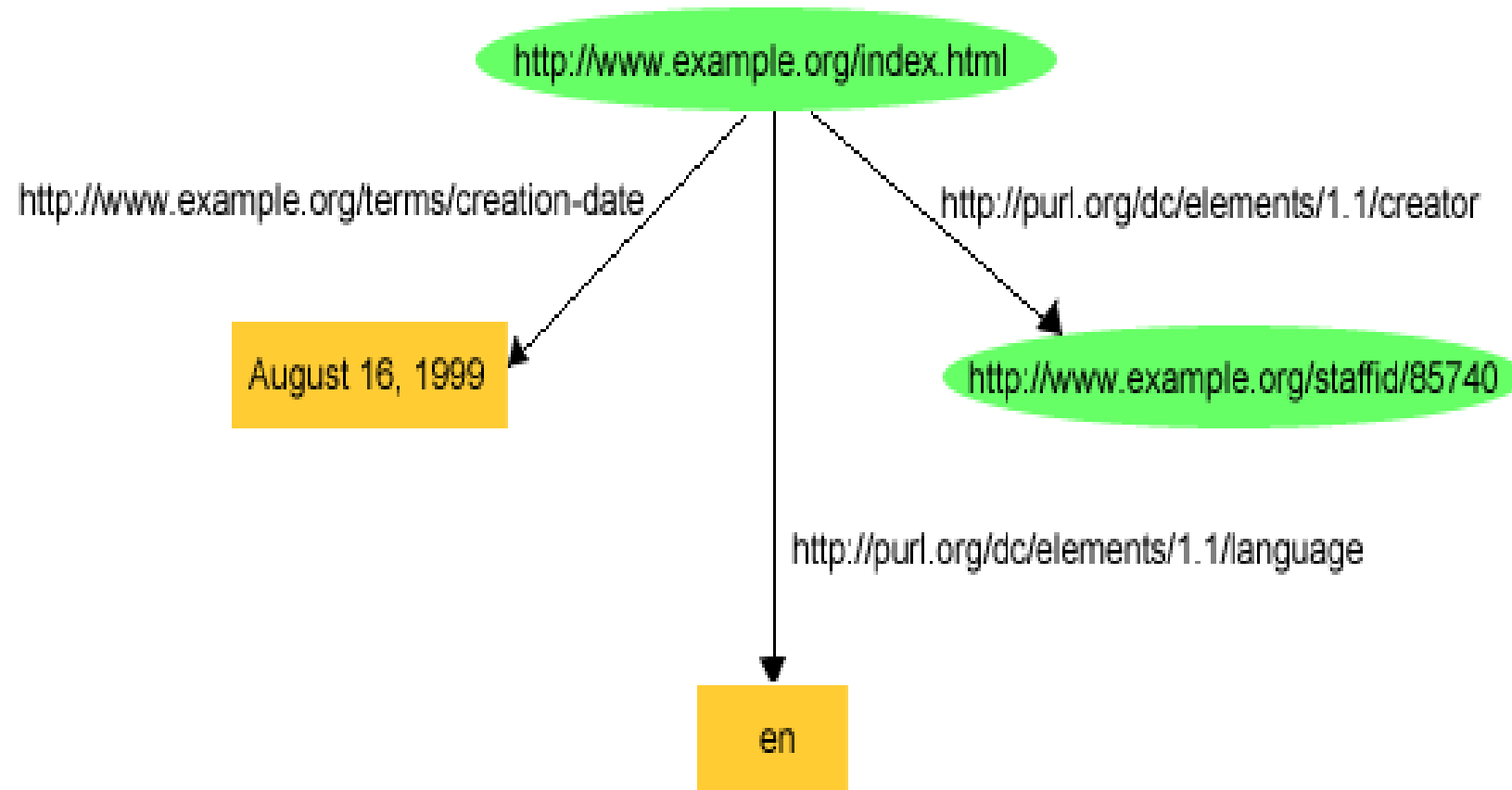
```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
```

```
@prefix exterms: <http://www.example.org/terms/>.
```

```
<http://www.example.org/index.html> exterms:creation-date "August 16, 1999" .
```

- **Σημείωση:** Το `@prefix` ορίζει τα namespaces `rdf:` and `exterms:`

Παράδειγμα II



Παράδειγμα II με Triples

`<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/creator> <http://www.example.org/staffid/85740> .`

`<http://www.example.org/index.html> <http://www.example.org/terms/creation-date> "August 16, 1999" .`

`<http://www.example.org/index.html> <http://purl.org/dc/elements/1.1/language> "en" .`

