

Τ.Ε.Ι. ΠΑΤΡΑΣ
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
ΑΝΑΠΤΥΞΗ ΕΝΟΣ ΟΝΤΟΛΟΓΙΚΟΥ PORTAL

ΓΚΕΝΕ ΜΑΡΙΝΑ
ΛΕΟΥ ΤΡΙΑΝΤΑΦΥΛΛΙΑ
ΜΠΑΝΙΩΤΗ ΜΑΡΙΑ

ΕΠΟΠΤΕΥΩΝ ΚΑΘΗΓΗΤΗΣ :ΠΑΠΑΙΩΝΝΟΥ ΒΑΙΟΣ

ΠΑΤΡΑ - 2009

Copyright © Γκενέ Μ., Λέου Τ., Μπανιώτη Μ., 2009
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της πτυχιακής εργασίας από το τμήμα Ε.Σ.Π.Σ. της Σ.Δ.Ο. του
Α.Τ.Ε.Ι. Πατρών δεν υποδηλώνει απαραίτητως και την αποδοχή των απόψεων
των συγγραφέων εκ μέρους του τμήματος.

ΠΡΟΛΟΓΟΣ

Η οντολογία στην πληροφορική αφορά σε αναπαραστάσεις οντοτήτων, ιδεών, γεγονότων μαζί με τις ιδιότητες και τις σχέσεις τους σύμφωνα με ένα σύστημα κατηγοριών. Έχει να κάνει με τον όρο οντολογία στην φιλοσοφία, με την έννοια της αναζήτησης της «ουσίας» ή της «ύπαρξης». Δηλαδή κατά πόσο η ταξινόμηση και η δομή των οντολογικών web portals συνιστά όχι απλά ένα μηχανισμό, αλλά έναν πολύπλοκο «οργανισμό».

ΠΕΡΙΕΧΟΜΕΝΑ

<i>ΕΙΣΑΓΩΓΗ - ΣΤΟΧΟΣ ΕΡΓΑΣΙΑΣ - ΠΕΡΙΓΡΑΦΗ ΕΡΓΑΣΙΑΣ</i>	1
<i>ΚΕΦΑΛΑΙΟ 1: ΟΝΤΟΛΟΓΙΕΣ</i>	5
1.1 Ορισμός Οντολογιών.....	5
1.2 Ανάγκες δημιουργίας οντολογιών	7
1.3 Βασικά πλεονεκτήματα/δυνατότητες των οντολογιών.....	9
1.4 Βασικά μειονεκτήματα / προβλήματα των οντολογιών	10
1.5 Τα κύρια συστατικά των οντολογιών.....	11
1.5.1 Κλάσεις (Classes ή Concepts).....	11
1.5.2 Άτομα (Individuals) ή Στιγμιότυπα (Instances).....	15
1.5.3 Σχέσεις (Relations).....	15
1.5.4 Συναρτήσεις (Functions).....	18
1.5.5 Αξιώματα (Axioms)	19
1.6 Κατηγορίες οντολογιών.....	19
1.7 Κριτήρια σχεδιασμού οντολογιών.....	20
1.8 Τεχνολογίες δημιουργίας οντολογιών	22
1.8.1 Μεθοδολογία ανάπτυξης οντολογιών – Κατασκευή της οντολογίας : 22	
1.8.2 Σημασιολογικά Εργαλεία	23
1.8.3 Έξυπνοι Πράκτορες.....	25
1.8.4 Μηχανές Αναζήτησης	27
1.8.5 Robust Hyperlinks.....	28
1.9 Μηχανική οντολογιών (Engineering ontology).....	29
1.10 Ευθυγράμμιση - Συσχέτιση οντολογιών	34

1.11 Συγχώνευση οντολογιών	37
1.12 Μερικές εφαρμογές των οντολογιών	40
1.12.1 Οι οντολογίες στο ηλεκτρονικό εμπόριο	40
1.12.2 E-learning ontology (Ηλεκτρονική μάθηση οντολογιών).....	41
1.12.3 Ψηφιακές Βιβλιοθήκες.....	42
1.12.4 Ιατρική – βιολογία – βιοϊατρική	45
1.13 Το Protégé ως εργαλείο εφαρμογής οντολογιών	46
ΚΕΦΑΛΑΙΟ 2: ΓΛΩΣΣΕΣ ΑΝΑΠΑΡΑΣΤΑΣΗΣ ΟΝΤΟΛΟΓΙΩΝ.....	49
2.1 Που διακρίνονται οι γλώσσες αναπαράστασης οντολογιών	49
2.2 Που διαφέρουν οι γλώσσες αναπαράστασης οντολογιών	50
2.3 Web – based γλώσσες	51
2.4 Λίγα λόγια για την γλώσσα HTML	52
2.5 Η γλώσσα XML	53
2.5.1 Πλεονεκτήματα χρήσης της XML	55
2.5.2 Αδυναμίες χρήσης της XML.....	56
2.6 Η γλώσσα RDF/S- (Resource Description Framework- Πλαίσιο Περιγραφής Πόρων)	56
2.6.1 Γραφική απεικόνιση RDF/ S.....	58
2.7 Πότε εμφανίζεται η OWL – Γλώσσα Οντολογίας του Παγκόσμιου Ιστού	60
2.8 Η γλώσσα OWL (Ontology Web Language- Γλώσσα Οντολογίας του Παγκόσμιου Ιστού)	61
2.8.1 Οι τρεις υπογλώσσες της OWL.....	62
2.8.2 Βασικά δομικά στοιχεία της OWL.....	64

2.8.3	Στοιχεία που επεκτείνουν την OWL Lite σε OWL DL και OWL Full	70
2.9	Η γλώσσα OIL (Ontology Interchange Language- Στρώμα Διεπαφής Οντολογίας).....	72
2.9.1	Βασική δομή διαστρωμάτωσης της OIL.....	72
2.10	Η DAML-OIL γλώσσα (Darpa Agent Markup Language).....	74
2.10.1	Που συναντάμε την DAML+OIL	75
2.11	γλώσσα XOL (Ontology eXchange Language)	76
2.12	Η SHOE γλώσσα (Simple Html Ontology Extensions).....	77
2.12.1	Που βρίσκει εφαρμογές η SHOE γλώσσα	78
2.13	Άλλες γλώσσες βιβλιοθηκών οντολογιών.....	78
2.14	Αξιολόγηση γλωσσών αναπαράστασης οντολογιών	80
ΚΕΦΑΛΑΙΟ 3: ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ – ΠΑΓΚΟΣΜΙΟΣ ΙΣΤΟΣ		82
3.1	Το Διαδίκτυο (Ιστός) σήμερα.....	82
3.2	Ορισμός σημασιολογικού ιστού (Semantic Web).....	83
3.3	Σχέση σημασιολογικού με παγκόσμιο ιστό	85
3.4	Οι πράκτορες Λογισμικού	87
3.4.1	Χαρακτηριστικά πρακτόρων.....	91
3.4.2	Σύγκριση ειδών πρακτόρων	92
3.5	Η αρχιτεκτονική του σημασιολογικού ιστού	94
3.5.1	Κωδικοποίηση χαρακτήρων.....	94
3.5.2	Διευθυνσιοδότηση εγγράφων.....	95
3.5.3	Extensible Markup Language (XML).....	96

3.5.4	Ανταλλαγή δεδομένων (RDF – Resource Description Framework)	96
3.5.5	Resource Description Framework Schema (RDFS)	97
3.5.6	Οντολογίες : OWL	98
3.5.7	Κανόνες	98
3.5.8	Αιτήματα	99
3.5.9	Ενοποιημένη Λογική.....	99
3.5.10	Αποδείξεις	100
3.5.11	Αξιοπιστία	100
3.6	Τα κέρδη που θα επιφέρει η σημασιολογία.....	101
ΚΕΦΑΛΑΙΟ 4: ΑΝΑΠΤΥΞΗ ΤΟΥ ΟΝΤΟΛΟΓΙΚΟΥ PORTAL		103
4.1	Σκοπός εφαρμογής.....	103
4.2	Τεχνολογίες – πλατφόρμες που επιλέχθηκαν.....	103
4.3	Σχεδιασμός του Portal – βασικά τμήματα.....	114
4.4	Υλοποίηση Portal – Παραμετροποίηση Joomla!	115
4.5	Παρουσίαση υλοποίησης – οθόνες από τα κύρια τμήματα του portal	121
ΣΥΜΠΕΡΑΣΜΑΤΑ.....		127
ΑΝΑΦΟΡΕΣ – REFERENCES		131
ΠΗΓΕΣ ΑΠΟ INTERNET.....		134
ΕΡΓΑΣΙΕΣ -ΔΗΜΟΣΙΕΥΣΕΙΣ.....		137

ΕΙΣΑΓΩΓΗ - ΣΤΟΧΟΣ ΕΡΓΑΣΙΑΣ - ΠΕΡΙΓΡΑΦΗ ΕΡΓΑΣΙΑΣ

Τα τελευταία χρόνια μια μεγάλη ερευνητική κοινότητα έχει ως αντικείμενο τις Οντολογίες και τον τρόπο με το οποίο αυτές μπορούν να αξιοποιηθούν στην επιστήμη των υπολογιστών. Τα αντικείμενα και οι εφαρμογές της έρευνας αυτής ποικίλουν από τον ορισμό των ίδιων των οντολογιών, των μοντέλων αναπαράστασης τους, της μηχανικής των οντολογιών μέχρι και την ανάπτυξη αυτόνομων οντολογικών μοντέλων για συγκεκριμένες εφαρμογές, π.χ. στις οντολογίες στο ηλεκτρονικό εμπόριο, στην οντολογία ηλεκτρονικής μάθησης (e-learning ontology), στις ψηφιακές βιβλιοθήκες, στην ιατρική – βιολογία – βιοϊατρική κ.λπ. και τέλος έρευνα στην σύγκριση και την συγχώνευση οντολογιών.

Γιατί όμως χρειαζόμαστε τις οντολογίες; Σύμφωνα με τη μελέτη μας καθώς έχουμε αύξηση της ψηφιακής πληροφορίας, γίνεται όλο και δυσκολότερη η εύρεση και η σύνοψη της πληροφορίας. Ένα άλλο πρόβλημα είναι ότι μεγαλώνει το κενό μεταξύ της σημασίας της πληροφορίας και της καταχωρημένης πληροφορίας και έτσι γίνεται εξαιρετικά δύσκολη η ενοποίηση της πληροφορίας και η διαμοίραση και διαχείριση της γνώσης. Αντίθετα με τις οντολογίες είναι δυνατόν να αναπτύξουμε μία κοινή κατανόηση της δομής της πληροφορίας τόσο μεταξύ ανθρώπων όσο και μεταξύ μηχανών. Επιπλέον μπορούμε να επιτύχουμε την επαναχρησιμοποίηση της γνώσης του πεδίου εφαρμογής ειδικά με την εισαγωγή προτύπων τα οποία θα μας επιτρέψουν την διαλειτουργικότητα.

Ειδικότερα, με την έλευση του Παγκόσμιου Ιστού ο οποίος έχει αλλάξει σημαντικά και ραγδαία τον τρόπο που οργανώνεται και διαμοιράζεται η πληροφορία ήδη σχεδιάζεται η επόμενη γενιά του Παγκόσμιου Ιστού, ο Σημασιολογικός Ιστός, ο οποίος επιδιώκει να καταστήσει την πληροφορία πιο

κατανοητή για τους υπολογιστές με την εισαγωγή μιας αυστηρότερης δομής βασισμένης στις οντολογίες.

