



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Μηχανική Γνώσης και Συστήματα Γνώσης

Περιγραφή πόρων Ιστού: RDF

Μανώλης Μαραγκουδάκης

Τμήμα Μηχανικών Πληροφοριακών και Επικοινωνιακών Συστημάτων



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

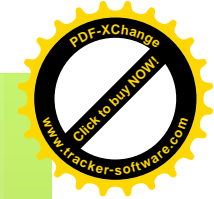


ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ



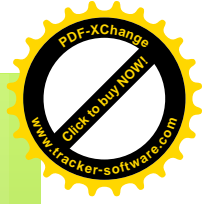
Ενότητα 3

Περιγραφή πόρων Ιστού: RDF

Διαφάνειες από το βιβλίο A Semantic
Web Primer

των

- Grigoris Antoniou
- Frank van Harmelen



Μειονεκτήματα της XML

- Η XML είναι καθολική μεταγλώσσα για να ορίζει σήμανση
- Παρέχει ένα ενιαίο πλαίσιο για ανταλλαγή δεδομένων και μεταδεδομένων μεταξύ εφαρμογών
- Ωστόσο, η XML δεν παρέχει κάποιο τρόπο καθορισμού της δημασιολογίας (σημασίας) των δεδομένων
- Π.χ. , δεν υπάρχει κάποιο νόημα που συνδέεται με την ένθεση των ετικετών.



Παράδειγμα σε XML

David Billington is a lecturer of Discrete Maths

```
<course name="Discrete Maths">
```

```
  <lecturer>David Billington</lecturer>
```

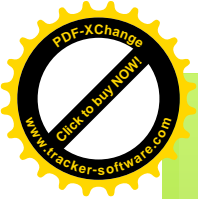
```
</course>
```

```
<lecturer name="David Billington">
```

```
  <teaches>Discrete Maths</teaches>
```

```
</lecturer>
```

Ανάποδη ένθεση, ίδια πληροφορία!



Βασικές ιδέες της RDF

- Βασικό δομικό μπλόκ: **αντικείμενο-ιδιότητα-τιμή** (τριπλέτα)
 - Ονομάζεται **πρόταση**
 - Η φράση για τον Billington είναι μια τέτοια πρόταση
- RDF έχει συντακτικό που βασίζεται στην XML
 - Κληρονομεί τα πλεονεκτήματα της XML
 - Είναι δυνατές και άλλες αναπαραστάσεις σε RDF



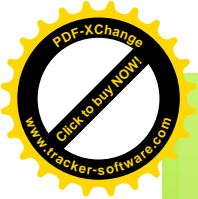
Βασικές ιδέες της RDF (2)

- Οι θεμελιώσεις έννοιες της RDF είναι:
 - Πόροι (resources)
 - Ιδιότητες (properties)
 - Προτάσεις (statements)



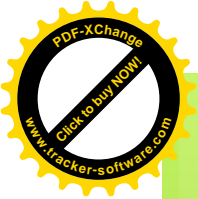
Πόροι

- Μπορούμε να δούμε έναν πόρο ως ένα αντικείμενο, ένα «πράγμα» για το οποίο θέλουμε να μιλήσουμε
 - Π.χ. συγγραφείς, βιβλία, εκδότες, μέρη, άνθρωποι, ξενοδοχεία, κτλ.
- Κάθε πόρος έχει ένα **URI**, έναν Universal Resource Identifier
- ένα URI μπορεί να είναι
 - Ένα URL (μια Web διεύθυνση) ή
 - Κάποιο άλλο είδος μοναδικού αναγνωριστικού



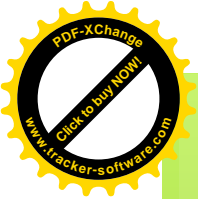
Ιδιότητες

- Αποτελούν ειδική περίπτωση πόρων
- Περιγράφουν σχέσεις μεταξύ των πόρων
 - Π.χ. “γραμμένο από”, “ηλικία”, “τίτλος”, κτλ.
- Και οι ιδιότητες καθορίζονται απο URIs
- Πλεονεκτήματα χρήσης URIs:
 - Ένα παγκόσμιο, μοναδικό σχήμα ονοματολογίας
 - Μειώνει σημαντικά το πρόβλημα της ομωνυμίας, που «μαστίζει» την αναπαράσταση των κατανεμημένων δεδομένων