Παράδειγμα II με Turtle

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
```

```
@prefix dc: <http://purl.org/dc/elements/1.1/#>.
```

```
@prefix exterm: <http://www.example.org/terms/>.
```

```
<http://www.example.org/index.html>
```

```
  exterm:creation-date "August 16, 1999";
```

```
  dc:language "en";
```

```
  dc:creator <http://www.example.org/staffid/85740>.
```

- **Notation:** Turtle χρησιμοποιεί το semi-colon (;) για να μην επαναλαμβάνει το **subject** και να χωρίζει **predicate-object ζευγών** του ίδιου **subject**. Μια λίστα τέτοιων ζευγών τερματίζεται με τελεία (.)

Σύνοψη RDF

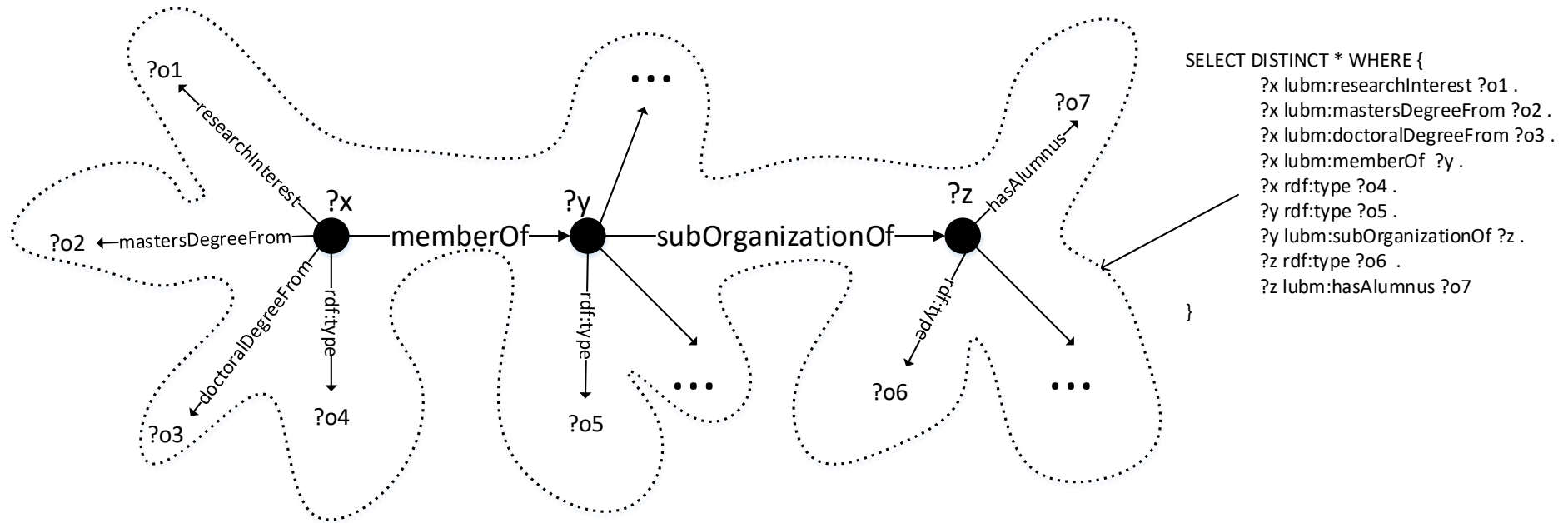
- Στην RDF, **δηλώσεις** για ένα τομέα κωδικοποιούνται με τη μορφή **τριπλετών (triples)**.
- Οι **τριπλέτες** έχουν τη μορφή
subject predicate object .
- Το **subject** πρέπει να είναι ένα **URI** ή ένα **blank node**.
- Το **predicate** πρέπει να είναι ένα **URI**.
- Το **object** of a triple πρέπει να είναι ένα **URI**, ένα **literal** ή ένα **blank node**.
- **Ορισμός:** Ένας **RDF graph** είναι ένα σύνολο από RDF triples.

SPARQL

SPARQL Basic Queries

- **SPARQL** βρίσκει **graph patterns**
- Το πιο απλό **graph pattern** είναι ένα **triple pattern**
- Triple Pattern = Σαν ένα RDF triple, αλλά ...
 - μπορούμε να βάλουμε μια μεταβλητή αντί για
 - URI ή λεκτικό στα subject, predicate, or object positions
- Η SPARQL συνδυάζει triple patterns για να προσδιορίσει ένα graph pattern
 - Exact match σε ένα υποσύνολο του γράφου είναι η απάντηση σε ένα ερώτημα

Motivation



Παράδειγμα

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
SELECT ?c
```

```
WHERE
```

```
{
```

```
    ?c rdf:type rdfs:Class .
```

```
}
```

Βρίσκει όλα τα triples για τα οποία:

- Το predicate είναι rdf:type
- Το object είναι rdfs:Class

Δλδ. βρίσκει και επιστρέφει όλες τις κλάσεις μέσα στον RDF graph.