Τι όμως εννοούμε με τον όρο «Οντολογία»; Τον όρο αρχικά τον βρίσκουμε στη φιλοσοφία, με πολλές εφαρμογές, στην προσπάθεια να απαντηθούν κάποια φιλοσοφικά ερωτήματα σχετικά με την ουσία και την ύπαρξη των πραγμάτων και εννοιών. Στην επιστήμη των υπολογιστών μια οντολογία είναι μια αυστηρή, τυπική προδιαγραφή μιας κοινά αποδεκτής εννοιολογικής μορφοποίησης. Με το όρο «εννοιολογική μορφοποίηση» εννοούμε μια αφηρημένη απλοποιημένη αντίληψη του κομματιού εκείνου του πεδίου εφαρμογής που επιθυμούμε να αναπαραστήσουμε. Είναι μια αυστηρά τυπική αναπαράσταση γιατί πρέπει να είναι κατανοητή από τη μηχανή και είναι κοινά αποδεκτή γιατί καταγράφει κοινά αποδεκτή γνώση. Γνώση εδώ είναι ένα σύνολο από δεδομένα με σημασιολογικό περιεχόμενο γι' αυτό άλλωστε και οι οντολογίες χρησιμοποιούνται για την αναπαράσταση αυτής. Με άλλα λόγια πρόκειται για την τυπική περιγραφή μίας κοινής αντίληψης για τον κόσμο.

Ο όρος αυτός είναι αποδεκτός από ερευνητικές κοινότητες της Τεχνητής Νοημοσύνης δηλαδή της Μηχανικής Γνώσης, Επεξεργασίας Φυσικής Γλώσσας και Αναπαράστασης Γνώσης ενώ αργότερα στα πεδία της Ευφυούς Ενοποίησης της Πληροφορίας της Διαχείρισης Γνώσης κ.α. Οι οντολογίες αναπτύχθηκαν με σκοπό την διευκόλυνση της γνώσης. Μια οντολογία είναι ένα σύνολο από έννοιες, ρόλους, αξιώματα και οντότητες. Οι έννοιες περιγράφουν αντικείμενα του πεδίου εφαρμογής, οι ρόλοι περιγράφουν σχέσεις μεταξύ των εννοιών, τα αξιώματα θέτουν περιορισμούς στις σχέσεις αυτές και τέλος οι οντότητες είναι τα πραγματικά αντικείμενα του πεδίου εφαρμογής που καταγράφονται με τις έννοιες.

Συνεπώς τα χαρακτηριστικά των οντολογιών που πρέπει να προδιαγραφούν είναι ο ορισμός του λεξιλογίου από έννοιες που περιγράφουν τον κόσμο μας, ο ορισμός της δομής των εννοιών (σχέσεις, ιεραρχίες, κ.λπ.), ορισμός ιδιοτήτων των εννοιών και τέλος τα αξιώματα-κανόνες.

Η μηχανική των οντολογιών είναι το πεδίο της επιστήμης το οποίο μελετά μεθόδους και μεθοδολογίες ανάπτυξης οντολογιών. Μελετά τις διαδικασίες ανάπτυξης, του κύκλου ζωής των οντολογιών καθώς επίσης και εργαλεία τα οποία να υποστηρίζουν τις μεθοδολογίες και τις μεθόδους ανάπτυξης οντολογιών.

Προκειμένου να αναπαραστήσουμε τις οντολογίες χρησιμοποιούμε μια γλώσσα μοντελοποίησης οντολογιών. Υπάρχουν πολλές τέτοιες γλώσσες που έχουν αναπτυχθεί από συγκεκριμένες ερευνητικές ομάδες είτε έχουν παγιωθεί ως πρότυπα έκφρασης όπως η KIF, η OWL, η DAML+OIL, η RDFS η Cyc, η IDEF5 και η Ontolingua.

Ακόμα περισσότερα είναι τα παραδείγματα οντολογιών και των συλλογών από οντολογίες (βιβλιοθήκες οντολογιών – ontology libraries) οι οποίες καλύπτουν ένα ευρύ φάσμα εφαρμογών.

Ο στόχος της εργασίας αυτής είναι διπλός. Ο πρώτος στόχος είναι η μελέτη και η καταγραφή του πεδίου της έρευνας των οντολογιών, ένα πεδίο αρκετά διαδεδομένο και πολύ μεγάλο σε εύρος. Ο δεύτερος στόχος είναι η ανάλυση, ο σχεδιασμός και η υλοποίηση ενός δυναμικού ιστότοπου – portal, σύμφωνα με τα σύγχρονα πρότυπα και εργαλεία ανάπτυξης διαδικτυακών εφαρμογών προκειμένου να γίνει η ενσωμάτωση και παρουσίαση της καταγεγραμμένης πληροφορίας σχετικά με το πεδίο έρευνας των οντολογιών ως ένα κεντρικό σημείο διαδικτυακής αναφοράς.

Η ανάπτυξη αυτή έγινε με την χρήση του διαδικτυακού λογισμικού Joomla! και αναρτήθηκε σε συγκεκριμένη διαδικτυακή διεύθυνση που καταχωρήσαμε για τους σκοπούς της εργασίας αυτής.

Joomla! είναι ένα ελεύθερο ανοικτό σύστημα διαχείρισης περιεχομένου για το περιεχόμενο στο Παγκόσμιο ιστό. Το σύστημα περιλαμβάνει τα χαρακτηριστικά γνωρίσματα όπως η εναποθήκευση σελίδων για να βελτιώσει την απόδοση, τις εκτυπώσιμες εκδόσεις των σελίδων, τις λάμπσεις ειδήσεων,

blogs, τις ψηφοφορίες, την έρευνα ιστοχώρου, και τη γλωσσική διεθνοποίηση¹. Το Joomla! είναι διάδοχος του Mambo και χρησιμοποιεί τη βάση δεδομένων MySQL εξ ορισμού.

Στην παρούσα εργασία ασχοληθήκαμε αρχικά με τη μελέτη των οντολογιών γενικά και κατόπιν με συγκεκριμένα οντολογικά πεδία έρευνας. Τέλος προχωρήσαμε στην ανάλυση, σχεδίαση και υλοποίηση του διαδικτυακού portal σχετικά με τις οντολογίες το οποίο και αναρτήσαμε στη διαδικτυακή διεύθυνση <http://www.ontologyportal.net>.

Η δομή της εργασίας μας έχει ως εξής:

Στο πρώτο κεφάλαιο, αναφερόμαστε στην χρήση των οντολογιών στην πληροφορική όπως και στις ανάγκες δημιουργίας, τα κύρια συστατικά και τις κατηγορίες τους. Αναλύουμε τις τεχνολογίες δημιουργίας οντολογιών, την μηχανική, την ευθυγράμμιση, την συγχώνευση και παραθέτουμε μερικές εφαρμογές τους.

Στο δεύτερο κεφάλαιο, αναλύουμε τις βασικότερες γλώσσες αναπαράστασης οντολογιών (XML, RDF, RDF-S, OWL, OIL, DAML-OIL, XOL, SHOE). Που διακρίνονται, που διαφέρουν και πως αξιολογούνται.

Στο τρίτο κεφάλαιο, αναπτύσσουμε και συγκρίνουμε τις έννοιες του σημασιολογικού και παγκόσμιου ιστού, αναλύουμε την χρησιμότητα πρακτόρων λογισμικού στον ιστό σήμερα. Παρουσιάζουμε την αρχιτεκτονική σημασιολογικού ιστού και τα κέρδη που επιφέρει η σημασιολογία.

Τέλος στο τέταρτο κεφάλαιο, περιγράφουμε την διαδικασία που ακολουθήσαμε για την δημιουργία του οντολογικού portal, όπου ήταν και ο στόχος της πτυχιακής μας διατριβής.

¹ Είναι η διαδικασία εφαρμογής λογισμικού που μπορεί να προσαρμοστεί στις διάφορες γλώσσες με την προσθήκη συγκεκριμένων συστατικών για την μετάφραση κειμένων.

ΚΕΦΑΛΑΙΟ 1: ΟΝΤΟΛΟΓΙΕΣ

Οι οντολογίες αποτελούν το δομικό στοιχείο του Σημασιολογικού Ιστού (ο οποίος περιγράφεται στο τρίτο κεφάλαιο). Ωστόσο, ο όρος «οντολογία» έχει γίνει δεκτός από αρκετές ερευνητικές κοινότητες της Τεχνητής Νοημοσύνης² (Artificial Intelligence). Αυτές αρχικά ήταν οι κοινότητες της Μηχανικής Γνώσης, Επεξεργασίας Φυσικής Γλώσσας, και Αναπαράστασης Γνώσης.

Στα τέλη του 1990 η έννοια των οντολογιών διαδόθηκε σε πεδία όπως Ευφυής Ενοποίηση Πληροφορίας, Ανάκτηση Πληροφορίας στο Διαδίκτυο, Διαχείριση Γνώσης στο Ηλεκτρονικό Εμπόριο και σε πολλούς ακόμα τομείς. Οι οντολογίες αναπτύχθηκαν με σκοπό την διευκόλυνση της κοινής χρήσης και επαναχρησιμοποίησης της γνώσης (Knowledge sharing and reuse).

1.1 Ορισμός Οντολογιών

Η οντολογία είναι μια έννοια που χρησιμοποιήθηκε για πρώτη φορά από τους αρχαίους Έλληνες φιλοσόφους στην προσπάθειά τους να απαντήσουν σε κάποια φιλοσοφικά ερωτήματα σχετικά με την ουσία και την ύπαρξη κάποιων πραγμάτων και εννοιών. Ωστόσο ο όρος οντολογία συναντάται και σε άλλα επιστημονικά πεδία όπως στους τομείς της τεχνητής νοημοσύνης, της γνώσης κτλ. Με τον όρο οντολογία εννοούμε την ακριβή περιγραφή πραγμάτων και εννοιών καθώς και των σχέσεων που υπάρχουν ανάμεσα τους.

² Η νοημοσύνη των μηχανών και ο κλάδος της πληροφορικής που στοχεύει να την δημιουργήσει

Ο πιο γνωστός ορισμός για την οντολογία, στην επιστήμη των υπολογιστών, πάνω στον οποίο στηρίχτηκαν και άλλοι ορισμοί, δόθηκε από τον Gruber (Gruber, 1993) και είναι ο ακόλουθος:

- *An ontology is a formal explicit specification of a shared conceptualization.*

Επεξηγηματικά: Μια οντολογία είναι μια τυπική³ (formal), κατηγορηματική⁴ (explicit) προδιαγραφή μιας διαμοιρασμένης⁵ (shared) εννοιολογικής αναπαράστασης⁶ (conceptualization) – Gruber

Παρακάτω αναφέρονται μερικοί ακόμα ενδεικτικοί ορισμοί που έχουν δοθεί για να περιγράψουν την έννοια της οντολογίας:

- *A logical theory which gives an explicit, partial account of a conceptualization.*
- *A set of logical axioms designed to account for the intended meaning of a vocabulary.*
- *An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base.*

Ο Guarino κατονομάζει διάφορα ερευνητικά πεδία της επιστήμης των υπολογιστών που έχουν «αγκαλιάσει» τις οντολογίες (Guarino, 1998) περιλαμβάνοντας σε αυτά τη μηχανική γνώσης, την αναπαράσταση γνώσης, τη

³ Ο όρος ‘τυπική’ (formal) αναφέρεται στο ότι η οντολογία πρέπει να είναι μηχανικά αναγνώσιμη.

⁴ Ο όρος ‘κατηγορηματική’ (explicit) σημαίνει ότι το είδος των εννοιών που χρησιμοποιούνται, και οι περιορισμοί που αφορούν την χρήση αυτών των εννοιών είναι προσδιορισμένα με σαφήνεια.

⁵ Ο όρος ‘διαμοιρασμένη’ (shared) αναφέρεται στο ότι η οντολογία πρέπει να αποτυπώνει γνώση κοινής αποδοχής στα πλαίσια μιας κοινότητας.

⁶ Ο όρος ‘εννοιολογική αναπαράσταση’ (conceptualization) αναφέρεται σε ένα αφηρημένο μοντέλο φαινομένων του κόσμου στο οποίο έχουν προσδιοριστεί οι έννοιες που σχετίζονται με τα φαινόμενα αυτά.