Προτάσεις

- Οι προτάσεις βεβαιώνουν τις ιδιότητες των πόρων.
- Μια πρόταση είναι μια τριάδα object-attribute-value (αντικείμενο-ιδιότητα-τιμή)
 - Αποτελείται από έναν πόρο, μια ιδιότητα και μια τιμή
- Είναι τιμές μπορεί να είναι πόροι ή **λεκτικά** (literals)
 - Literals = ατομικές τιμές (strings)



3 Θεωρήσεις μιας πρότασης

- Μια τριπλέτα
- Ένα κομμάτι ενός γράφου
- Ένα κομμάτι XML κώδικα

Άρα ένα έγγραφο RDF μπορεί να θεωρηθεί ως:

- Ένα σύνολο τριάδων
- Ένας γράφος (ή σημασιολογικό δίκτυο-semantic net)
- Ένα XML έγγραφο

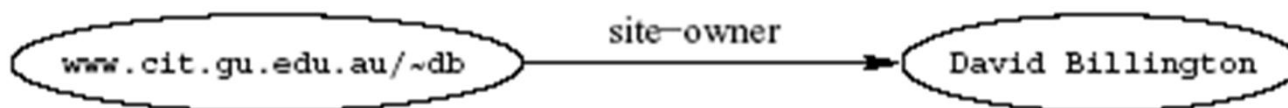


Προτάσεις σαν τριάδες

(<http://www.cit.gu.edu.au/~db>,
<http://www.mydomain.org/site-owner>,
#David Billington)

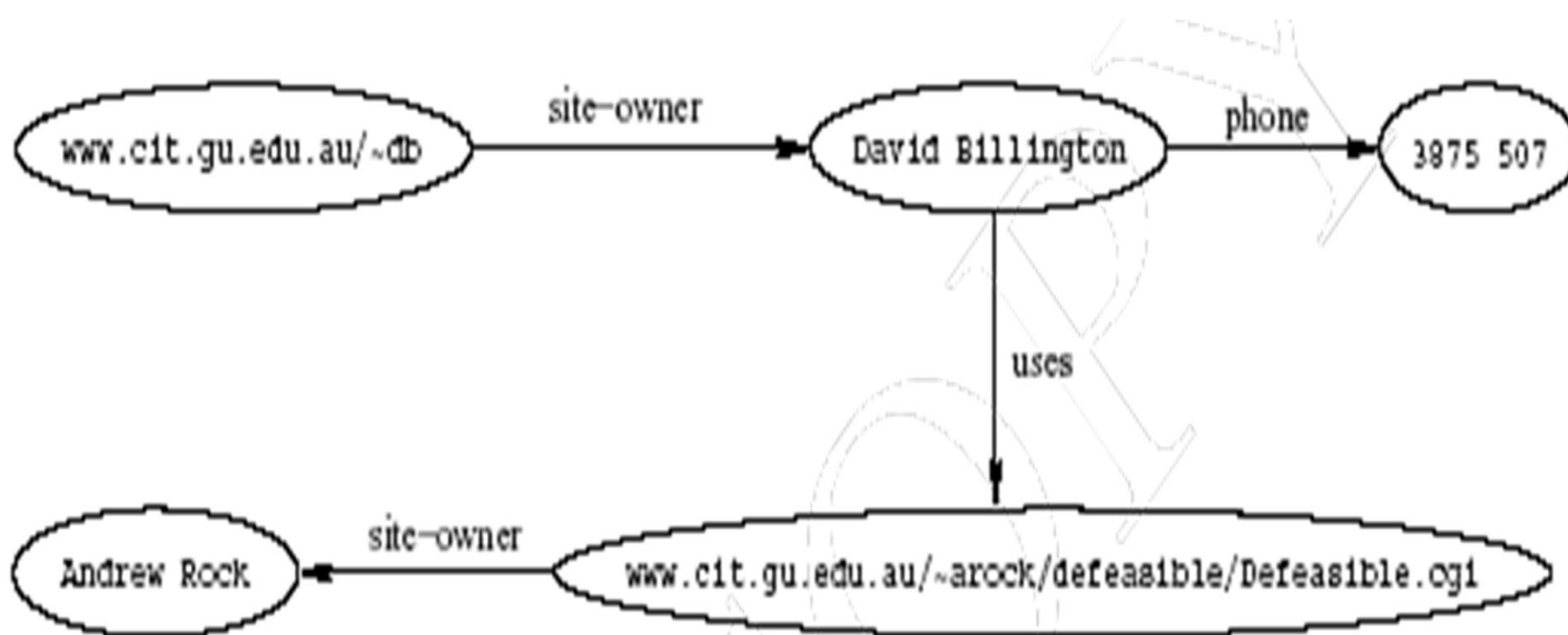
- Η τριάδα (x, P, y) είναι σαν ένα λογικό τύπο $P(x, y)$
 - Το δυαδικό κατηγορημα P σχετίζει το αντικείμενο x με το αντικείμενο y
 - Η RDF προσφέρει μόνο **δυαδικά κατηγορήματα** (ιδιότητες)

