

Εργαστήριο Σημασιολογικού Ιστού

Ενότητα 6: RDF Schema (RDFS)

Μ.Στεφανιδάκης

29-3-2018

Τι μπορούμε να εκφράσουμε με την RDF;

- ▶ Δηλώσεις σε μορφή τριάδων (s,p,o)
 - ▶ Χωρίς οποιαδήποτε έννοια **δομής**...
- ▶ Παράδειγμα:

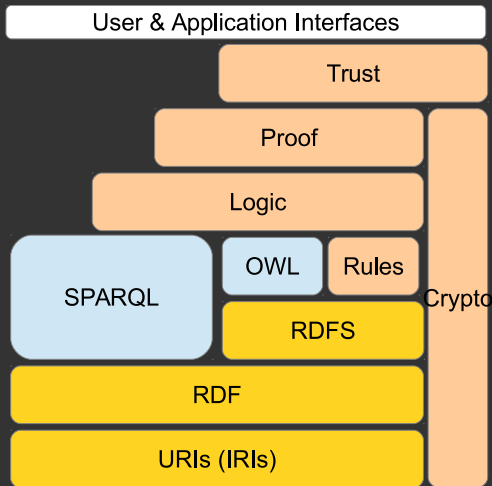
`<http://ex.com/SemWeb> <http://ex.com/vocab#teaches> <http://ex.com/E-Gov> .`

- ▶ Θεωρώντας ότι το `<http://ex.com/vocab#teaches>` σημαίνει “διδάσκει”
 - ▶ Έχει σημασία το παραπάνω;
 - ▶ Ή το επόμενο;

`<http://ex.com/Αίθουσα2> <http://ex.com/vocab#startsAt> "12:00:00.0"^^xsd:time .`

- ▶ Τι λείπει;
 - ▶ Ένας τρόπος να εκφράσουμε την οργάνωση (το “**σχήμα**”) των δεδομένων!

Τα επίπεδα του Σημασιολογικού Ιστού



RDFS: πρότυπο για την περιγραφή της οργάνωσης των δεδομένων (data-modelling)

RDF Schema (RDFS)

- ▶ Πρότυπο **σημασιολογικής επέκτασης** της RDF
 - ▶ Στην πραγματικότητα, μέρος του ενιαίου προτύπου RDF
- ▶ Παρέχει **λεξιλόγιο** (vocabulary) για να κάνουμε δηλώσεις
 - ▶ Για την περιγραφή **ομάδων** οντοτήτων
 - ▶ Και την περιγραφή **σχέσεων** μεταξύ των ομάδων αυτών
- ▶ Το “σχήμα” RDFS βασίζεται στις **κλάσεις** (classes) και τις **ιδιότητες** (properties)
- ▶ Μπορεί να επεκταθεί με πιο εξειδικευμένα λεξιλόγια

Χώροι ονομάτων RDFS

- ▶ Χρησιμοποιούμε τον χώρο ονομάτων

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

- ▶ τον οποίο κατα σύμβαση συμβολίζουμε με το πρόθεμα **rdfs:**

- ▶ Καθώς επίσης και τον χώρο ονομάτων

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

- ▶ το γνωστό πρόθεμα **rdf:**
- ▶ Και οι δύο χώροι χρησιμοποιούνται με ενιαίο τρόπο για την οντολογική περιγραφή των δεδομένων
 - ▶ Διατηρούνται για λόγους συμβατότητας με το παρελθόν

RDFS και Κλάσεις

- ▶ Οι οντότητες μπορούν να χωριστούν σε ομάδες, οι οποίες ονομάζονται **κλάσεις** (classes)
 - ▶ Τα μέλη κάθε κλάσης ονομάζονται **στιγμιότυπα** (instances) της κλάσης
 - ▶ Η σχέση **rdf:type** συνδέει τα στιγμιότυπα με την κλάση τους
- ▶ Οι κλάσεις είναι και αυτές οντότητες και περιγράφονται με URIs
- ▶ Μια οντότητα μπορεί να ανήκει σε περισσότερες από μία κλάσεις
 - ▶ Επίσης: δύο κλάσεις μπορούν να έχουν το ίδιο σύνολο στιγμιοτύπων και να παραμένουν ξεχωριστές!

