



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΑΝΟΙΚΤΑ ακαδημαϊκά
μαθήματα ΠΠ

Αναπαράσταση Γνώσης στον Παγκόσμιο Ιστό

Ενότητα 1: Εισαγωγή

Το όραμα του Σημασιολογικού Ιστού

Ιωάννης Χατζηλυγερούδης

Πολυτεχνική Σχολή

Τμήμα Μηχ/κών Η/Υπολογιστών & Πληροφορικής

Περιεχόμενα ενότητας

1. Περιεχόμενα Μαθήματος
2. Ο σημερινός Ιστός (Syntactic Web)
3. Μηχανές Αναζήτησης
4. Σημαντικός Ιστός (Semantic Web)
5. Διαχείριση Γνώσης - Ιστός
6. Διαχείριση Γνώσης - Σημ. Ιστός



Περιεχόμενα ενότητας

7. Ηλεκτρονικό Εμπόριο - Ιστός
8. Ηλεκτρονικό Εμπόριο - Σημ. Ιστός
9. Τεχνολογίες Σημαντικού Ιστού
10. Μεταδεδομένα
11. Σύγκριση HTML - XML
12. RDF (Οντότητα - Ιδιότητα - Τιμή)
13. Πυραμίδα Σημαντικού Ιστού



Περιεχόμενα Μαθήματος

- Εισαγωγή
- XML-XML Schema
- RDF-RDF Schema
- Οντολογίες (γλώσσα OWL)
- Περιγραφικές Λογικές
- Συλλογιστές (Reasoners)
- Κανόνες



Ο σημερινός Ιστός (Syntactic Web)

- Πολύ σημαντική εξέλιξη - Αλλαγή προοπτικής των Η/Υ
- Βασικές εξυπηρετήσεις είναι: αναζήτηση και χρήση πληροφοριών, αναζήτηση και επαφή με άλλα άτομα, παραγγελία προϊόντων κλπ
- Δεν υπάρχει επαρκής υποστήριξη από εργαλεία λογισμικού- κύρια εργαλεία χρήσης: **μηχανές αναζήτησης**



Μηχανές Αναζήτησης

- Βασίζονται σε λέξεις-κλειδιά
- Προβλήματα
 - Υψηλή ανάκληση, χαμηλή ακρίβεια
 - Χαμηλή ή καθόλου ανάκληση (σπάνια)
 - Εξάρτηση από τις λέξεις
 - Επιστρέφουν μεμονωμένες σελίδες
- Μετά την αναζήτηση δεν υπάρχει άλλη βοήθεια (π.χ. υποστήριξη ανάκλησης πληροφοριών)



Σημαντικός Ιστός (Semantic Web)

- Κίνητρο: Περιεχόμενο ιστοσελίδων μη άμεσα επεξεργάσιμο από μηχανές
- Πιθανή λύση: Χρήση υπαρχουσών μεθόδων TN-δύσκολη και πολύπλοκη
- Προτεινόμενη: Δημιουργία νέων μεθόδων για την περιγραφή/αναπαράσταση του περιεχόμενου
- Οργανισμός: Κοινοπραξία Παγκοσμίου Ιστού (WWW Consortium-W3C)



Διαχείριση Γνώσης - Ιστός

- Αναζήτηση πληροφοριών: εξαρτάται από τις μηχανές αναζήτησης δηλ. λέξεις-κλειδιά.
- Εξαγωγή πληροφοριών: Απαιτείται ανθρώπινη προσπάθεια και χρόνος.
- Συντήρηση πληροφοριών: Προβλήματα ασυνέπειας σε ορολογίες.
- Αποκάλυψη πληροφοριών: Γίνεται με μεθόδους εξόρυξης δεδομένων που δεν είναι πάντα αποτελεσματικές (π.χ. κατανεμημένη πληροφορία).
- Προβολή (view) πληροφοριών: ο περιορισμός προσπέλασης είναι δύσκολο να υλοποιηθεί στον ιστό.



