

Προγραμματισμός Σημασιολογικού Ιστού

Ενότητα 7: Χρήση RDFS και λεξιλογίων RDF

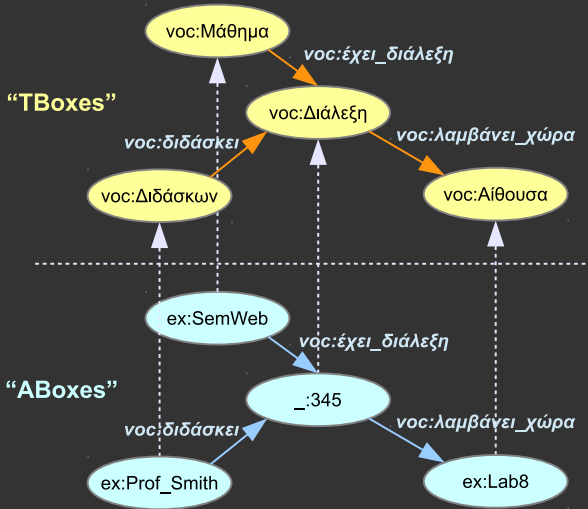
Μ.Στεφανιδάκης

3-4-2016

Επανάληψη: RDFS και λεξιλόγια RDF

- ▶ **RDFS: σημασιολογική επέκταση της RDF**
 - ▶ Το “σχήμα” RDFS βασίζεται στις **κλάσεις** (classes) και τις **ιδιότητες** (properties)
- ▶ **Λεξιλόγια RDF (RDF Vocabularies)**
 - ▶ Παρέχουν όρους (“λέξεις”) για την περιγραφή των κλάσεων και σχέσεων (του “σχήματος”, δηλαδή) των RDF δεδομένων μας
 - ▶ όλα είναι URIs σε κάποιον **χώρο ονομάτων** (namespace)
- ▶ Κλειδί για την επιτυχία του Σημασιολογικού Ιστού είναι η **(επανα)χρησιμοποίηση** κοινών λεξιλογίων
 - ▶ Η αξία ενός λεξιλογίου RDF αυξάνεται ανάλογα με τη χρήση του
 - ▶ **Πάντα:** πρέπει να αναζητήσετε υπάρχοντα λεξιλόγια RDF πριν φτιάξετε ένα δικό σας!

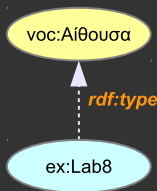
Δύο επίπεδα: Σχήμα και Δεδομένα



TBox (terminology) και **ABox** (assertion): όροι από την επιστήμη της Αναπαράστασης Γνώσης

Η σύνδεση Σχήματος και Δεδομένων

- ▶ Το σχήμα (στη μορφή λεξιλογίου RDF) συνήθως βρίσκεται σε ξεχωριστό έγγραφο
 - ▶ Οι τριάδες RDF περιγράφουν οντολογικές έννοιες (TBoxes)
- ▶ Τα “κανονικά” δεδομένα αποτελούν άλλο σετ τριάδων RDF (ABoxes)
 - ▶ Όπου χρησιμοποιούνται τα URIs του λεξιλογίου ως κατηγορήματα ή αντικείμενα
- ▶ Η σύνδεση επιτυγχάνεται μέσω του **rdf:type**



`ex:Lab` **rdf:type** `voc:Αίθουσα` .

Τι μπορεί να κάνει το RDFS για τις εφαρμογές μας;

- ▶ Επίσημα: **τίποτα!**
 - ▶ Δηλαδή: το πρότυπο RDFS **δεν ορίζει** πώς θα χρησιμοποιηθεί από μια εφαρμογή!
 - ▶ Ορίζει μόνο τη σημασιολογία...
- ▶ Οι χρήσεις όμως είναι πολλές και ενδιαφέρουσες:
 - ▶ Εφαρμογές που δρουν με βάση τη σημασιολογία των δεδομένων RDFS
 - ▶ Η ροή ελέγχου καθορίζεται από τα (μέτα)δεδομένα (data-driven applications)
 - ▶ Ανακάλυψη και έλεγχος συνοχής σε σημασιολογικά δεδομένα
 - ▶ Εάν ένα σετ δεδομένων τηρεί ένα “σχήμα”
 - ▶ Εξαγωγή νέας γνώσης πέραν των δηλωμένων τριάδων
 - ▶ Διαδικασία εξαγωγής συμπερασμάτων (inference)