Σημασιολογικά Δίκτυα



- Ένας κατευθυνόμενος γράφος με επισημειωμένους κόμβους και ακμές
 - Από τον πόρο (το υποκείμενο της πρότασης)
 - Προς την τιμή (το αντικείμενο της πρότασης)
- Στην Τεχνητή Νοημοσύνη είναι γνωστό ως *Σημασιολογικό Δίκτυο*
- Η τιμή μιας πρότασης μπορεί να είναι ένας πόρος
 - Άρα μπορεί να συνδέεται και με άλλους πόρους

Παράδειγμα





Οι προτάσεις σε XML

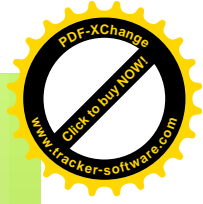
- Οι γράφοι είναι ισχυρά εργαλεία για κατανόηση από ανθρώπους **αλλά**
- Ο Σημασιολογικός Ιστός οραματίζεται αναπαραστάσεις που είναι προσπελάσιμες και επεξεργάσιμες από μηχανές
- Υπάρχει και μια Τρίτη αναπαράσταση βασισμένη σε XML
 - Όμως η XML είναι σειριακή διάταξη, δεν ταιριάζει με το μοντέλο της RDF
 - Άρα το πραγματικό μοντέλο της RDF είναι οι γράφοι



Προτάσεις σε XML (2)

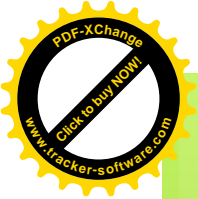
```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:mydomain="http://www.mydomain.org/my-rdf-ns">

  <rdf:Description
    rdf:about="http://www.cit.gu.edu.au/~db">
    <mydomain:site-owner
      rdf:resource="#David Billington"/>
    </rdf:Description>
  </rdf:RDF>
```

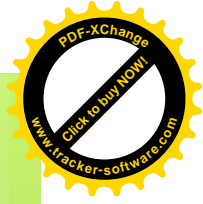
Προτάσεις σε XML (3)

- Ένα έγγραφο RDF ξεκινά με την ετικέτα `rdf:RDF`
- Το περιεχόμενο αυτού του στοιχείου είναι ένας αριθμός από **περιγραφές**, που χρησιμοποιούν ετικέτες `rdf:Description`.
- Κάθε περιγραφή κάνει μια πρόταση για έναν πόρο, ο οποίος προσδιορίζεται με 3 τρόπους:
 - Ένα χαρακτηριστικό (attribute) `about`, που αναφέρεται σε έναν υπάρχοντα πόρο
 - Ένα χαρακτηριστικό `ID`, που δημιουργεί έναν νέο πόρο
 - Χωρίς ονομασία, που δημιουργεί έναν ανώνυμο πόρο



Προτάσεις σε XML (4)

- Το στοιχείο `rdf:Description` κάνει μια πρόταση για τον πόρο `http://www.cit.gu.edu.au/~db`
- Μέσα στην περιγραφή
 - Η ιδιότητα χρησιμοποιείται ως ετικέτα (`mydomain:site-owner`)
 - Το περιεχόμενο της ετικέτας είναι η τιμή της ιδιότητας)



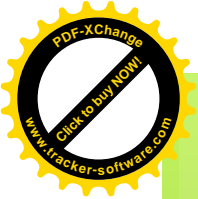
Υποστασιοποίηση

- Στην RDF είναι εφικτό να κάνουμε προτάσεις σχετικές με προτάσεις
 - **Ο Γρηγόρης πιστεύει ότι ο David Billington είναι ο δημιουργός της σελίδας**
<http://www.cit.gu.edu.au/~db>
- Η προσέγγιση είναι να θέτουμε ένα μοναδικό identifier σε κάθε πρόταση
 - Χρησιμοποιείται για να αναφερόμαστε στην πρόταση αυτή