Παράδειγμα(2)

- Βρες όλα τα στιγμιότυπα μιας συγκεκριμένης κλάσης (π.χ. uni:course) :
(το rdf prefix έχει παραληφθεί για συντομία)

```
PREFIX uni: <http://www.mydomain.org/uni-ns#>
```

```
SELECT ?i
```

```
WHERE
```

```
{
```

```
    ?i rdf:type uni:course .
```

```
}
```

Select-from-where στην SPARQL

- Όπως στην SQL, SPARQL queries έχουν μια **SELECT-FROM-WHERE** δομή:
 - **SELECT** προσδιορίζει τον **αριθμό και τη σειρά των μεταβλητών** που θα επιστραφούν
 - **FROM** προσδιορίζει την **πηγή (=γράφο)** τον οποίο ρωτάμε (optional)
 - **WHERE** εφαρμόζει περιορισμούς και φίλτρα στα δεδομένα που θα επιστραφούν με τη **μορφή graph pattern templates και boolean** συνθήκες (=, <, >, IN, κτλ)
- Βρες όλους τους τηλ. αριθμούς του προσωπικού του πανεπιστημίου:

```
SELECT ?x ?y
WHERE
{ ?x uni:phone ?y . }
```

- ?x και ?y είναι μεταβλητές
- ?x uni:phone ?y αναπαριστά μια triple pattern

Implicit Join in SPARQL

- Βρές όλους τους διδάσκοντες και τα τηλ. τους:

```
SELECT ?x ?y
WHERE
{ ?x rdf:type uni:Lecturer ;
  uni:phone ?y . }
```

- Implicit join: Περιορίζουμε το 2^ο pattern μόνο σε αυτές τις triples, των οποίων το subject είναι η μεταβλητή ?x
 - Semicolon(;) σημαίνει ότι η επόμενη triple έχει το ίδιο subject με την προηγούμενη (για λόγους συντομίας)

Implicit join (2)

- Το προηγούμενο ερώτημα μπορεί να γραφτεί και ως.

```
SELECT ?x ?y
```

```
WHERE
```

```
{
```

```
    ?x rdf:type uni:Lecturer .
```

```
    ?x uni:phone ?y .
```

```
}
```

Φίλτρα

- Βρες το όνομα όλων των μαθημάτων που διδάσκονται από τον διδάσκοντα με ID 949352 και το όνομα τους περιέχουν τη λέξη BigData

```
SELECT ?name
```

```
WHERE
```

```
{
```

```
    ?x rdf:type uni:Course ;
```

```
        uni:isTaughtBy :949352 ;
```

```
        uni:name ?name .
```

```
    FILTER regex(?name, "BigData")
```

```
}
```

Προαιρετικά (Optional) Patterns

```
<uni:lecturer rdf:about="949352">
```

```
  <uni:name>George Papastefanatos</uni:name>
```

```
</uni:lecturer>
```

```
<uni:lecturer rdf:about="94318">
```

```
  <uni:name>David Billington</uni:name>
```

```
  <uni:email>david@work.example.org</uni:email>
```

```
</uni:lecturer>
```

- Για τον ένα διδάσκοντα περιέχει μόνο το όνομα
- Για τον άλλο περιέχει και το email

Προαιρετικά (Optional) Patterns

- Βρες όλους τους διδάσκοντες και τα email τους.

```
SELECT ?name ?email
```

```
WHERE
```

```
{  
    ?x rdf:type uni:Lecturer ;  
        uni:name ?name ;  
        uni:email ?email .  
}
```

Προαιρετικά (Optional) Patterns

- Το αποτέλεσμα του προηγούμενου query είναι:

?name	?email
David Billington	david@work.example.org

- Ένας διδάσκων (George Papastefanatos) δεν επιστρέφεται σαν αποτέλεσμα αν δεν έχει e-mail

Προαιρετικά (Optional) Patterns

Χρήση του Keyword OPTIONAL:

```
SELECT ?name ?email
```

```
WHERE
```

```
{  
    ?x rdf:type uni:Lecturer ;  
        uni:name ?name .  
    OPTIONAL { x? uni:email ?email }  
}
```

Προαιρετικά (Optional) Patterns

- Δηλαδή “βρες τα ονόματα των διδασκόντων και αν υπάρχει το email τους.”