Η διαφορά από τον αντικειμενοστρεφή προγραμματισμό

- ▶ **Προσοχή:** εδώ οι κλάσεις δεν προσδιορίζουν τις ιδιότητες που επιτρέπεται να έχουν τα στιγμιότυπα!
 - ▶ **όχι π.χ.** “κάθε στιγμιότυπο της κλάσης `Document` έχει ιδιότητα `author` που το συνδέει με ένα στιγμιότυπο της κλάσης `Person`”
- ▶ **Αντιθέτως,** εδώ περιγράφονται οι ιδιότητες – σε ποιες κλάσεις εφαρμόζονται
 - ▶ **δηλαδή π.χ.** “η ιδιότητα `author` συνδέει ένα στιγμιότυπο της κλάσης `Document` με ένα στιγμιότυπο της κλάσης `Person`”
- ▶ Επεκτάσιμο σχήμα, χωρίς να απαιτείται ο επαναπροσδιορισμός των κλάσεων!

Κλάσεις, Υποκλάσεις και Στιγμιότυπα

- ▶ Ιδιαιτερότητα κλάσεων RDFS
 - ▶ Μια κλάση μπορεί να είναι και στιγμιότυπο –ακόμα και του εαυτού της!
- ▶ Η κλάση όλων των RDFS κλάσεων ονομάζεται `rdfs:Class`
 - ▶ και είναι στιγμιότυπο του εαυτού της..
- ▶ Όταν μια κλάση C είναι υποκλάση μιας άλλης κλάσης C', τότε όλα τα στιγμιότυπα της C είναι επίσης στιγμιότυπα της C'
 - ▶ χρήση ιδιότητας `rdfs:subClassOf`

Κλάσεις που ορίζονται στο RDFS

- ▶ **rdfs:Resource**: κλάση των πάντων
 - ▶ Τα πάντα είναι στιγμιότυπα αυτής της κλάσης
 - ▶ Κάθε άλλη κλάση είναι υποκλάση της
 - ▶ Στιγμιότυπο της **rdfs:Class**
- ▶ **rdfs:Class**: η κλάση όλων των RDFS κλάσεων
 - ▶ Είναι στιγμιότυπο του εαυτού της
- ▶ **rdfs:Literal**: η κλάση όλων των σταθερών τιμών literals
 - ▶ Στιγμιότυπο της **rdfs:Class** και υποκλάση της **rdfs:Resource**

Κλάσεις που ορίζονται στο RDFS (2)

- ▶ **rdfs:Datatype**: η κλάση των τύπων δεδομένων
 - ▶ Κάθε στιγμιότυπό της είναι ένας τύπος (μια κλάση) δεδομένων (int, string, langString κ.ο.κ)
 - ▶ Κάθε στιγμιότυπό της είναι υποκλάση της **rdfs:Literal**
 - ▶ Είναι και στιγμιότυπο και υποκλάση της **rdfs:Class**
 - ▶ εφόσον τα στιγμιότυπά της είναι κλάσεις
- ▶ **rdf:Property**: η κλάση όλων των ιδιοτήτων RDF
 - ▶ Στιγμιότυπο της **rdfs:Class**

Ιδιότητες RDF/RDFS

- ▶ **Ιδιότητα** (property): μια **δυαδική** σχέση μεταξύ οντοτήτων
- ▶ Η έννοια των **Υπο-ιδιοτήτων** (subproperties)
 - ▶ Εάν μια ιδιότητα P είναι υπο-ιδιότητα μιας άλλης P', τότε όλα τα ζεύγη οντοτήτων που σχετίζονται μέσω της P, σχετίζονται επίσης μέσω της P'
 - ▶ η σχέση δηλώνεται μέσω της ιδιότητας **`rdfs:subPropertyOf`**

Ιδιότητες RDF/RDFS – πρακτικά

Δήλωση (τριάδα)	Σημασιολογία
C rdf:type rdfs:Class	η οντότητα C είναι κλάση RDF
P rdf:type rdf:Property	η οντότητα P είναι ιδιότητα RDF
R rdf:type C	η οντότητα R είναι στιγμιότυπο της κλάσης C
C rdfs:subClassOf C'	η κλάση C είναι υποκλάση της κλάσης C'
P rdfs:subPropertyOf P'	η ιδιότητα P είναι υπο-ιδιότητα της ιδιότητας P'
P rdfs:domain C	όλα τα υποκείμενα (subjects) που συνδέονται με την ιδιότητα P είναι στιγμιότυπα της κλάσης C
P rdfs:range C	όλα τα αντικείμενα (objects) που συνδέονται με την ιδιότητα P είναι στιγμιότυπα της κλάσης C

Παρατηρήσεις

- ▶ R **rdf:type** C
 - ▶ υπονοεί $R = \text{rdfs:Resource}$, $C = \text{rdfs:Class}$
- ▶ C **rdfs:subClassOf** C'
 - ▶ υπονοεί $C, C' = \text{rdfs:Class}$
 - ▶ μεταβατική ιδιότητα
- ▶ P **rdfs:subPropertyOf** P'
 - ▶ υπονοεί $P, P' = \text{rdf:Property}$
 - ▶ μεταβατική ιδιότητα
- ▶ P **rdfs:domain** C
- ▶ P **rdfs:range** C
 - ▶ υπονοούν $P = \text{rdf:Property}$, $C = \text{rdfs:Class}$
 - ▶ έχουν καθολική ισχύ, δεν μπορούν να περιοριστούν σε συγκεκριμένες κλάσεις μόνο!