γλωσσολογική τεχνολογία, το σχεδιασμό βάσεων δεδομένων, την ανάκτηση και εξαγωγή πληροφορίας, και τη διαχείριση και οργάνωση της γνώσης. Σήμερα, ιδιαίτερη σημασία δίνεται στην χρήση των οντολογιών Σηματολογικού Ιστού, κάτι που μπορεί να θεωρηθεί και σα διαχείριση γνώσης σε παγκόσμια κλίμακα, συνδυάζοντας ερευνητικά πεδία, όπως αυτά που αναφέρθηκαν.

Μία οντολογία είναι ένα σύνολο από έννοιες (concepts), ρόλους (roles), αξιώματα (axioms), και οντότητες (entities ή individuals). Οι έννοιες περιγράφουν σύνολα από αντικείμενα του σύμπαντος (πραγματικού κόσμου), οι ρόλοι περιγράφουν τις σχέσεις που υπάρχουν μεταξύ των εννοιών, τα αξιώματα περιορίζουν τις σχέσεις αυτές με διάφορους περιορισμούς, και οι οντότητες είναι τα πραγματικά αντικείμενα του σύμπαντος που καταγράφονται με τις έννοιες.

1.2 Ανάγκες δημιουργίας οντολογιών

Οι λόγοι, για τους οποίους είναι χρήσιμη η ανάπτυξη μιας οντολογίας, είναι αρκετοί και ιδιαίτερα σημαντικοί:

- Διαμοιρασμός της κοινής αντίληψης της δομής της πληροφορίας ανάμεσα σε ανθρώπους ή υπολογιστικούς πράκτορες: Είναι ίσως ο πιο σημαντικός στόχος στην ανάπτυξη οποιασδήποτε οντολογίας. Για παράδειγμα, αν υποθέσουμε ότι διάφοροι ιστότοποι περιέχουν ιατρική πληροφορία, διαμοιράζονται και δημοσιεύουν υλικό που βασίζεται πάνω στην ίδια οντολογία, τότε οι πράκτορες μπορούν να ανακτήσουν και να συνθέσουν πληροφορία από όλες αυτές τις διαφορετικές πηγές.
- Επαναχρησιμοποίηση της πληροφορίας του τομέα: Μια από τις κινητήριες δυνάμεις που βρίσκονται πίσω από τις τρέχουσες εξελίξεις στην έρευνα για τις οντολογίες. Εάν μια ομάδα ερευνητών αναπτύσσει μια οντολογία, κάποιες

άλλες ομάδες μπορούν απλά να την επαναχρησιμοποιήσουν στο δικό τους αντικείμενο. Επιπρόσθετα, εάν χρειαστεί να αναπτυχθεί μια ευρύτερη οντολογία, αυτό μπορεί να γίνει με τη συνένωση ήδη υπαρχόντων επιμέρους οντολογιών.

- Δημιουργία ρητών υποθέσεων για κάθε γνωστικό τομέα, έτσι ώστε αν χρειαστεί να αλλάξει κάποια από τις υποθέσεις, εφόσον αλλάξει η γνώση που έχουμε για τον τομέα αυτό, να μπορεί να υλοποιηθεί πολύ εύκολα η αλλαγή αυτή. Επιπλέον, βοηθούν το νέο χρήστη με περιορισμένη γνώση, ώστε να μπορέσει να αντιληφθεί ευκολότερα και με σαφήνεια την έννοια και τη σημασία των όρων του τομέα.
- Διαχωρισμός της γνώσης του κάθε τομέα από τη λειτουργική γνώση: Μπορεί, για παράδειγμα, να γίνει η περιγραφή μιας εργασίας για τη συναρμολόγηση ενός προϊόντος από τα συστατικά του σύμφωνα με κάποιες προδιαγραφές, και να υλοποιηθεί και ένα πρόγραμμα το οποίο κάνει τη διαδικασία συναρμολόγησης και ρύθμισης ανεξάρτητη από τα εκάστοτε συστατικά.
- Ανάλυση της γνώσης του κάθε τομέα: Αυτό μπορεί να γίνει πολύ εύκολα σε περίπτωση που είναι διαθέσιμες ξεκάθαρες προδιαγραφές για την ορολογία που χρησιμοποιείται.

Επιπλέον προβλήματα σύμφωνα με τον Tim Berners-Lee που οδήγησαν στη δημιουργία των οντολογιών είναι (Berners-Lee et al., 2001).

- Το μεγαλύτερο ποσοστό των πληροφοριών των οργανισμών είναι κωδικοποιημένο σε μορφή κειμένου,
- Το μη δομημένο ή ημι-δομημένο κείμενο το οποίο είναι δύσκολο να ανακτηθεί, να ερμηνευτεί και να αναλυθεί,
- Οι ανολοκλήρωτες περιγραφές,
- Οι άσκοπα επαναλαμβανόμενες τιμές,
- Τα ασύμβατα έγγραφα,

- Οι διαφορετικές ορολογίες / γλώσσες / συντομογραφίες.

Οι οντολογίες σχεδιάζονται έτσι ώστε να είναι δυνατός ο διαμοιρασμός γνώσης μεταξύ των πρακτόρων. Αναλυτικότερα η αύξηση της ψηφιακής πληροφορίας, η ενοποίηση, η δυσκολότερη πρόσβαση, εύρεση και σύνοψη αυτής καθώς και το μεγάλο κενό μεταξύ της σημασίας της πληροφορίας και της καταχωρημένης πληροφορίας σε συνδυασμό με τη δύσκολη διαμοίραση και διαχείριση της γνώσης (δε μιλούν όλοι οι άνθρωποι την ίδια γλώσσα) είχαν σαν αποτέλεσμα την είσοδο της οντολογίας στο σημασιολογικό ιστό.

1.3 Βασικά πλεονεκτήματα/δυνατότητες των οντολογιών

Οι δυνατότητες που προσφέρονται με την δημιουργία οντολογιών στους διαφόρους τομείς της πληροφορίας είναι οι ακόλουθες:

- Η αυξημένη ακρίβεια αναζήτησης μέσω της άρσης προβλημάτων όπως διαφορούμενες έννοιες, ανεπαρκής πληροφορία από τον χρήστη κ.λπ.,
- Η δυνατότητα σύνδεσης των σχετικών πληροφοριών,
- Η μείωση του θορύβου της πληροφορίας,
- Η διευκόλυνση των διαδικασιών επίλυσης προβλημάτων ή/και της λήψης αποφάσεων,
- Η συμμετοχή και η επαναχρησιμοποίηση της γνώσης,
- Η αυτοματοποίηση,
- Οι προσαρμοζόμενες υπηρεσίες,
- Η ανάπτυξη αποδοτικότερων εργαλείων εξαγωγής/ διαχείρισης γνώσης,
- Η εκμετάλλευση πολυμεσικών δεδομένων.

1.4 Βασικά μειονεκτήματα / προβλήματα των οντολογιών

Λόγω του ότι οι μηχανές αναζήτησης του σημασιολογικού ιστού είναι ακόμη αναπτυσσόμενες και μη ολοκληρωμένες, παρουσιάζονται προβλήματα σε σχέση με τις δυνατότητες που προσφέρουν οι οντολογίες. Τα πιο βασικά προβλήματα που μπορούν να παρουσιαστούν είναι τα ακόλουθα:

- Υπάρχει δυσκολία στην παραγωγή μεταδεδομένων⁷ και οντολογιών από τους χρήστες,
- Ο ανακριβής ορισμός των οντολογιών οδηγεί πολύ συχνά σε προβλήματα ολοκλήρωσης,
- Ο μη ανοικτός και ευέλικτος ορισμός των οντολογιών οδηγεί σε προβλήματα ευχρηστίας,
- Η μη ακριβής κατανόηση του επιχειρηματικού μοντέλου και η λανθασμένη ολοκλήρωση των υπηρεσιών που βασίζονται σε οντολογίες οδηγούν σε αδυναμία εκμετάλλευσης των δυνατοτήτων τους,
- Συχνά δημιουργείτε ανομοιογένεια των υπαρχόντων σημασιολογικών μοντέλων (ontology interoperability),
- Προς το παρόν είναι μη εφικτή η εφαρμογή ενός συνολικού οντολογικού μοντέλου για όλες τις εφαρμογές του ηλεκτρονικού εμπορίου (δηλαδή e-Business, e-Commerce, e-CRM κτλ).

⁷ Τα μεταδεδομένα είναι δομημένα και κωδικοποιημένα δεδομένα τα οποία περιγράφουν χαρακτηριστικά πληροφοριακών οντοτήτων, αποσκοπώντας στην ταύτιση, αναγνώριση, ανακάλυψη, αξιολόγηση και διαχείριση των οντοτήτων που περιγράφονται.

1.5 Τα κύρια συστατικά των οντολογιών

Η γνώση στις οντολογίες, ανεξαρτήτως της τεχνολογίας η οποία τις περιγράφει, τυποποιείται χρησιμοποιώντας πέντε κατηγορίες συστατικών: τις κλάσεις (classes), τα στιγμιότυπα (instances), τις σχέσεις (relations), τις συναρτήσεις (functions) και τέλος τα αξιώματα (axioms). Παρακάτω θα γίνει αναφορά σε αυτά τα συστατικά, για να ξεκαθαριστεί ακόμη περισσότερο η μορφή μία οντολογίας.

1.5.1 Κλάσεις (Classes ή Concepts)

Οι κλάσεις είναι σύνολα από στιγμιότυπα και συνήθως είναι οργανωμένες σε μια ιεραρχία, η οποία είναι γνωστή και ως ταξινόμια (taxonomy). Οι κλάσεις αναπαριστούν έννοιες, είτε αφηρημένες ή συγκεκριμένες. Μπορούν να εμπεριέχουν άτομα, άλλες κλάσεις ή και συνδυασμό των δύο. Παραδείγματος χάριν, για να αναπαραστήσουμε μια ταξινόμια από τους φοιτητές ενός πανεπιστημίου (προπτυχιακοί, μεταπτυχιακοί, υποψήφιοι διδάκτορες κτλ). Σε αυτήν την ταξινόμια η κλάση «Προπτυχιακός Φοιτητής» είναι υποκλάση της κλάσης «Φοιτητής».

Μερικά παραδείγματα κλάσεων είναι:

- αριθμός, αναπαριστά όλους τους αριθμούς,
- αυτοκίνητο, αναπαριστά όλα τα αυτοκίνητα,
- αετός, αναπαριστά όλους τους αετούς,
- ψάρι, αναπαριστά όλα τα ψάρια,
- κλάση, αναπαριστά όλες τις κλάσεις.

Μία κλάση μπορεί να εμπεριέχει άλλες κλάσεις, μπορεί να ανήκει στον εαυτό της ή μπορεί να είναι μία ολότητα - δηλαδή να περιέχει τα πάντα, κτλ. Τέτοιου είδους διαφορές εντοπίζονται στα διάφορα ήδη οντολογιών. Πολλές

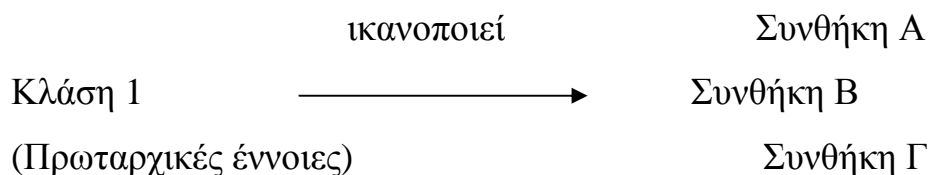
φορές χρειάζεται να τεθούν περιορισμοί στις παραπάνω περιπτώσεις, για να αποφευχθεί η δημιουργία ορισμένων ευρέως γνωστών παραδόξων, όπως για παράδειγμα το να δηλωθεί σε μία ιεραρχία που περιγράφει ανθρώπους, ότι ένας συγκεκριμένος άνθρωπος είναι ταυτόχρονα και άντρας και γυναίκα.