?name	?email
George Papastefanatos	
David Billington	david@work.example.org

SPARQL 1.1

- New features in **Query** language:
 - **Aggregate functions**: COUNT, SUM, MIN, MAX, AVG, GROUP_CONCAT, and SAMPLE.
 - **Subqueries**: Nest the results of a query within another query
 - **Negation**: Check the absence of triples in a graph
 - **Expressions in SELECT**: Introduce new variables in the SELECT clause
 - **Property paths**: Search graphs through structures that involve arbitrary-length paths
 - **Assignment**: BIND keyword and expressions in SELECT and GROUP_BY
 - **Short form for CONSTRUCT**
 - **Expanded functions and operators**: EXISTS, NOT EXISTS, SUBSTR, etc.
- <https://www.w3.org/TR/sparql11-query/>

Main References

- W3C SPARQL Working Group

<http://www.w3.org/2001/sw/DataAccess/>

- SPARQL validator:

<http://www.sparql.org/query-validator.html>

- SPARQL implementations:

<http://www.w3.org/topic/SparqlImplementations>

- SPARQL Endpoints

<http://www.w3.org/topic/SparqlEndpoints>

- SPARQL in Dbpedia

<http://dbpedia.org/sparql>

Πως δημοσιεύουμε Linked Data on the Web

PUBLISHING LINKED DATA

Scenario

- Online whisky shop: *Wiskii.com*
- New business venture, founded by *Jeff*
- For the whisky connoisseur
- Detailed background information from experts
- Contributions from customers
- Custom web app, relational backend
- Simultaneous publication in HTML and RDF

5 Βήματα για δημοσίευση Linked Data

1. Κατανόηση των βασικών αρχών
2. Κατανόηση των δεδομένων – πόροι και σχέσεις μεταξύ τους
3. Επιλογή URIs για τους πόρους
4. Δημιουργία της υποδομής για την αποθήκευση, πρόσβαση και εφαρμογή επερωτήσεων
5. Σύνδεση με άλλα σύνολα δεδομένων

1. Κατανόηση των βασικών αρχών

PUBLISHING LINKED DATA

5-star Open Data plan

- ★ make your stuff available on the Web (whatever format) under an open license¹
- ★★ make it available as structured data (e.g., Excel instead of image scan of a table)²
- ★★★ make it available in a non-proprietary open format (e.g., CSV instead of Excel)³
- ★★★★ use URIs to denote things, so that people can point at your stuff⁴
- ★★★★★ link your data to other data to provide context⁵

2. Κατανόηση των δεδομένων – πόροι και σχέσεις μεταξύ τους

PUBLISHING LINKED DATA

2. Κατανόηση των δεδομένων

- Ποιες είναι οι βασικές έννοιες στα δεδομένα;
 - People?
 - Places?
 - Books?
 - Films?
 - Musicians?
 - Concepts?
 - Photos?
 - Comments?
 - Reviews?
 - ...

2. Κατανόηση των δεδομένων

- Things in the *Wiskii.com* database
 - Distilleries
 - Regions and Locations
 - Founders
 - Owners
 - Brands
 - Products
 - Photos
 - Reviews
 - Comments
 - Prices/Offers



2. Κατανόηση των δεδομένων

- Ποια λεξιλόγια μπορούν να τα περιγράψουν?

☞ Αρχές

- Επαναχρησιμοποιήστε - μην ανακαλύπτετε τον τροχό.
- Συνδυάστε όπου χρειάζεται διαφορετικά λεξιλόγια

☞ Π.χ.

- GeoNames
- FOAF
- SKOS
- DC
- CIDOC CRM
- ...

3. Επιλογή URIs για τους πόρους

PUBLISHING LINKED DATA

3. Επιλογή URIs για τους πόρους – Αρχές

- Χρησιμοποιήστε HTTP URIs.
 - Επιτρέπουν στους χρήστες να "αναζητήσουν" (dereference) ένα URI για να αποκτήσουν πρόσβαση σε μια αναπαράσταση του πόρου που προσδιορίστηκε από αυτό το URI
- Μην χρησιμοποιείται άλλων οργανισμών τα namespaces
 - <http://www.imdb.com/title/tt0441773/>
 - <http://www.imdb.com/title/tt0441773/thing>
 - <http://myfilms.com/tt0441773>
 - <http://myfilms.com/tt0441773/html>
- Μην βάζετε λεπτομέρειες υλοποίησης μέσα στο URI
 - <http://dbpedia.org/resource/Berlin>
 - ~~http://www4.wiwiss.fu-berlin.de:2020/demos/dbpedia/cgi-bin/resources.php?id=Berlin~~
- Ένα URI δεν πρέπει να περιέχει έννοιες και αντικείμενα που αλλάζουν με το χρόνο (π.χ. Session token)
→ Persistent URIs
- Hash or Slash
 - <http://mydomain.com/foaf.rdf#me>
 - <http://mydomain.com/id/me>