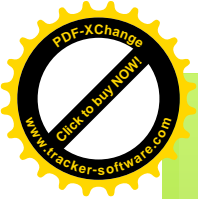
Υποστασιοποίηση (2)

- Εισάχουμε ένα βοηθητικό αντικείμενο (π.χ. belief1)
- Το συσχετίζουμε με κάθε ένα από το 3 μέρη της αρχικής πρότασης μέσω των ιδιοτήτων subject, predicate και object
- Στο προηγούμενο παράδειγμα
 - **subject** του belief1 = David Billington
 - **predicate** του belief1 = δημιουργός
 - **object** του belief1 = <http://www.cit.gu.edu.au/~db>



Τύποι Δεδομένων

- Οι τύποι δεδομένων χρησιμοποιούνται στον προγραμματισμό για επιτρέψουν την κατανόηση των μεταβλητών
- Στην RDF, χρησιμοποιούμε τυποποιημένα λεκτικά
 - **Π.χ.** (#David Billington,[http://www.mydomain.org/age,\"27\"^http://www.w3.org/2001/XMLSchema#integer](http://www.mydomain.org/age,\))



Τύποι Δεδομένων (2)

- ^^-σηματοδοτεί τον τύπο του λεκτικού
- Στην πράξη, χρησιμοποιούμε XMLSchema
- XML Schema υποστηρίζει μια μεγάλη γκάμα τύπων δεδομένων
 - Π.χ. Booleans, integers, floating-point numbers, times, dates, κτλ.

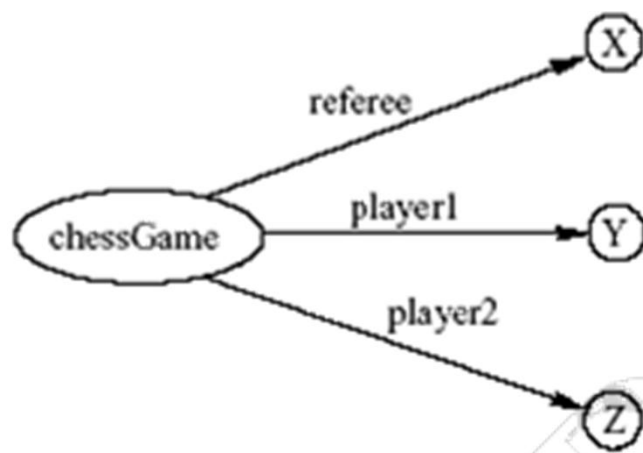


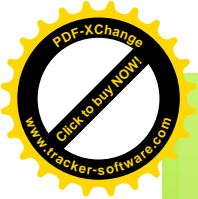
Κριτική της RDF

- Η RDF χρησιμοποιεί μόνον δυαδικές ιδιότητες
 - Σοβαρό μειονέκτημα
 - Όμως όλα μπορούν να αναπαρασταθούν με δυαδικά κατηγορήματα
- Παράδειγμα: referee(X,Y,Z)
 - X είναι διαιτητής σε ένα αγώνα μεταξύ των Y και Z

Κριτική της RDF (2)

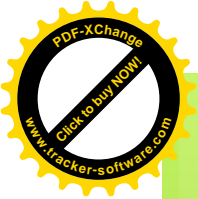
- εισάγουμε:
 - Έναν νέο βοηθητικό πόρο chessGame
 - Τα δυαδικά κατηγορήματα είναι τα ref, player1, και player2
- Το referee(X,Y,Z) γίνεται:





XML Σύνταξη της RDF

- Ένα έγγραφο RDF αποτελείται από ένα στοιχείο `rdf:RDF`
 - Το περιεχόμενο αυτού του στοιχείου είναι ένας αριθμός από περιγραφές



Παράδειγμα από μαθήματα πανεπιστημίου

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:uni="http://www.mydomain.org/uni-ns">
```

```
  <rdf:Description rdf:about="949318">
```

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

```
    <uni:title>Associate Professor</uni:title>
```

```
    <uni:age rdf:datatype="&xsd:integer">27</uni:age>
```

```
  </rdf:Description>
```



Παράδειγμα από μαθήματα πανεπιστημίου (2)