Διαχείριση Γνώσης - Σημ. Ιστός

- Γνώση οργανωμένη σε εννοιολογικές περιοχές.
- Αυτοματοποίηση συντήρησης γνώσης.
- Ερωτήματα αντί λέξεων-κλειδιών.
- Οι απαντήσεις θα παρουσιάζονται με τρόπο φιλικό και δομημένο.
- Υποστήριξη απαντήσεων ερωτημάτων σε περισσότερα από ένα έγγραφα.
- Δυνατότητα καθορισμού δικαιωμάτων πρόσβασης σε έγγραφα.



Ηλεκτρονικό Εμπόριο - Ιστός

- Στις περισσότερες περιπτώσεις ο καταναλωτής επισκέπτεται τα διάφορα καταστήματα και «ξεφυλλίζει» τις τιμές, τις προσφορές κλπ ώστε να κάνει την επιλογή του.
- Η δημιουργία shopbots διευκολύνει, αλλά έχει αρκετές δυσκολίες και ανάπτυξης και εκπλήρωσης των αναγκών των πελατών.
- Η εξαγωγή πληροφοριών στηρίζεται σε ανάλυση λέξεων-κειμένων και ευρετικές προσεγγίσεις (εξαγωγή πληροφοριών περιορισμένης έκτασης)



Ηλεκτρονικό Εμπόριο - Σημ. Ιστός

- Πληροφορίες τιμολόγησης και προϊόντων μπορούν να εξάγονται σωστά, χωρίς λάθη.
- Πρόσθετες πληροφορίες π.χ. σχετικά με την υπόληψη των ηλεκ-καταστημάτων μπορούν να ανακτώνται από ανεξάρτητες πηγές.
- Το λογισμικό των shobots θα είναι λιγότερα πολύπλοκο.
- Πιο προχωρημένοι πράκτορες θα μπορούν να κατασκευαστούν, που θα διεξάγουν ακόμη και διαπραγματεύσεις.



Τεχνολογίες Σημαντικού Ιστού

- Μεταδεδομένα (Metadata)
- Οντολογίες (Ontologies)
- Λογική (Logic)
- Πράκτορες (Agents)



Μεταδεδομένα (1 από 3)

- Υπάρχουσες ιστοσελίδες σε HTML
 - Προσβάσιμες από τον άνθρωπο μόνο
 - Ακατάλληλες για αυτόματη πρόσβαση από μηχανές
 - Ακόμη και προχωρημένα συστήματα (π.χ. υπερ-ευφυείς πράκτορες) δεν μπορούν να έχουν επαρκή πρόσβαση
 - Προσβασιμότητα → Κατανοησιμότητα



Μεταδεδομένα (2 από 3)

- Παράδειγμα-Προσέγγιση Web (HTML)

```
<H1>Alex Physiotherapy Center</H1>
```

```
Welcome to the home page of Alex Physiotherapy  
Center. Our experienced therapists Mr John Carry  
and Miss Maria Stand as well as our secretary Miss  
Linda Johnson are available to serve you.
```

```
<H2>Consultation hours</H2>
```

```
Monday 11am - 7pm <br>
```

```
Tuesday 11am - 7pm <br>
```

```
...
```

```
We do not offer consultation during the week of  
<a href = ...> Panachaikoi Agones </a> games
```



Μεταδεδομένα (3 από 3)

- Παράδειγμα-Προσέγγιση Semantic Web (XML)

```
<company>
  <treatmentOffered>Phisiotherapy</treatmentOffered>
  <companyName>Alex Phisiotherapy Center</companyName>
  <staff>
    <therapist>John Carry</therapist>
    <therapist>Maria Stand</therapist>
    <secretary>Linda Johnson</secretary>
  </staff>
</company>
```