3. Επιλογή URIs: Συνήθεις πρακτικές

- http://dbpedia.org/resource/New_York_City ← Thing
- http://dbpedia.org/data/New_York_City ← RDF data
- http://dbpedia.org/page/New_York_City ← HTML page

- <http://revyu.com/people/tom> ← Thing
- <http://revyu.com/people/tom/about/rdf> ← RDF data
- <http://revyu.com/people/tom/about/html> ← HTML page

- <http://kmi.open.ac.uk/people/tom/> ← Thing
- <http://kmi.open.ac.uk/people/tom/rdf> ← RDF data
- <http://kmi.open.ac.uk/people/tom/html> ← HTML page

- <http://mydomain.com/thing> ← Thing
- <http://mydomain.com/thing.rdf> ← RDF data
- <http://mydomain.com/thing.html> ← HTML page

3. Επιλογή URIs: Wiskii.com

- <http://wiskii.com/regions/speyside>
- <http://wiskii.com/distilleries/talisker>
- <http://wiskii.com/brands/talisker>
- <http://wiskii.com/products/talisker-20-yo>
- <http://wiskii.com/products/glenmorangie-lasanta>
- <http://wiskii.com/people/william-matheson>
- <http://wiskii.com/photos/58>
- <http://wiskii.com/reviews/271>

3. Επιλογή URIs : Wiskii.com

- <http://wiskii.com/distilleries/talisker>
- <http://wiskii.com/distilleries/talisker/rdf>
- <http://wiskii.com/distilleries/talisker/html>

- <http://wiskii.com/brands/talisker>
- <http://wiskii.com/brands/talisker/rdf>
- <http://wiskii.com/brands/talisker/html>

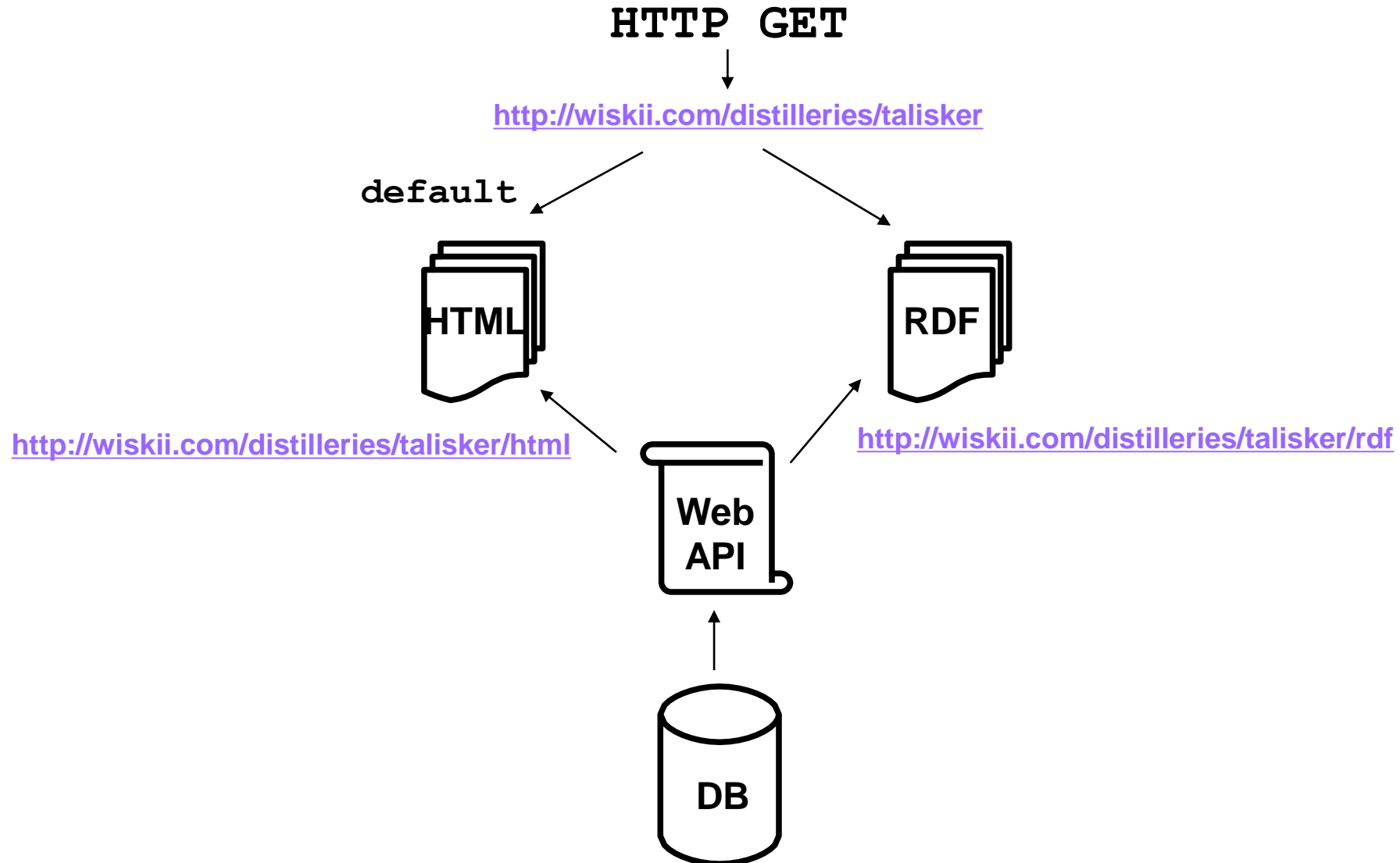
- <http://wiskii.com/people/william-matheson>
- <http://wiskii.com/people/william-matheson/rdf>
- <http://wiskii.com/people/william-matheson/html>