Παράδειγμα εξαγωγής νέας γνώσης

Έστω το σετ δεδομένων RDF:

voc:Course rdf:type rdfs:Class .

voc:UndergraduateCourse rdf:type rdfs:Class .

voc:UndergraduateCourse rdfs:subClassOf voc:Course .

ex:SemWeb rdf:type voc:Course .

ex:CSIntro rdf:type voc:UndergraduateCourse .

Στην ερώτηση: “δώσε τα μαθήματα (voc:Course)”, τα **δηλωμένα** δεδομένα δίνουν:

- ▶ ex:SemWeb

Αν το πρόγραμμά μας υλοποιεί τους κανόνες RDFS, η απάντηση θα είναι:

- ▶ ex:SemWeb
- ▶ ex:CSIntro

Παράδειγμα εξαγωγής νέας γνώσης (2)

Έστω το σετ δεδομένων RDF:

```
voc:teaches rdfs:domain voc:Teacher .  
ex:Prof_Smith voc:teaches ex:Semweb .
```

Στην ερώτηση: “δώσε τους διδάσκοντες (voc:Teacher)”, τα **δηλωμένα** δεδομένα δεν δίνουν καμία απάντηση!

- ▶ Θα υπήρχε λύση μόνο αν δηλώναμε:

```
ex:Prof_Smith rdf:type voc:Teacher .
```

Αν το πρόγραμμά μας υλοποιεί τους κανόνες RDFS, η απάντηση θα είναι:

- ▶ ex:Prof_Smith

Μια στιγμή!

- ▶ Μήπως το λεξιλόγιο σας επιτρέπει μια διάλεξη να γίνεται ταυτόχρονα σε πολλές αίθουσες;
 - ▶ Πώς θα εκφράσουμε ότι μια συγκεκριμένη διάλεξη γίνεται **το πολύ σε μία** αίθουσα την ίδια στιγμή;
- ▶ Γενικότερα, αν ένα υποκείμενο μπορεί να συνδεθεί **ακριβώς N φορές** μέσω κατηγορήματος r με διάφορα αντικείμενα;
 - ▶ Το RDFS δεν μπορεί να εκφράσει αυτόν τον περιορισμό!

Επίσης...

- ▶ Πώς θα καθησυχάσετε τους “θεωρητικούς” πως
 - ▶ “Ό,τι είναι Αίθουσα δεν είναι Διδάσκων”
 - ▶ (πρακτικά το γνωρίζουμε, βέβαια)
- ▶ Το RDFS δεν μπορεί να εκφράσει το γεγονός ότι οι κλάσεις Αίθουσες και Διδάσκοντες δεν έχουν κοινά στοιχεία!

Επιπλέον...

- ▶ Αν είχατε κλάσεις LabLecture και TheoreticalLecture
- ▶ Και αντίστοιχα κλάσεις LabRoom και TeachingRoom
- ▶ Πώς θα δηλώνατε σε λεξιλόγιο RDF ότι:
 - ▶ όταν η ιδιότητα takesPlaceAt έχει domain LabLecture, τότε έχει range LabRoom και
 - ▶ όταν η ιδιότητα takesPlaceAt έχει domain TheoreticalLecture, τότε έχει range TeachingRoom
- ▶ Με το RDFS μπορούμε να ορίσουμε μόνο καθολικά το domain και range μιας ιδιότητας, όπως η takesPlaceAt, όχι ανα κλάση
- ▶ Για όλα τα παραπάνω, απαιτούνται πιο σύνθετα λεξιλόγια!