Σύγκριση HTML - XML

HTML

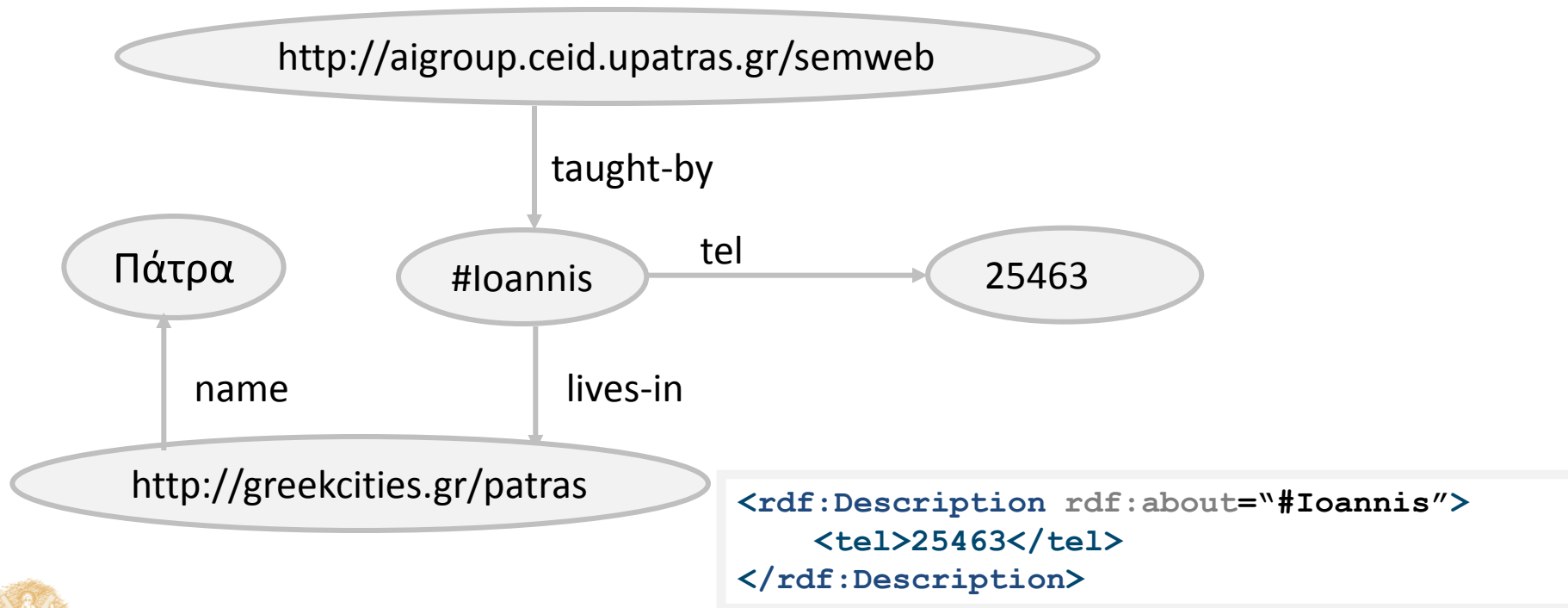
```
<H1>Αναπαράσταση Γνώσης</H1>  
<UL>  
  <LI>Καθηγητής: Ιωάννης Χατζής </LI>  
  <LI>Επίπεδο: Προπτυχιακό </LI>  
</UL>
```