- <http://wiskii.com/photos/58>

4. Δημιουργία της υποδομής

PUBLISHING LINKED DATA

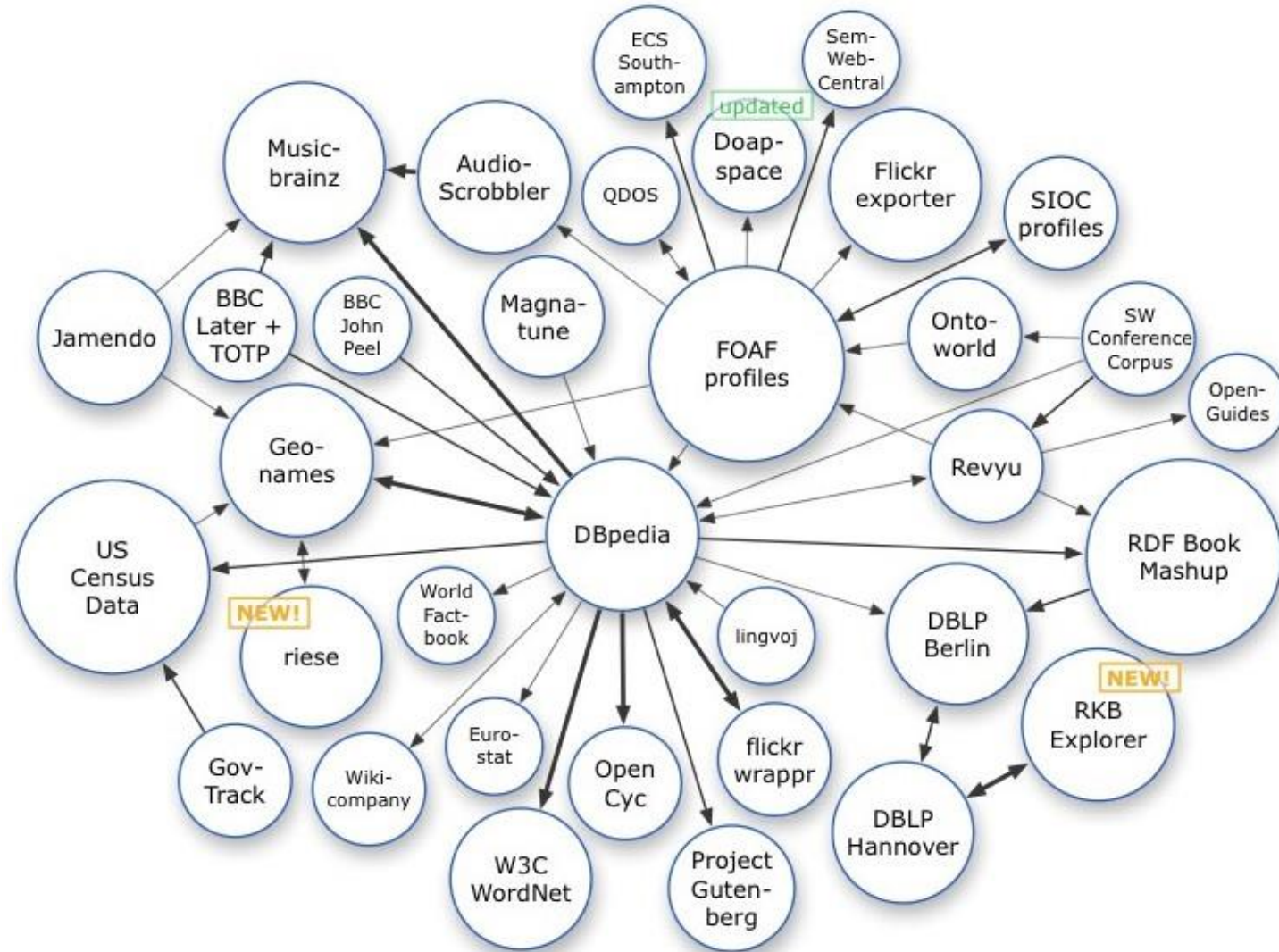
4. Δημιουργία της υποδομής



5. Συνδεθείτε με άλλα σύνολα δεδομένων

PUBLISHING LINKED DATA

Other Available Data Sets



5. Συνδεθείτε με άλλα σύνολα δεδομένων

- Χρήση γνωστών Predicates για interlinking

🌀 owl:sameAs

🌀 foaf:homepage

🌀 foaf:topic

🌀 foaf:based_near

🌀 foaf:maker/foaf:made

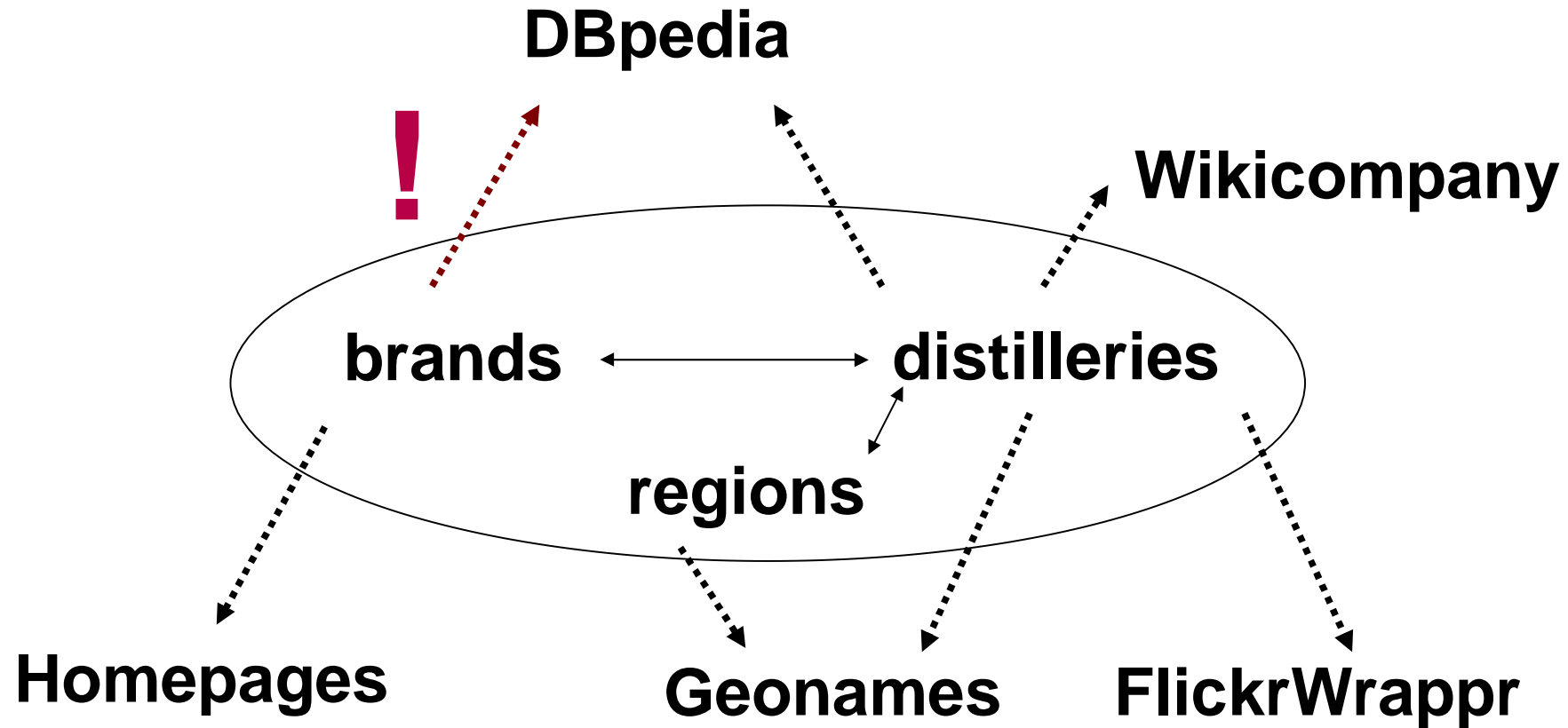
🌀 foaf:depiction

🌀 foaf:page

🌀 foaf:primaryTopic

🌀 rdfs:seeAlso

5. Συνδεθείτε με άλλα σύνολα δεδομένων



5. Συνδεθείτε με άλλα σύνολα δεδομένων

- Αλγόριθμοι διασύνδεσης

- ☞ String Matching

- e.g. Lexical Distance between labels

- ☞ Common Key Matching

- e.g. ISBN, Musicbrainz IDs

- ☞ Property-based Matching

- 2 αντικείμενα έχουν τον ίδιο τίτλο, τύπο και συντεταγμένες.

- Στοχεύστε σε αμφίδρομες σχέσεις

Εφαρμογές

- DBPedia - <http://wiki.dbpedia.org/>
- Geonames <http://www.geonames.org/> -
<http://www.geonames.org/ontology/documentation.html>
- EU Open Data - <https://data.europa.eu/data/sparql?locale=en>
- Europeana - <https://pro.europeana.eu/page/europeana-lod-pilot-edm>

Επιπλέον υλικό για μελέτη

- Chapters 2 and 3 of the Semantic Web Primer available from <http://www.csd.uoc.gr/~hy566/SWbook.pdf> .
- The following material from the Semantic Web Activity Web page on RDF <http://www.w3.org/RDF/> :