Οι κλάσεις αναπαριστώνται με τις έννοιες. Οι έννοιες χρησιμοποιούνται με την ευρεία σημασία τους. Μια έννοια μπορεί να είναι οτιδήποτε για κάτι που λέγεται και γι' αυτό το λόγο θα μπορούσε επίσης να είναι η περιγραφή μιας εργασίας, μιας λειτουργίας, μιας ενέργειας, μιας ιδέας, μιας κρίσης κλπ. Οι έννοιες είναι δυνατόν να διαιρεθούν σε δύο κατηγορίες:

α) τις *πρωταρχικές έννοιες* (primitive concepts), οι οποίες περιγράφονται μόνο από αναγκαίες συνθήκες (necessary conditions). Αυτό σημαίνει πως αν κάποιο αντικείμενο είναι μέλος μιας κλάσης, τότε είναι αναγκαίο να ικανοποιεί τις αντίστοιχες συνθήκες.

β) τις *έννοιες εξ' ορισμού* (defined concepts), οι οποίες περιγράφονται από ικανές και αναγκαίες συνθήκες (necessary and sufficient conditions). Αυτό σημαίνει ότι αν κάποιο αντικείμενο είναι μέλος μιας κλάσης, τότε είναι αναγκαίο να ικανοποιεί τις αντίστοιχες συνθήκες και αν ένα αντικείμενο ικανοποιεί τις συνθήκες τότε πρέπει να ανήκει στην αντίστοιχη κλάση.

ΑΝΑΓΚΑΙΕΣ ΣΥΝΘΗΚΕΣ



ΙΚΑΝΕΣ ΚΑΙ ΑΝΑΓΚΑΙΕΣ ΣΥΝΘΗΚΕΣ



Εικόνα 1.1 Συνθήκες

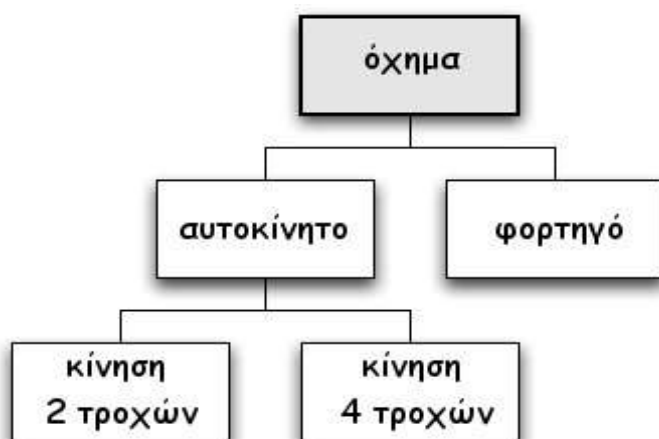
Οι κλάσεις μίας οντολογίας μπορεί να είναι υποθετικές (intensional) ή συμπερασματικές (extensional). Μία υποθετική κλάση είναι μία περιγραφή με ιδιότητες, από την οποία μπορούν να περιγραφούν πολλά στιγμιότυπα. Για παράδειγμα η κλάση άνθρωπος μπορεί να θεωρηθεί ως μία περιγραφή (με ιδιότητες όπως το ονοματεπώνυμο, η ημερομηνία γέννησης, το χρώμα ματιών, κτλ), από την οποία μπορούν να προκύψουν πολλά στιγμιότυπα. Για αυτό το λόγο το σχήμα δημιουργίας μίας σχεσιακής βάσης δεδομένων (εντολές SQL) ονομάζεται υποθετική βάση δεδομένων (intensional database), ενώ οι εγγραφές των δεδομένων ονομάζονται συμπερασματική βάση δεδομένων (extensional database): Από το ίδιο σχήμα, μπορούν να προκύψουν διαφορετικά αποτελέσματα (εγγραφές). Η κλάση άνθρωπος μπορεί όμως να αντιμετωπιστεί και ως συμπερασματική, όπου σε αντιστοιχία με τη μαθηματική θεωρία των συνόλων, αποτελεί το στοιχείο ενός συνόλου, π.χ. του άνθρωπος = {Μαρίνα, Μαρία, Φυλιώ}.

Η διάκριση μεταξύ των δύο ειδών κλάσεων είναι ιδιαίτερα λεπτή. Η υπόθεση (intension) σχετίζεται με κάποια ιδιότητα, ενώ το συμπέρασμα (extension) σχετίζεται με τα αντικείμενα που έχουν αυτή την ιδιότητα εντός ενός προκαθορισμένου πλαισίου (context). Ένα καλό παράδειγμα αποτελεί η ιδιότητα “κόκκινο”. Πολλά μπορούν να ειπωθούν για τη συγκεκριμένη ιδιότητα, για παράδειγμα ότι πρόκειται για ένα φάσμα συχνοτήτων συγκεκριμένου μήκους κύματος, ότι είναι ορατή στο ανθρώπινο μάτι, κτλ. Αυτό είναι ένα

παράδειγμα υπόθεσης. Αντίθετα, μπορεί να γίνει αναφορά σε συγκεκριμένα αντικείμενα τα οποία έχουν τη συγκεκριμένη ιδιότητα (κόκκινο χρώμα), π.χ. μία κόκκινη μπάλα, ένα κόκκινο τριαντάφυλλο, κτλ. Αυτό είναι ένα παράδειγμα συμπεράσματος.

Αν και οι συμπερασματικές κλάσεις παρουσιάζουν καλύτερη συμπεριφορά και είναι περισσότερο κατανοητές από μαθηματικής απόψεως, δεν επιτρέπουν τη δημιουργία λεπτών διακρίσεων που συνήθως κρίνονται απαραίτητες για την κατασκευή οντολογιών. Για παράδειγμα, μία οντολογία ίσως χρειαστεί να υλοποιήσει σαφή διαχωρισμό μεταξύ μίας κλάσης όλων των πλασμάτων τα οποία έχουν νεφρά, και μίας κλάσης όλων των πλασμάτων με καρδιά, ακόμη κι αν τυχαίνει οι δύο αυτές κλάσεις να έχουν ακριβώς τα ίδια μέλη.

Ένα επίσης σημαντικό χαρακτηριστικό, είναι ότι μία κλάση μπορεί να εντάσσει (subsume) ή να εντάσσεται σε άλλες κλάσεις. Για παράδειγμα, στην εικόνα 1.2, η κλάση όχημα εντάσσει την κλάση αυτοκίνητο, εφόσον οποιοδήποτε αυτοκίνητο εντάσσεται και στην ευρύτερη κατηγορία των οχημάτων. Η σχέση της ένταξης χρησιμοποιείται ώστε να δημιουργηθεί μία ιεραρχία κλάσεων, με την γενικότερη να βρίσκεται στην κορυφή (π.χ. αυτοκίνητο), και τις ειδικότερες στα κάτω σημεία (π.χ. κίνηση 2 τροχών).



Εικόνα 1.2 Μερικό διάγραμμα ιεραρχίας κλάσεων οντολογίας.

Μία διανομή (partition) είναι ένα σύνολο συσχετιζόμενων κλάσεων και κανόνων, που επιτρέπουν στα αντικείμενα να τοποθετούνται στην κατάλληλη κλάση. Για παράδειγμα, στην εικόνα 1.2, υπάρχει μία διανομή για την κλάση αυτοκίνητο που την χωρίζει σε δύο υποκλάσεις (ένα επίπεδο κατώτερες στην ιεραρχία): η μία αναπαριστά όλα τα αυτοκίνητα που υποστηρίζουν κίνηση δύο τροχών, και η άλλη όλα όσα υποστηρίζουν κίνηση τεσσάρων τροχών. Εάν η διανομή εγγυάται ότι ένα αυτοκίνητο δεν μπορεί να υπάρχει και στις δύο υποκλάσεις, τότε ονομάζεται διαχωρίσιμη (disjoint). Εάν οι κανόνες της διανομής εξασφαλίζουν ότι κάθε διακριτό αντικείμενο της υπερκλάσης (της ανώτερης κατά ένα επίπεδο στην ιεραρχία) είναι στιγμιότυπο τουλάχιστον μίας εκ των υποκλάσεων, η διανομή ονομάζεται εξαντλητική (exhaustive).

1.5.2 Άτομα (Individuals) ή Στιγμιότυπα (Instances)

Σύμφωνα με την Wikipedia (Wiki, 2008), τα άτομα ή στιγμιότυπα (instances) είναι τα βασικά συστατικά “χαμηλού επιπέδου” μίας οντολογίας. Αποτελούν διακριτά αντικείμενα-μέλη των κλάσεων και εκφράζουν συγκεκριμένα στοιχεία. Για παράδειγμα οι “Άνδρας” και “Ελλάδα” αποτελούν στιγμιότυπα των κλάσεων “Άνθρωπος” και “Χώρα” αντίστοιχα.

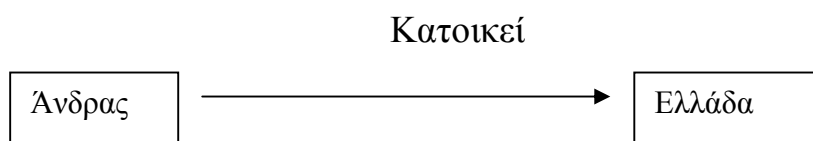
Μία οντολογία δεν είναι απαραίτητο να περιέχει άτομα, άλλα ένας από τους βασικούς σκοπούς της είναι να παρέχει ένα τρόπο ταξινόμησης των ατόμων, ακόμη κι αν αυτά δεν αποτελούν τμήμα της.

1.5.3 Σχέσεις (Relations)

Οι σχέσεις αναπαριστούν δυαδικές συσχετίσεις ανάμεσα στα στιγμιότυπα των κλάσεων. Οι σχέσεις μπορεί να έχουν ένα πεδίο ορισμού (domain) και ένα πεδίο τιμών (range) ενώ συνδέουν στιγμιότυπα από το πεδίο ορισμού με στιγμιότυπα από το πεδίο τιμών. Το πεδίο τιμών μπορεί να είναι είτε μια κλάση

ή ένας τύπος δεδομένων. Ανάλογα με το είδος του πεδίου τιμών οι σχέσεις διακρίνονται στις ακόλουθες κατηγορίες:

- **Relations:** Στην κατηγορία αυτή ανήκουν οι σχέσεις που συνδέουν ένα στιγμιότυπο με ένα άλλο στιγμιότυπο.
- **Attributes:** Στην κατηγορία αυτή ανήκουν οι σχέσεις που συνδέουν ένα στιγμιότυπο με έναν κυριολεκτικό τύπο δεδομένων (π.χ. αριθμητικό, συμβολοσειρά, κτλ).



Άνδρας: στιγμιότυπο της κλάσης Άνθρωπος

Ελλάδα: στιγμιότυπο της κλάσης Χώρα

Εικόνα 1.3 Διαφορετικά είδη σχέσεων

Στην εικόνα 1.3, για παράδειγμα, η κλάση “Άνθρωπος” είναι το πεδίο ορισμού της σχέσης “Κατοικεί”, ενώ η κλάση “Χώρα” αποτελεί το πεδίο τιμών της ίδιας σχέσης. Όπως και οι κλάσεις, έτσι και οι σχέσεις είναι δυνατόν να οργανωθούν σε ιεραρχίες. Για παράδειγμα η σχέση “έχει μητέρα” αποτελεί υπο-σχέση (sub-property) της σχέσης “έχει πατέρα”.

Οι σχέσεις εκφράζουν ένα είδος αλληλεπίδρασης μεταξύ των εννοιών ενός πεδίου (π.χ. subclass-of, is-a).