```
<rdf:Description rdf:about="CIT1111">  
  <uni:courseName>Discrete  
  Maths</uni:courseName>  
  <uni:isTaughtBy>David  
  Billington</uni:isTaughtBy>  
</rdf:Description>
```

```
<rdf:Description rdf:about="CIT2112">  
  <uni:courseName>Programming  
  III</uni:courseName>  
  <uni:isTaughtBy>Michael  
  Maher</uni:isTaughtBy>  
</rdf:Description>
```



Στοιχεία ιδιοτήτων

- Το περιεχόμενο των `rdf:Description` στοιχείων

```
<rdf:Description rdf:about="CIT3116">  
  <uni:courseName>Knowledge  
  Representation</uni:courseName>  
  <uni:isTaughtBy>Grigoris Antoniou</uni:isTaughtBy>  
</rdf:Description>
```

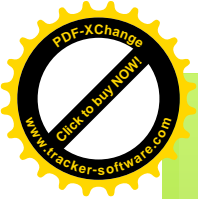
- `uni:courseName` και `uni:isTaughtBy` ορίζουν 2 ζεύγη ιδιότητας-τιμής για το CIT3116 (2 RDF προτάσεις)



Τύποι δεδομένων

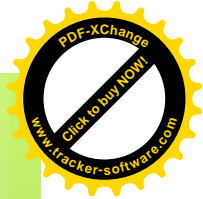
- Η ιδιότητα `rdf:datatype="&xsd:integer"` χρησιμοποιείται για να μας δείξει τον τύπο δεδομένων της ιδιότητας `age`

```
<rdf:Description rdf:about="949318">  
  <uni:name>David Billington</uni:name>  
  <uni:title>Associate Professor</uni:title>  
  <uni:age rdf:datatype="&xsd:integer">27</uni:age>  
</rdf:Description>
```



Η ιδιότητα `rdf:resource`

- Οι σχέσεις μεταξύ μάθημάτων και διδασκόντων δεν έχουν επίσημα ορισθεί αλλά υπάρχουν έμμεσα μέσω της σχέσης του ίδιου ονόματος
- Η χρήση του ίδιου ονόματος μπορεί όμως να είναι μια σύμπτωση για τη μηχανή
- Μπορούμε να σημειώσουμε ότι 2 οντότητες είναι ίδιες με τη χρήση της ιδιότητας `rdf:resource`



Η ιδιότητα `rdf:resource` (2)

```
<rdf:Description rdf:about="CIT1111">  
  <uni:courseName>Discrete  
  Mathematics</uni:courseName>  
  <uni:isTaughtBy rdf:resource="949318"/>  
</rdf:Description>  
  
<rdf:Description rdf:about="949318">  
  <uni:name>David Billington</uni:name>  
  <uni:title>Associate Professor</uni:title>  
</rdf:Description>
```



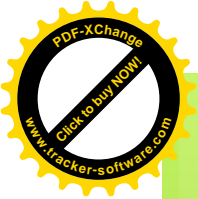
Αναφερόμενοι σε εξωτερικά ορισμένους πόρους

- Π.χ., για να αναφερθούμε στον εξωτερικά ορισμένο πόρο CIT1111:
<http://www.mydomain.org/uni-ns#CIT1111>
ως τιμή της ιδιότητας `rdf:about`
- www.mydomain.org/uni-ns είναι το URI που υπάρχει ο ορισμός του CIT1111
- Ένα `rdf:ID` ορίζει ένα τμήμα του URI, που μπορεί να χρησιμοποιηθεί για να αναφέρει την οριζόμενη περιγραφή



Παράδειγμα εμφωλευμένων περιγραφών

```
<rdf:Description rdf:about="CIT1111">  
  <uni:courseName>Discrete Maths</uni:courseName>  
  <uni:isTaughtBy>  
    <rdf:Description rdf:ID="949318">  
      <uni:name>David Billington</uni:name>  
      <uni:title>Associate Professor</uni:title>  
    </rdf:Description>  
  </uni:isTaughtBy>  
</rdf:Description>
```



Εμφωλευμένες περιγραφές

- Οι περιγραφές μπορούν να ορίζονται μέσα σε άλλες περιγραφές
- Άλλα μαθήματα όπως το CIT3112, μπορούν να αναφερθούν στο πόρο με ID 949318
- Αν η περιγραφή γίνει μέσα σε μια άλλη περιγραφή, η εμβέλεια παραμένει καθολική!