 - RDF Primer. The version on the above Web page uses RDF/XML; don't forget to see the version based on Turtle at <http://www.w3.org/2007/02/turtle/primer/> .
 - Resource Description Framework (RDF): Concepts and Abstract Syntax
- Check out the content published at the RDF namespace URI:
 - <http://www.w3.org/1999/02/22-rdf-syntax-ns#> where you will find an RDF Schema description of the RDF vocabulary given in RDF/XML!
- The DBpedia project (<http://dbpedia.org/About>), a nice application of RDF and Linked Data (<http://linkeddata.org/>).
- The specifications for choosing right URIs <https://www.w3.org/TR/cooluris/>

Thank you

The screenshot shows the GUNET-2 website interface for the 'Big Data Management' course. The header includes the GUNET-2 logo and a search bar. The main content area is titled 'Big Data Management' and features a 'Περιγραφή' (Description) section with the following text: 'Introduction - review of relational and object-relational databases. Modern trends in database design. Non-traditional data types (text, multimedia, spatial information). Non-traditional database architecture (sensor networks, data streams, distributed, in the cloud). The "big data" era (MapReduce architecture, etc.). Lab hours with MongoDB, Spark (Batch Processing, Streaming, MLib).' To the right of the description is a 'Ταυτότητα Μαθήματος' (Course Identity) section