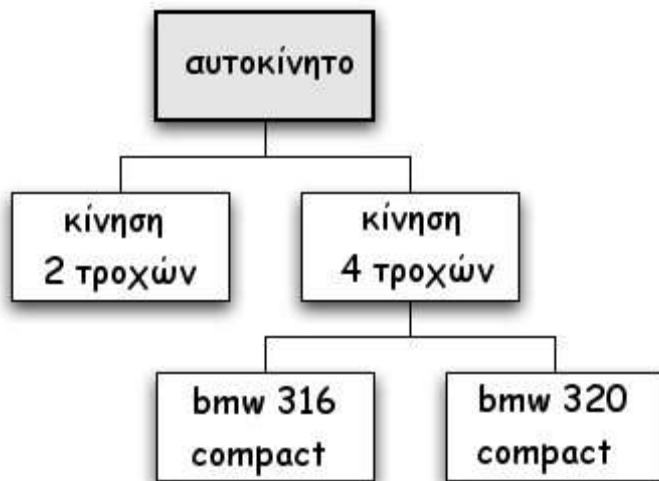
Σύμφωνα με τη Wikipedia (Wiki, 2008), οι ιδιότητες είναι χρήσιμες για την περιγραφή των σχέσεων (relationships ή relations) μεταξύ των αντικειμένων μίας οντολογίας. Μία σχέση είναι μία ιδιότητα η τιμή της οποίας είναι ένα άλλο αντικείμενο της οντολογίας. Στην εικόνα 1.4, μία πιθανή ιδιότητα του αντικειμένου bmw 316 compact είναι η:

διάδοχος: bmw 320 compact.

Αυτή η πληροφορία γνωστοποιεί ότι το bmw 320 compact είναι το μοντέλο που διαδέχθηκε το μοντέλο bmw 316 compact.

Οι οντολογίες αντλούν μεγάλη δύναμη από την ικανότητα τους για περιγραφή τέτοιου είδους σχέσεων. Το σύνολο όλων των σχέσεων μίας οντολογίας περιγράφουν τη σημασιολογία της.

Η σημαντικότερη σχέση μίας οντολογίας είναι η ένταξη (γνωστή και ως *είναι-ένα*, *είναι-υποτύπος*, *είναι-υποκλάση*). Αυτή η σχέση καθορίζει ποια αντικείμενα είναι μέλη σε μία ιεραρχία κλάσεων.



Εικόνα 1.4 Ιεραρχία κλάσεων οντολογίας αυτοκινήτου.

Παραδείγματος χάριν, στην εικόνα 1.4 φαίνεται ότι το bmw 316 compact είναι-ένα όχημα με κίνηση τεσσάρων τροχών, και με τη σειρά του ένα όχημα με κίνηση τεσσάρων τροχών είναι- ένα αυτοκίνητο, το οποίο με τη σειρά του είναι- ένα όχημα, κτλ.

Η χρήση της σχέσης *είναι-ένα* δημιουργεί μία ιεραρχική ταξινόμηση (hierarchical taxonomy) μία δενδρική δομή που απεικονίζει με σαφήνεια πως σχετίζονται μεταξύ τους τα αντικείμενα . Σε μία τέτοιου είδους ιεραρχία, κάθε αντικείμενο είναι το “παιδί” μίας “γονικής” κλάσης (εκτός φυσικά από τη ρίζα ή τις ρίζες του δέντρου).

Μία άλλη σχέση είναι η σχέση μέρους (meronymy (WikiMer, 2008)) που γράφεται ως *είναι-μέρος*. Η σχέση μέρους αναπαριστά πώς αντικείμενα συνδυάζονται μεταξύ τους ώστε να διαμορφώσουν σύνθετα (composite) αντικείμενα. Λόγου χάριν, αν επεκταθεί το παράδειγμα της εικόνας 1.4 ώστε η οντολογία να περιέχει αντικείμενα όπως το αντικλεπτικό σύστημα bmw, είναι γνωστό ότι “το αντικλεπτικό σύστημα bmw *είναι-μέρος* του bmw 316 compact”.

Σε περίπτωση που εισαχθούν σχέσεις του τύπου *είναι-μέρος* σε μία οντολογία, παρατηρείται ότι η απλή και κομψή δενδρική δομή μετατρέπεται σε σύνθετη και επομένως είναι δυσκολότερο να ερμηνευτεί. Η εξήγηση του λόγου για τον οποίο εμφανίζεται αυτό το πρόβλημα δεν είναι δύσκολη. Μία οντότητα που *είναι-μέρος* μίας οντότητας, μπορεί να *είναι-μέρος* και κάποιας άλλης γενικότερης οντότητας, κτλ.

Πέρα από τις πρότυπες σχέσεις του τύπου *είναι-ένα* και *είναι-μέρος*, νέες σχέσεις μπορούν να προκύψουν στα πλαίσια επίλυσης ενός προβλήματος. Για παράδειγμα, στον τομέα των αυτοκινήτων θα μπορούσε να καθοριστεί η σχέση *προέρχεται-από*, η οποία περιγράφει τη χώρα προέλευσης ενός αυτοκινήτου. Έτσι, για το παράδειγμα της εικόνας 1.4 μπορεί να αναφερθεί ότι το bmw 316 compact *προέρχεται-από* το Βερολίνο, το οποίο *είναι-μία* από τις πόλεις της Γερμανίας. Το λογισμικό το οποίο θα χρησιμοποιούσε αυτή την οντολογία θα ήταν ικανό να απαντήσει σε ερωτήσεις του τύπου “ποια αυτοκίνητα *προέρχονται-από* τη Γερμανία;”

1.5.4 Συναρτήσεις (Functions)

Οι συναρτήσεις εκπροσωπούν μια ειδική περίπτωση σχέσης. Για παράδειγμα, η τιμή-μεταχειρισμένου-αυτοκινήτου μπορεί να προσδιορίζεται ως συνάντηση της αρχικής τιμής του καινούριου αυτοκινήτου, του μοντέλου και των χαρακτηριστικών του αυτοκινήτου, καθώς και των χιλιομέτρων που έχει διανύσει.

1.5.5 Αξιώματα (Axioms)

Τα αξιώματα χρησιμοποιούνται προκειμένου να αναπαριστούν προτάσεις που είναι πάντοτε αληθείς. Για παράδειγμα, αν ο Χ είναι δευτεροετής φοιτητής, τότε μπορεί να παρακολουθήσει το επιλεγόμενο μάθημα Μ.

1.6 Κατηγορίες οντολογιών

Μία οντολογία, σύμφωνα με το βαθμό της τυπικότητας (formality) της αναπαράστασής της, μπορεί να είναι:

- Άτυπη (highly informal), εκφρασμένη δηλαδή σε μια φυσική γλώσσα.
- Ημι-άτυπη (semi-informal), για παράδειγμα διατυπωμένη σε ένα περιορισμένο και δομημένο υποσύνολο κάποιας φυσικής γλώσσας.
- Ημι-τυπική (semi-formal), διατυπωμένη σε μια τεχνητή και αυστηρά ορισμένη γλώσσα.
- Αυστηρά τυπική (rigorously formal): ορισμοί όρων με αυστηρή σημασιολογία, θεωρήματα και αποδείξεις ιδιοτήτων.

Οι Gomez-Perez και Benjamins (1999) ταξινομούν τις οντολογίες ως ακολούθως:

- Οντολογίες αναπαράστασης γνώσης (knowledge representation ontologies): παρέχουν οντότητες αναπαράστασης χωρίς να προσδιορίζουν τι συγκεκριμένο αναπαριστούν.
- Γενικές ή κοινές οντολογίες (general / common ontologies): στοχεύουν στο να αποτυπώσουν γενική γνώση γύρω από τον κόσμο, παρέχοντας βασικές έννοιες όπως ο χρόνος, ο χώρος, τα συμβάντα κλπ.

- Οντολογίες ανώτερου επιπέδου (top-level ontologies): παρέχουν γενικές έννοιες κάτω από τις οποίες συσχετίζονται όλοι οι όροι σε ήδη υπάρχουσες οντολογίες.
- Οντολογίες μεταδεδομένων (metadata ontologies): παρέχουν ένα λεξιλόγιο για την περιγραφή του περιεχομένου πληροφορίας, η οποία είναι ηλεκτρονικά διαθέσιμη.
- Οντολογίες πεδίου ορισμού (domain ontologies): αναπαριστούν γνώση γύρω από ένα συγκεκριμένο πεδίο, π.χ. ιατρική κλπ.
- Οντολογίες μεθοδολογίας ή εργασιών (method or task ontologies): παρέχουν όρους που αναφέρονται σε συγκεκριμένες εργασίες, π.χ. διάγνωση κλπ.

1.7 Κριτήρια σχεδιασμού οντολογιών

Οι τυπικές οντολογίες είναι σχεδιασμένες. Όταν επιλέγουμε πώς να απεικονίσουμε κάτι σε μια οντολογία, λαμβάνουμε αποφάσεις σχεδίου. Για να καθοδηγήσουμε και να αξιολογήσουμε τα σχέδιά μας, χρειαζόμαστε αντικειμενικά κριτήρια που δημιουργούνται με σκοπό να προκύψει το αντικείμενο, παρά με βάση τις προγενέστερες έννοιες της φυσικότητας ή της αλήθειας. Εδώ προτείνουμε ένα προκαταρκτικό σύνολο κριτηρίων σχεδιασμού για τις οντολογίες των οποίων ο σκοπός είναι διανομή και λειτουργικότητα γνώσης μεταξύ των προγραμμάτων βασισμένων σε μια κοινή σύλληψη. (Gruber T., 1993)

1. *Σαφήνεια και αντικειμενικότητα (clarity and objectivity)*. Μια οντολογία πρέπει αποτελεσματικά να αποδώσει την προοριζόμενη έννοια των καθορισμένων όρων. Οι ορισμοί πρέπει να είναι αντικειμενικοί. Ενώ το κίνητρο για τον καθορισμό μιας έννοιας μπορεί να προκύψει από τις κοινωνικές καταστάσεις ή τις υπολογιστικές απαιτήσεις, ο καθορισμός πρέπει να είναι

ανεξάρτητος από το κοινωνικό ή υπολογιστικό πλαίσιο. Ο φορμαλισμός είναι μέσο για αυτόν τον σκοπό. Ο ορισμός μπορεί να δηλωθεί στα λογικά αξιώματα. Όπου είναι δυνατόν, ένας πλήρης ορισμός (ένα κατηγορημα που καθορίζεται από απαραίτητους και ικανοποιητικούς όρους) προτιμάται πέρα από έναν μερικό ορισμό (που καθορίζεται από μόνο απαραίτητους ή ικανοποιητικούς όρους). Όλοι οι ορισμοί πρέπει να τεκμηριωθούν με φυσική γλώσσα. Τέλος θα πρέπει να υπάρχουν παραδείγματα που θα βοηθούν τον αναγνώστη να καταλάβει ορισμούς, από τους οποίους θα λείπουν οι απαραίτητες και επαρκείς συνθήκες για την κατανόησή τους.

2. *Συνοχή (Coherence)*. Μια οντολογία πρέπει να είναι συνεκτική, δηλαδή πρέπει να εγκρίνει τα συμπεράσματα που είναι σύμφωνα με τους ορισμούς. Τα οριζόμενα αξιώματα θα πρέπει, τουλάχιστον, να έχουν μια λογική συνέχεια και συνέπεια. Η συνοχή πρέπει επίσης να ενισχύσει τις έννοιες που καθορίζονται ανεπίσημα, όπως αυτές που περιγράφονται στην τεκμηρίωση φυσικής γλώσσας και στα παραδείγματα. Εάν μια πρόταση που μπορεί να προκύψει από τα αξιώματα έρχεται σε αντίθεση με τον ορισμό ή το παράδειγμα που δίνεται ανεπίσημα, τότε η οντολογία είναι ασυνάρτητη.