XML

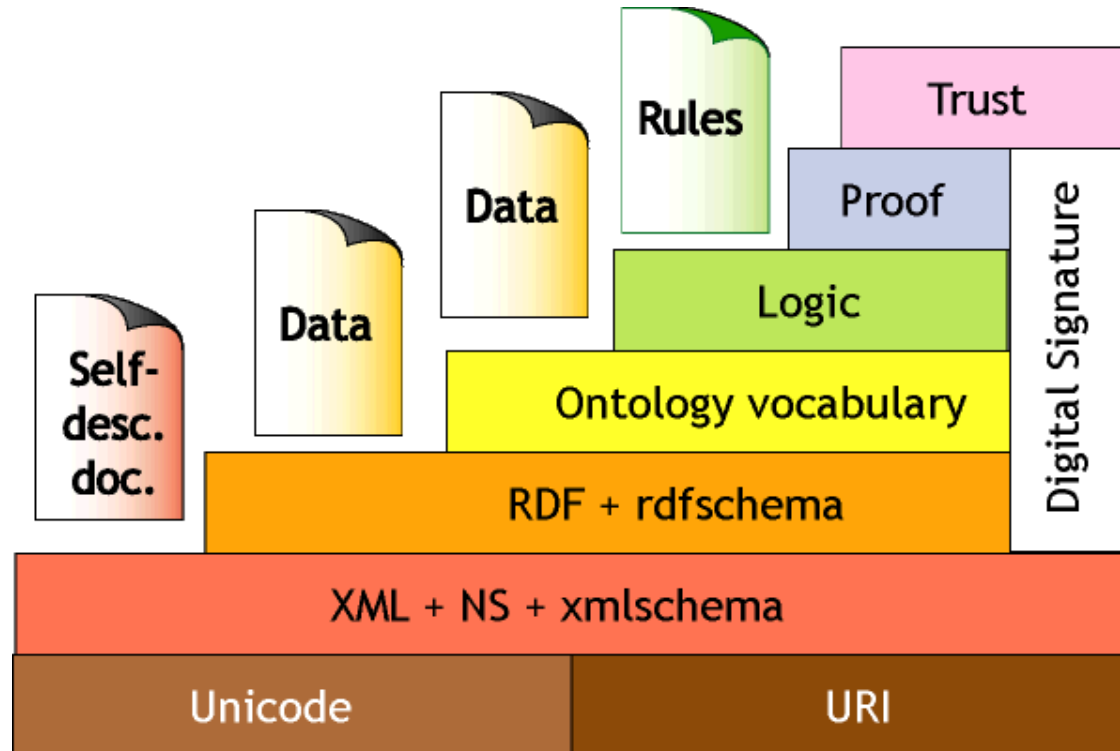
```
<course>  
  <title> Αναπαράσταση Γνώσης </title>  
  <teacher> Ιωάννης Χατζής </teacher>  
  <level> Προπτυχιακό </level>  
</course>
```


RDF (Οντότητα - Ιδιότητα - Τιμή)

- RDF: παρέχει πληροφορίες για πόρους στο διαδίκτυο
- **Object -> Attribute-> Value**