listing course code (CDS110), lecturers (N. Pelékis, G. Pappas), and other details. A sidebar on the left contains various navigation and utility tools. The footer includes the GUNET logo and the text 'Πληροφορίες Πνευματικών Δικαιωμάτων' and 'POWERED BY OPEN@CLASS'.



Αρχική Σελίδα » CDS110- Big Data Management » Σύνδεσμοι

Επιλογές Μαθήματος

- Ανακοινώσεις
- Ασκήσεις
- Ατζέντα
- Έγγραφα
- Πληροφορίες Μαθήματος
- Σύνδεσμοι

CDS110- Big Data Management

Σύνδεσμοι

- Γενικοί σύνδεσμοι
 - Ιστοσελίδα Data Science Lab. (DataStories)

Κατηγοριοποιημένοι σύνδεσμοι

- Books**
 - Bailis P, et al. (eds.) (2015) Readings in Database Systems
 - Codd EF (1990) The relational model for database management: version 2
 - Liu L, Özsu MT (eds.) (2009) Encyclopedia of Database Systems
- Papers**
 - Abadi D, et al. (2013) The Beckman report on database research
 - Abadi D, et al. (2018) The Seattle report on database research
 - Abiteboul S, et al. (2003) The Lowell database research self assessment
 - Agrawal R, et al. (2008) The Claremont report on database research
 - Codd EF (1970) A relational model of data for large shared data banks
- Posts**
 - Big Data Architecture: A Complete and Detailed Overview
 - HPI Genealogy of Relational Database Management Systems
- Videos, Tutorials etc.**
 - Learn PostgreSQL Tutorial - Full Course for Beginners
 - History of Databases