3. *Επεκτασιμότητα (Extendibility)*. Η αναπαράσταση θα πρέπει να σχεδιαστεί κατά τέτοιο τρόπο, ώστε κάποιος να έχει τη δυνατότητα να επεκτείνει και να εξειδικεύει την οντολογία μονοτονικά. Μια οντολογία πρέπει να έχει σχεδιαστεί με σκοπό την χρήση κοινού λεξιλογίου. Πρέπει να προσφέρει ένα εννοιολογικό ίδρυμα για μια σειρά προσδοκώμενων στόχων, και η απεικόνιση πρέπει να επεξεργαστεί έτσι ώστε κάποιος να μπορεί να επεκτείνει και να ειδικεύσει την οντολογία. Με άλλα λόγια, κάποιος να μπορεί να είναι σε θέση να καθορίσει νέους όρους για πρόσθετες χρήσεις βασισμένες στο υπάρχον λεξιλόγιο, που με τέτοιο τρόπο ώστε να μην απαιτεί αναθεώρηση των υπάρχοντων ορισμών.

4. *Ελάχιστη οντολογική δέσμευση (Minimal ontological commitment)*. Μία οντολογία θα πρέπει να απαιτεί την ελάχιστη οντολογική δέσμευση, η οποία θα είναι ικανή να υποστηρίζει τις επιθυμητές λειτουργίες διαμοιρασμού της γνώσης. Θα πρέπει να έχει κάποιες αξιώσεις σχετικά με το χώρο ο οποίος μοντελοποιείται, επιτρέποντας στα μέρη που δεσμεύονται στην ελευθερία της οντολογίας να εξειδικεύουν και να εγκαθιστούν την οντολογία όπως πραγματικά χρειάζεται.

5. *Ελάχιστο εύρος κωδικοποίησης (Minimal encoding bias)*. Η κωδικοποίηση της οντολογίας, θα πρέπει να γίνεται με τον απλούστερο δυνατό τρόπο, απαιτώντας όσο το δυνατόν μικρότερους και απλούστερους κώδικες για την αναπαράστασή της. Η σύλληψη πρέπει να ειδικευτεί στο επίπεδο γνώσης χωρίς να εξαρτάται από συγκεκριμένο επίπεδο συμβόλων κωδικοποίησης. Εμφανίζεται όταν μια απεικόνιση επιλογών γίνεται καθαρά για την ευκολία της παρατήρησης ή της εφαρμογής. Η πόλωση της κωδικοποίησης πρέπει να ελαχιστοποιηθεί, επειδή οι πράκτορες διανομής γνώσης μπορούν να διευκολύνουν με διαφορετική απεικόνιση τα συστήματα και μορφές της παρουσίας.

1.8 Τεχνολογίες δημιουργίας οντολογιών

1.8.1 Μεθοδολογία ανάπτυξης οντολογιών – Κατασκευή της οντολογίας :

Δεν υπάρχουν τυποποιημένες μεθοδολογίες για τη ανάπτυξη οντολογιών παρά μόνο εμπειρικοί κανόνες. Στην εργασία των M. Uschold και M. Gruninger (1996) προτείνονται οι ακόλουθες φάσεις για την ανάπτυξη οντολογιών:

- Προσδιορισμός σκοπιμότητας και πεδίου εφαρμογής
- Κατασκευή της οντολογίας

Σύλληψη (capture) :το στάδιο αυτό περιλαμβάνει τις ενέργειες προσδιορισμού των βασικών εννοιών και των μεταξύ τους σχέσεων, καθώς και την παραγωγή σαφών προδιαγραφών σε μορφή κειμένου των παραπάνω προσδιορισμών. Επιπλέον στο στάδιο της σύλληψης γίνεται συμφωνία για τους όρους με τους οποίους θα αναφερόμαστε στις έννοιες και σχέσεις.

Κωδικοποίηση (coding) : το στάδιο αυτό αναπαριστά την σύλληψη του προηγούμενου σταδίου της σύλληψης σε μια τυπική γλώσσα.

Ενοποίηση (integration) υπάρχουσών οντολογιών : το στάδιο αυτό αναφέρεται στο πως θα πρέπει να χρησιμοποιηθούν υπάρχουσες οντολογίες (ή τμήματα αυτών)

- Αξιολόγηση (evaluation) : το στάδιο αυτό περιλαμβάνει την έκφραση τεχνικών κρίσεων σχετικά με τις οντολογίες, το σχετιζόμενο με αυτές περιβάλλον λογισμικού και τη τεκμηρίωση σε σχέση μ' ένα πλαίσιο αναφοράς , όπως έχει αναφέρει ο Gomez- Perez ,(1995).
- Τεκμηρίωση (documentation) :στο στάδιο αυτό προσδιορίζονται όλες οι σημαντικές παραδοχές τόσο αναφορικά με τις βασικές έννοιες που ορίζονται στην οντολογία όσο και με τα βασικά στοιχεία που χρησιμοποιήθηκαν για την έκφραση των εννοιών στην οντολογία και πρέπει να τεκμηριωθούν.

Σημαντικό να ξεκαθαριστεί το γιατί να κατασκευαστεί η οντολογία και ποιες είναι οι πιθανές χρήσεις της. Χρήσιμο είναι επίσης να προσδιοριστεί ποιοι πρόκειται να είναι οι πιθανοί χρήστες της οντολογίας.

1.8.2 Σημασιολογικά Εργαλεία

Τα σημασιολογικά εργαλεία είναι κάποια εργαλεία που προσπαθούν να καλύψουν το κενό στην επικοινωνία ανθρώπου μηχανής. Ενώ οι περισσότερες από τις τρέχουσες μορφές περιεχομένου του ιστού σχεδιάζονται με σκοπό να

είναι κατανοητές από τον άνθρωπο, δεν είναι κατανοητές από Η/Υ. Το περιεχόμενο του σημασιολογικού ιστού είναι σχεδιασμένο έτσι ώστε να το κατανοούν τόσο ο άνθρωπος όσο και οι Η/Υ.

Για να επιτευχθεί αυτό γίνεται μια προσπάθεια ανάπτυξης εννοιών που να μπορεί να τις «καταλάβει» ένας Η/Υ. Αυτό με λίγα λόγια γίνεται με την οργάνωση της επιστημονικής γνώσης και ορολογίας σε μια τέτοια μορφή που να μπορεί με κάποιες τεχνικές να την κατανοεί ένας ηλεκτρονικός υπολογιστής.

Τέτοια εργαλεία είναι οι :

- *Term lists*: Λίστες όρων όπως authority files, γλωσσάρια και λεξικά.
- *Metadata-like models*: όπως κατάλογοι και γεωγραφικά λεξικά.
- *Classification and categorization schemes*: Σχέδια ταξινόμησης και κατηγοριοποίησης όπως θεματικές επικεφαλίδες, σχήματα ταξινόμησης και κατηγοριοποίησης.
- *Relationship models*: Σχεσιακά Μοντέλα όπως θησαυροί, σημασιολογικά δίκτυα, χάρτες εννοιών και οντολογίες.
- *Metadata content standards*: Πρότυπα για το περιεχόμενο των μεταδεδομένων και ειδικά αυτά που ασχολούνται με την αντιπροσώπευση της γνώσης.
- *Domain-specific content markup languages*: όπως τα DTD (Document Type Definitions) ή τα σχήματα XML.
- *General knowledge representation languages*: Γενικές γλώσσες αντιπροσώπευσης της γνώσης, όπως ο υπολογισμός του πρώτου κατηγορήματος διαταγής, λογικές περιγραφής, εννοιολογικές γραφικές παραστάσεις, γενικά συστήματα πλαισίων.

Πολλά από τα σημασιολογικά εργαλεία που αναφέρονται παραπάνω θεωρούνται συμβατικά συστήματα οργάνωσης της γνώσης όπως οι Term lists, τα Classification and categorization schemes και κάποια από τα σχεσιακά μοντέλα. Η συμβατική αναπαράσταση εννοιών έχει αξία σε μαθησιακά περιβάλλοντα, αφού παρέχουν για παράδειγμα πρόσβαση σε κείμενα και

εφημερίδες, για τα οποία υπάρχει αναπαράσταση εννοιών βασισμένη σε όρους. Πολλά από αυτά τα εργαλεία έχουν μια σημαντική αδυναμία σε περιβάλλον Ψηφιακών Βιβλιοθηκών λόγω των σχετικά αδύναμων δομών τους. Δηλαδή καταλήγουν τελικά να παίρνουν την μορφή γλωσσικών όρων. Πολλά από τα σημαντικά χαρακτηριστικά των επιστημονικών εννοιών όμως, όπως η αναπαράστασή τους, οι έννοιες, οι ιδιότητες, οι σχέσεις με άλλες έννοιες δεν μπορούν να αναπαρασταθούν με τόσο απλούς γλωσσικούς όρους.

Ο σημασιολογικός ιστός θα χρησιμοποιεί σημασιολογικά εργαλεία, όπως οντολογίες για να περιγράψει τους διάφορους πόρους στο δίκτυο. Έτσι, ενώ μέχρι σήμερα το δίκτυο ήταν ένας χώρος που περιέχει μεγάλη ποσότητα ανοργάνωτης, κατά κύριο λόγο, πληροφορίας, πλέον ο ιστός θα γίνει φορέας γνώσης. Αυτό θα επιτευχθεί με την λογική δόμηση της πληροφορίας, του περιεχομένου δηλαδή του ιστού με ένα σημασιολογικό τρόπο, γεγονός που θα επιφέρει πολλές αλλαγές σε σχέση με τον τρόπο που λειτουργεί τώρα ο ιστός. Τέλος να σημειωθεί ότι εκτός από τα σημασιολογικά εργαλεία και τις γλώσσες ορισμού των οντολογιών, υπάρχουν εργαλεία με γραφικό περιβάλλον που διευκολύνουν τη δημιουργία οντολογιών.

Μερικά από αυτά είναι : Protégé, WebODE, OntoEdit, κ.ά, που αναφέρουμε και σε παρακάτω ενότητες.

1.8.3 Έξυπνοι Πράκτορες

Στον σημασιολογικό ιστό οι έξυπνοι πράκτορες θα αλλάξουν τον τρόπο με τον οποίο οι χρήστες πλοηγούνται, βρίσκουν και συλλέγουν πληροφορίες στον ιστό. Αυτό θα γίνει γιατί πλέον θα είναι σε θέση να κατανοήσουν και να αξιολογήσουν την πληροφορία που συναντούν με την βοήθεια των οντολογιών. Επίσης οι πράκτορες πλέον θα μπορούν να συνεννοούνται και να συνεργάζονται

καλύτερα αφού ο σημασιολογικός ιστός θα είναι στην ουσία μια συλλογή από υπηρεσίες που θα περιγράφονται από οντολογίες όπως η DAMLS. Αυτό σημαίνει ότι ο κάθε πράκτορας θα είναι σε θέση να αναγνωρίσει έναν άλλο πράκτορα και την λειτουργία του. Ακόμα περισσότερο ένας πράκτορας θα μπορεί να βάλει έναν άλλο πράκτορα να επιτελέσει μέρος του δικού του έργου.

Οι έξυπνοι πράκτορες πλέον θα μπορούν να κινούνται από τον ένα ιστότοπο στον άλλο και να επιτελούν την διεργασία που τους έχει ανατεθεί αυτόνομα. Είναι προφανές ότι είναι απαραίτητοι στην ανάπτυξη του σημασιολογικού ιστού.

Με την χρήση έξυπνων πρακτόρων θα μειωθεί το φορτίο του δικτύου αφού οι διεργασίες θα γίνονται αυτόματα, πράγμα που σημαίνει ότι θα μειωθεί το φαινόμενο του να 'κρεμάει' το δίκτυο. Ακόμα, υπάρχει υποστήριξη για την ενσωμάτωση πρωτοκόλλων για την καλύτερη διαλειτουργικότητα. Οι πράκτορες εκτελούν τις διεργασίες τους αυτόνομα οπότε δεν χρειάζεται η παρέμβαση του χρήστη. Επίσης μαθαίνουν δυναμικά.