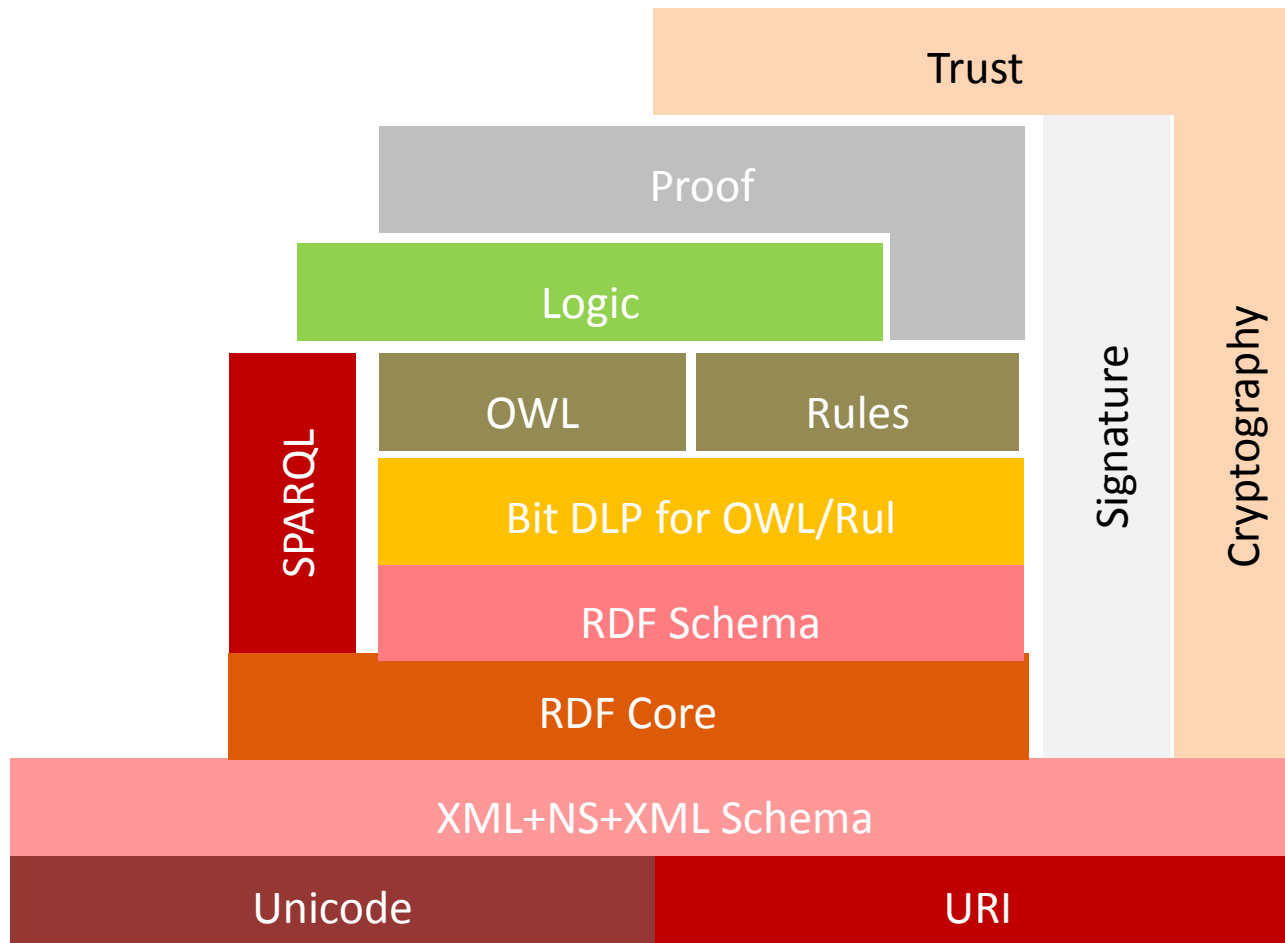


Πυραμίδα Σημαντικού Ιστού



Εικ. 1: The Berners-Lee Semantic WXML,” at XML eb ‘Birthday Cake’
Tim Berners-Lee, “Semantic Web on 2000, December 6, Washington, DC.
<http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>

Πυραμίδα Σημαντικού Ιστού



Εικ. 2: Semantic Web Architecture

Tim Berners-Lee, "Web for real people," April 2005.

<http://www.w3.org/2005/Talks/0511-keynote-tbl/>

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την 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.



Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.



Σημείωμα Αναφοράς

Copyright: Πανεπιστήμιον Πατρών, Ιωάννης Χατζηλυγερούδης, 2015.
«Αναπαράσταση Γνώσης στον Παγκόσμιο Ιστό. Εισαγωγή». Έκδοση: 1.0.
Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:
<https://eclass.upatras.gr/courses/CEID1098/>



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων (1/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνα 1: Tim Berners-Lee, “Semantic Web on XML,” at *XML 2000*, December 6, Washington, DC.

<http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide10-0.html>

Εικόνα 2: Tim Berners-Lee, “Web for real people,” April 2005.

<http://www.w3.org/2005/Talks/0511-keynote-tbl/>



Σημείωμα Χρήσης Έργων Τρίτων (2/2)

- Οι διαφάνειες είναι κατά μεγάλο μέρος βασισμένες στο βιβλίο "Εισαγωγή στον Σημασιολογικό Ιστό", των Grigoris Antoniou και Frank van Harmelen, Β' Έκδοση, 2009 (Ελληνική Έκδοση).