RDFS και λεξιλόγια RDF

► Λεξιλόγια RDF (RDF Vocabularies)

- Παρέχουν όρους (“λέξεις”) για την περιγραφή των κλάσεων και σχέσεων των RDF δεδομένων μας
 - όλα είναι URIs σε κοινό χώρο ονομάτων
- Περιγράφονται μέσω τριάδων RDF
 - συνήθως σε ξεχωριστό έγγραφο RDF
- Μαζί με οντολογική πληροφορία RDFS για τους όρους
 - πώς χρησιμοποιούνται οι όροι!
 - ποιος όρος είναι κλάση ή ιδιότητα (`rdf:type`, `rdfs:Class`, `rdf:Property`)
 - υποκλάσεις και υπο-ιδιότητες (`rdfs:subClassOf`, `rdfs:subPropertyOf`)
 - τι δέχεται ως υποκείμενο ή αντικείμενο η κάθε ιδιότητα (`rdfs:domain`, `rdfs:range`)

Παράδειγμα: λεξιλόγιο FOAF

- ▶ Friend-of-a-Friend Project
 - ▶ Μια πρώιμη προσπάθεια κοινωνικής δικτύωσης (2000)
 - ▶ Χρησιμοποιεί λεξιλόγιο RDF για την περιγραφή σχέσεων
 - ▶ ποιος ξέρει ποιον
 - ▶ Συνέπεσε με την ανάπτυξη του Σημασιολογικού Ιστού
 - ▶ Και χρησιμοποιήθηκε ως βασικό παράδειγμα από όσους υποστήριζαν ότι τα URIs πρέπει να είναι πραγματικές διευθύνσεις στο Web
- ▶ Σήμερα, το λεξιλόγιο FOAF έχει μεγαλύτερη σημασία από το ίδιο το "Project FOAF"!
 - ▶ <http://xmlns.com/foaf/spec/>
 - ▶ δώστε το ως είσοδο στον **W3C RDF Validator!**

Παράδειγμα: λεξιλόγιο FOAF (2)

- ▶ χώρος διευθύνσεων FOAF:
`http://xmlns.com/foaf/0.1/`
- ▶ “η οντότητα foaf:Person είναι κλάση”
`foaf:Person` `rdf:type` `rdfs:Class` .
- ▶ “η οντότητα foaf:mbox είναι ιδιότητα”
`foaf:mbox` `rdf:type` `rdf:Property` .
- ▶ “η κλάση foaf:Person είναι υποκλάση της foaf:Agent”
`foaf:Person` `rdfs:subClassOf` `foaf:Agent` .
- ▶ “η ιδιότητα foaf:knows δέχεται ως αντικείμενα στιγμιότυπα της κλάσης foaf:Person”
`foaf:knows` `rdfs:range` `foaf:Person` .
- ▶ Μπορείτε να βρείτε στο FOAF παραδείγματα χρήσης των `rdfs:domain` και `rdfs:subPropertyOf`;

Η σειρά σας!

- ▶ Δημιουργήστε το δικό σας λεξιλόγιο RDF για τα δεδομένα του ωρολογίου προγράμματος
 - ▶ Σε ξεχωριστό έγγραφο, σε μορφή N-triples
 - ▶ namespace <http://host/sw/you/myvocab#>
- ▶ Περιγράψτε μέσω του RDFS
 - ▶ τις κλάσεις Διδάσκων, Διάλεξη, Αίθουσα, Μάθημα
 - ▶ τις ιδιότητες που χρησιμοποιείτε ως κατηγορήματα
 - ▶ τι πρέπει να έχει κάθε ιδιότητα ως υποκείμενο και αντικείμενο
 - ▶ ότι ο Διδάσκων είναι υποκλάση της foaf:Person
- ▶ Μετατρέψτε σε μορφή rdfxml και δείτε μέσω του W3C RDF Validator

```
rapper -i ntriples -o rdfxml yourfile.nt > yourfile.rdf.xml
```

- ▶ **Μην ξεχνάτε:** σε κανονικές συνθήκες πρέπει να αναζητήσετε υπάρχοντα λεξιλόγια RDF πριν φτιάξετε ένα δικό σας!