Ακόμη από την φύση τους είναι διαφορετικοί αφού ο κάθε ένας έχει δημιουργηθεί για να διεκπεραιώνει διαφορετική λειτουργία. Αν συνδυάσουμε αυτό το γεγονός με την ικανότητά τους να συνεργάζονται ή να χρησιμοποιούν άλλους πράκτορες για να ολοκληρώσουν την δική τους αποστολή, είναι λογικό να υποθέσουμε ότι καλύπτουν την εργασία τους πιο ποιοτικά.

Τέλος οι πράκτορες είναι ανεκτικοί στα λάθη και αφού μαθαίνουν δυναμικά, τότε θα μπορούσαμε να πούμε ότι διαρκώς βελτιώνονται ώστε να επιτελέσουν το έργο τους με μεγαλύτερη ακρίβεια.

1.8.4 Μηχανές Αναζήτησης

Στον τομέα των μηχανών αναζήτησης ο σημασιολογικός ιστός θα προσφέρει πολλές καινοτομίες. Με την καλύτερη οργάνωση της πληροφορίας μέσω θησαυρών και οντολογιών μπορούν να παρέχουν καλύτερα και ποιοτικότερα αποτελέσματα χωρίς να φορτώνουν τόσο πολύ το δίκτυο. Επίσης ο χρήστης θα είναι πλέον ικανός να διαμορφώνει πιο ελεύθερα ερωτήματα, όχι απαραίτητα με λέξεις-κλειδιά ή με τελεστές Boolean⁸ και η μηχανή αναζήτησης να του φέρνει ικανοποιητικά αποτελέσματα αφού θα μπορεί να τα κατανοήσει. Αυτό θα επιτευχθεί προφανώς με την χρήση της XML και των οντολογιών, με τις οποίες θα μπορεί η μηχανή να κατανοήσει και να αξιολογήσει λογικά το περιεχόμενο μιας σελίδας.

Αυτή τη στιγμή οι μηχανές αναζήτησης στον ιστό έχουν δύο προσεγγίσεις, αυτή των μηχανών *μεγάλης κλίμακας* που στηρίζονται σε έξυπνους πράκτορες και αυτή των μηχανών *μικρής κλίμακας* που στηρίζονται στην ανασκόπηση σελίδων.

- Μηχανές αναζήτησης μεγάλης κλίμακας:

Σ' αυτή την περίπτωση έχουμε μεγάλη ανάκτηση αλλά μικρή ακρίβεια. Αυτό οφείλεται στο γεγονός ότι στην ουσία έξυπνοι πράκτορες συλλέγουν ιστοσελίδες και τις αποθηκεύουν σε μια μεγάλη κεντρική βάση δεδομένων (οι πράκτορες μπορούν να συλλέξουν σχεδόν όλες τις σελίδες του web) αλλά ο μεγάλος αριθμός των ιστοσελίδων μειώνει τον αριθμό των σχετικών αποτελεσμάτων, επομένως η ακρίβεια είναι σε χαμηλά επίπεδα.

- Μηχανές αναζήτησης μικρής κλίμακας:

⁸ επιτρέπει το συνδυασμό όρων , πραγματοποιούνται λογικές πράξεις με τελεστές τις λέξεις AND, OR,NOT και NEAR.

Στην δεύτερη περίπτωση, ακριβώς επειδή έχει γίνει ανασκόπηση της κάθε σελίδας , τα αποτελέσματα φανερόνουν μεγάλη σχετικότητα, άρα υψηλή ακρίβεια. Όμως, όπως είναι λογικό, αυτή η δουλεία γίνεται σε περιορισμένο όγκο ιστοσελίδων, επομένως έχουμε μικρό ποσοστό ανάκτησης.

Και τα δύο είδη μηχανών αναζήτησης είναι βασισμένα σε λέξεις-κλειδιά, πράγμα που σημαίνει ότι και στις δύο περιπτώσεις παρουσιάζεται πρόβλημα, όσον αφορά στο ερώτημα που διατυπώνεται, λόγω των πολλαπλών σημασιών που μπορεί να έχει μια λέξη. Έχουν αναπτυχθεί διάφορες τεχνικές για να ξεπεράσουν αυτό το πρόβλημα με αλγόριθμους που προσπαθούν να καθορίσουν τις συνώνυμες λέξεις για παράδειγμα, αλλά υπάρχει πάντα το πρόβλημα ότι πολλές λέξεις με άσχετη σημασία μπορούν να προέρχονται από κοινή ρίζα ή κοντινές στη σημασία λέξεις να έχουν εντελώς άσχετη ρίζα.

Η λύση σε αυτό το πρόβλημα είναι προφανώς να υπάρχει σημασιολογία στις ιστοσελίδες ώστε να καθορίζεται το νόημα τους, αλλά δεν έχει επικρατήσει κάποια γλώσσα που να καθιερώνει κάποιες σταθερές HTML ετικέτες.

Η XML φαίνεται πολλά υποσχόμενη αφού κρατάει το περιεχόμενο, την δομή και την παρουσίαση χωριστά και είναι πολύ πιο επαρκής για την αναπαράσταση γνώσης. Η συντακτική δομή της όμως την περιορίζει καθώς μπορεί να παρουσιάσει μόνο κάποιες σημασιολογικές ιδιότητες. Σε αυτόν τον τομέα θα εισχωρήσουν οι οντολογίες για την πιο κατανοητή περιγραφή των ιστοσελίδων.

Η πιο δημοφιλής εμποριοποιημένη μηχανή αναζήτησης είναι η Google (www.google.com) ενώ κάνουν την εμφάνιση τους και άλλες όπως η Cuil (www.cuil.com), η Toggle (www.toggle.com) και άλλες.

1.8.5 Robust Hyperlinks

Μία σημαντική καινοτομία στον σημασιολογικό ιστό είναι η δυνατότητα να παράγουμε ‘γερούς’ υπερσυνδέσμους. Αυτό σημαίνει ότι ένα έγγραφο θα είναι διαθέσιμο ακόμα και αν έχει αλλάξει ο υπερσύνδεσμός του ή και μέρος του περιεχομένου του.

Το να δημιουργηθεί ένα robust link είναι κάτι πολύ απλό ως σύλληψη. Για να γίνει ένας υπερσύνδεσμος robust πρέπει να συμπεριληφθεί σε αυτόν, μαζί με το URI (Uniform Resource Identifier) ένα μέρος από το περιεχόμενο του εγγράφου. Μιλάμε δηλαδή για μια lexical signature, μια ψηφιακή υπογραφή δηλαδή που αναγνωρίζει την συγκεκριμένη ιστοσελίδα από το περιεχόμενό της.

Έτσι ένας έξυπνος πράκτορας που γνωρίζει για τα robust hyperlinks, όταν αναζητά, αρχικά θα ψάξει για το URL αγνοώντας την λεξική υπογραφή. Αν όμως αυτή η τεχνική αποτύχει για οποιονδήποτε λόγο, τότε μπαίνει σε μια δεύτερη φάση κατά την οποία αναζητεί κοιτώντας για έγγραφα που η υπογραφή τους είναι κοντινή με την λεξική υπογραφή του αρχικού κειμένου. Για παράδειγμα σε μια ψηφιακή βιβλιοθήκη πολύ συχνά τα διάφορα έγγραφα είναι αποθηκευμένα σε απομακρυσμένα, ανεξάρτητα αποθετήρια. Με την χρήση λοιπόν αυτής της τεχνικής εξασφαλίζεται ότι το έγγραφο θα συνεχίσει να είναι διαθέσιμο αν αλλάξει ο αρχικός του υπερσύνδεσμος. Επομένως συνεισφέρει πολύ στην διαθεσιμότητα των διάφορων τεκμηρίων. Αυτή η τεχνική κατ’ αρχάς καταπολεμά αποτελεσματικά το φαινόμενο των broken links που συχνά απαντάται στο δίκτυο.

1.9 Μηχανική οντολογιών (Engineering ontology)

Διάφορες μεθοδολογίες (Uschold et al., 1995; Gruininger et al., 1995; Fernandez-Lopez et al., 1999; Sure et al., 2002; Pinto et al., 2004) έχουν προταθεί για την μηχανική των οντολογιών μέσα σε περιβάλλοντα διαχείρισης γνώσης. Από την ανάπτυξη (τον καθορισμό των στόχων και των απαιτήσεων),

στην υλοποίηση, στην διασύνδεση, αξιολόγηση, και συντήρηση των εννοιολογικών μορφοποιήσεων, ο κύκλος ζωής της οντολογίας πρέπει να είναι καθορισμένος με σαφήνεια και να υποστηρίζεται από εργαλεία (Sure et al., 2002). Οι πιο παλιές μεθοδολογίες μηχανικής οντολογιών έχουν σχεδιαστεί ώστε να υποστηρίζουν διαδικασίες μηχανικής γνώσης, ξεκινώντας από την ανάλυση και καθορισμό των απαιτήσεων ενός συστήματος βασισμένο στη γνώση, και προχωρώντας μέχρι την ανάπτυξη και την συντήρηση του. Σύμφωνα με όλες αυτές τις προσεγγίσεις, η ανάπτυξη και συντήρηση των οντολογιών είναι μια διαδικασία που καθοδηγείται από τον μηχανικό γνώσης, ο οποίος συλλέγει στοιχεία για την ενημέρωση της οντολογίας και ξεκινάει τη διαδικασία αλλαγής σε μία νέα έκδοση της οντολογίας ύστερα από έλεγχο για πιθανές ασυνέπειες για την εφαρμογή που αξιοποιεί την οντολογία (application-oriented). Μόνο πολύ πρόσφατες μεθοδολογίες μηχανικής οντολογιών (Pinto et al., 2004) δίνουν σημασία στον ρόλο του εργατή γνώσης και στην ενεργή και αποφασιστική συμμετοχή του στις διαδικασίες της οντολογικής μηχανικής που είναι ο καθορισμός απαιτήσεων, η εννοιολογική μορφοποίηση, ο τυπικός ορισμός, η κωδικοποίηση, η συντήρηση, η αξιολόγηση, η διαχείριση εκδόσεων (εξέλιξη) και η τεκμηρίωση (Pinto and Martins, 2004).

Πριν ξεκινήσει κάποιος να αναπτύσσει οντολογίες, ανεξάρτητα με το αν ο σκοπός του είναι να δημιουργήσει ένα σημασιολογικό μοντέλο για κάποια θεματική περιοχή ή να διευκολύνει την ενοποίηση των δεδομένων (data integration), πρέπει πρώτα να αποφασίσει ποιο μοντέλο ανάπτυξης θα ακολουθήσει. Τα βασικά μοντέλα ανάπτυξης οντολογιών είναι δύο: η *‘από πάνω προς τα κάτω’* ανάπτυξη (top-down) και η *‘από κάτω προς τα πάνω’* ανάπτυξη (bottom-up). Κάθε ένα από αυτά τα μοντέλα έχει τα πλεονεκτήματα και τα μειονεκτήματά του και το ποιο θα πρέπει να προτιμήσει ο μηχανικός οντολογιών εξαρτάται κυρίως από την εφαρμογή για την οποία αναπτύσσονται οι οντολογίες. (Τόγιας Κ.Ι., 2007)

Σύμφωνα με το από πάνω προς τα κάτω μοντέλο ανάπτυξης αρχικά αναπτύσσουμε μια οντολογία που αναπαριστά το θεματικό πεδίο στο οποίο εντάσσεται η εφαρμογή, χρησιμοποιώντας θεμελιώδεις (top-level) έννοιες του πεδίου. Στη συνέχεια επεκτείνουμε τη βασική θεματική αυτή οντολογία έτσι ώστε να αποτυπώνει πιο συγκεκριμένες έννοιες, μέχρι να φτάσουμε το σημείο η οντολογία μας να παρέχει μια πλήρη περιγραφή όλων των δεδομένων που πρέπει να καλύπτει. Μια τέτοια προσέγγιση απαιτεί από τους μηχανικούς και τους ειδικούς να χαρακτηρίσουν το θεματικό πεδίο και να προσδιορίσουν τις βασικές έννοιες και συσχετίσεις. Στην πράξη τις περισσότερες φορές είναι δύσκολο να υπάρξει μια τέτοια συμφωνία μεταξύ των ειδικών. Όταν όμως οι δυσκολία αυτή ξεπεραστεί, η παραγόμενη οντολογία αποτελεί ένα πλούσιο σημασιολογικά θεματικό μοντέλο και παρέχει μια σαφή και πλήρη αναπαράσταση του θεματικού πεδίου, που μπορεί να χρησιμοποιηθεί στη συνέχεια για τη συσχέτιση εννοιών που εντάσσονται σε διαφορετικές θεματικές περιοχές. (Τόγιας Κ.Ι., 2007)

Μια εναλλακτική προσέγγιση για την ανάπτυξη οντολογιών είναι η χρήση των δεδομένων ως σημείο εκκίνησης, αναπτύσσοντας αρχικά οντολογίες εστιασμένες στην περιγραφή συγκεκριμένων πηγών δεδομένων, και στη συνέχεια με την διασύνδεση και την επέκταση αυτών των οντολογιών να κινηθούμε προς όλο και μεγαλύτερου επιπέδου αφαίρεση. Αυτή η από κάτω προς τα πάνω προσέγγιση διασφαλίζει ότι τα δεδομένα θα περιγραφούν με ακρίβεια, μια και η οντολογία χτίζεται κυριολεκτικά πάνω σε αυτά. Θεωρητικά, εφόσον τα δεδομένα που έχουμε στα χέρια μας παρέχουν μια καλή αναπαράσταση του θεματικού πεδίου, μια τέτοια προσέγγιση θα οδηγήσει στην ανάπτυξη ενός ακριβούς και πλήρους θεματικού μοντέλου. Επειδή όμως οι πιθανές ασυνέπειες και οι λεπτομέρειες εφαρμογής γίνονται αντιληπτές μόνο στην υψηλού επιπέδου αφαίρεσης θεματική οντολογία, αυτή μπορεί να μην παρέχει πλήρη και ακριβή περιγραφή του θεματικού πεδίου και των σχέσεων

που αναπτύσσονται μέσα σε αυτό, αφού είναι εστιασμένη στις ιδιαιτερότητες των δεδομένων.

Μερικά από τα πιο γνωστά εργαλεία ανάπτυξης και συστημάτων διαχείρισης οντολογιών που μπορούν να χρησιμοποιηθούν είτε για την ανάπτυξη οντολογιών από μηδενικής βάσης, είτε για την επεξεργασία και επαναχρησιμοποίηση οντολογιών κατασκευασμένων από τρίτους, αναφέρονται παρακάτω: Apollo, OilEd, OntoEdit, Ontolingua Server, WebOnto, όπως και το Protégé που αναπτύσσουμε μερικώς και σε παρακάτω υποενότητα. (Τόγιας Κ.Ι., 2007)

Μετά την ολοκλήρωση της κατασκευής ενός συνόλου οντολογιών είναι σημαντική η δημιουργία διασυνδέσεων μεταξύ εννοιών που ανήκουν σε διαφορετικές οντολογίες του συνόλου. Η διασύνδεση μεταξύ των οντολογιών θα βοηθήσει στην επίτευξη της μέγιστης δυνατής σημασιολογικής ολοκλήρωσης ανεξάρτητα από τη μέθοδο ανάπτυξης που ακολουθήθηκε και το αν οι οντολογίες καλύπτουν την ίδια ή διαφορετικές θεματικές περιοχές. Η διασύνδεση μπορεί να πραγματοποιηθεί με δύο τρόπους: είτε με τη χρήση μιας οντολογίας αναφοράς ανώτερου επιπέδου αφαίρεσης, είτε με απευθείας διασύνδεση μεταξύ δυο οντολογιών με τη χρήση σχέσεων όπως η ισοδυναμία (equivalence) ή η υπο-κλάση (subclass) που παρέχονται από όλες τις γλώσσες αναπαράστασης οντολογιών. (Τόγιας Κ.Ι., 2007)

Η αξιολόγηση αποτελεί ένα σημαντικό τμήμα της διαδικασίας ανάπτυξης και (επανα)χρησιμοποίησης οντολογιών και πρέπει να συστηματοποιηθεί προκειμένου να καθιερωθεί η χρήση οντολογιών στο σημασιολογικό ιστό και γενικότερα στις σημασιολογικές εφαρμογές. Οι χρήστες που ανακτούν οντολογίες πρέπει να έχουν στη διάθεσή τους κάποια μέθοδο που να τους βοηθάει να τις αξιολογήσουν και να αποφασίσουν ποιες είναι αυτές που καλύπτουν καλύτερα τις απαιτήσεις τους. Ομοίως, όσοι κατασκευάζουν οντολογίες χρειάζονται έναν τρόπο να αξιολογήσουν τις προκύπτουσες οντολογίες προκειμένου να βελτιώσουν την διαδικασία παραγωγής ή να κάνουν

διορθώσεις. Επίσης, αυτόματες ή ημιαυτόματες τεχνικές ontology learning απαιτούν αποδοτικά μέτρα αξιολόγησης για την επιλογή της ‘καλύτερης’ οντολογίας από ένα σύνολο υποψήφιων οντολογιών, για την επιλογή των βέλτιστων τιμών για τις παραμέτρους του αλγόριθμου μάθησης, ή για την κατεύθυνση της διαδικασίας μάθησης.

Οι μέθοδοι αξιολόγησης που έχουν κατά καιρούς προταθεί μπορούν να κατηγοριοποιηθούν είτε με βάση τον τρόπο προσέγγισης που ακολουθείται, είτε με βάση το επίπεδο στο οποίο εστιάζεται η αξιολόγηση (Τόγιας Κ.Ι., 2007) . Γενικά οι περισσότερες μέθοδοι αξιολόγησης ακολουθούν κάποια από τις παρακάτω προσεγγίσεις :

- Σύγκριση της αξιολογούμενης οντολογίας με ένα ‘χρυσό πρότυπο’ (Golden Standard). Ως χρυσό πρότυπο μπορεί να χρησιμοποιηθεί και μια άλλη οντολογία.
- Χρήση της υπό εξέταση οντολογίας σε κάποια εφαρμογή και αξιολόγηση των αποτελεσμάτων στην εφαρμογή αυτής.
- Σύγκριση της οντολογίας με κάποια πηγή δεδομένων (π.χ.. μια συλλογή εγγράφων) σχετικών με το θεματικό πεδίο που καλύπτει η οντολογία.
- Αξιολόγηση από ανθρώπους που αποφαινόμενοι αν η οντολογία πληρεί κάποιο σύνολο προκαθορισμένων κριτηρίων, προτύπων ή απαιτήσεων.

Καθώς οι οντολογίες και οι σχετικές σημασιολογικές τεχνολογίες εφαρμόζονται σε όλο και μεγαλύτερες εφαρμογές τόσο μεγαλύτερη γίνεται η ανάγκη για εναρμονισμό των τεχνολογιών αυτών με τα βιομηχανικά στάνταρντ και την εξασφάλιση υψηλής απόδοσης. Η συστηματική αξιολόγηση των οντολογιών ίσως οδηγήσει σε προτυποποιημένα μέτρα σύγκρισης και αντίστοιχες πιστοποιήσεις ποιότητας. Τα πιο γνωστά από τα εργαλεία που έχουν αναπτυχθεί για την αξιολόγηση οντολογιών είναι τα εξής : OntoAnalyser, OntoGenerator, OntoClean, ONE-T. (Τόγιας Κ.Ι., 2007)

1.10

Ευθυγράμμιση - Συσχέτιση οντολογιών

Η επαναχρησιμοποίηση των ήδη υπάρχουσών οντολογιών δεν είναι συχνά εφικτή αν δεν καταβληθεί αξιόλογη προσπάθεια. Όταν κάποιος επιθυμεί να επαναχρησιμοποιήσει διαφορετικές οντολογίες μαζί, οι οντολογίες αυτές θα πρέπει κατά κάποιο τρόπο να συνδυαστούν. Αυτό είναι δυνατόν να επιτευχθεί με την *ενοποίηση* (integration) των οντολογιών, που σημαίνει ότι αυτές θα συνενωθούν σε μία νέα οντολογία, ή ακόμα ότι οι οντολογίες μπορεί να παραμείνουν και χωριστά. Και στις δύο περιπτώσεις όμως, οι οντολογίες θα πρέπει να ευθυγραμμιστούν, να έρθουν δηλαδή σε αμοιβαία συμφωνία. Η *ευθυγράμμιση* (alignment) των εννοιών ανάμεσα στις οντολογίες είναι εξαιρετικά δύσκολη, καθώς αυτό απαιτεί κατανόηση της σημασίας των εννοιών. Η ευθυγράμμιση δύο οντολογιών υποδηλώνει το γεγονός ότι θα υπάρξει αλλαγή σε μία τουλάχιστον από αυτές. Η αλλαγή αυτή σημαίνει αυτόματα ότι έχουμε μία νέα έκδοση της οντολογίας. Στην περίπτωση, δε, που οι οντολογίες δεν αναπαριστώνται στην ίδια γλώσσα, τότε συχνά κρίνεται αναγκαία η μετάφρασή τους.

Για να ενισχυθεί η σημασιολογία της ευθυγράμμισης έχει αναπτυχθεί ένα σύνολο οντολογιών που προορίζεται για να συλλάβει τη σημασιολογία και τις σχετικές βαθμίδες για τις αυτοματοποιημένες διαδικασίες οντολογιών, συμπεριλαμβανομένης της ευθυγράμμισης οντολογίας. Αυτές οι οντολογίες είναι μέρος μιας τρέχουσας προσπάθειας να στραφεί η κοινότητα ευθυγράμμισης οντολογιών σε ένα σύνολο κανόνων για την αντιμετώπιση προβλημάτων, ερευνητικούς στόχους και κριτήρια αξιολόγησης. Παρακάτω γίνεται αναφορά σε μερικές από τις βασικές ιδιότητες των οντολογιών

- *Ευθυγράμμιση και ισοδυναμία* : μια ευθυγράμμιση οντολογίας παρέχει μόνο μια σχέση μεταξύ των στοιχείων της δηλαδή σε κάθε περίπτωση εξαρτάται από τις ευθυγραμμίσεις μεταξύ των άλλων στοιχείων, έτσι ο ευθυγραμμιστής δεν μπορεί να πει με βεβαιότητα πως οποιαδήποτε συγκεκριμένη ευθυγράμμιση είναι ιδανική αλλά ότι είναι η καταλληλότερη των δεδομένων.
- *Αρχείο ευθυγράμμισης* : δηλώνει τις περιπτώσεις ότι ένα στοιχείο από μια οντολογία 'Α' αντιστοιχεί με ένα άλλο στοιχείο από μια άλλη οντολογία 'Β' με μερική πιθανότητα, αποδίδοντας στην έξοδο τους αλγορίθμους. Μπορεί να εφαρμοστεί σε πιο ασυνήθιστες ευθυγραμμίσεις όπως μεταξύ μιας κατηγορίας και μιας ιδιοκτησίας ή μίας περίπτωσης και μιας κατηγορίας.
- *Αρχείο αξιολόγησης ευθυγράμμισης* : είναι ένα νέο έγγραφο που προκύπτει από τη σύγκριση ενός αρχείου ευθυγράμμισης (βασισμένο στον αλγόριθμο που έχει αποδοθεί κατά την έξοδο) με το αρχικό αρχείο ευθυγράμμισης.

Η συνδυασμένη χρήση πολλαπλών οντολογιών παρεμποδίζεται από αρκετά προβλήματα, όπως οι αναντιστοιχίες που μπορεί να εμφανίζονται μεταξύ διαφορετικών οντολογιών. Αρχικά, διακρίνουμε δύο επίπεδα στα οποία είναι δυνατόν να παρουσιαστούν αναντιστοιχίες. Το πρώτο είναι το *επίπεδο της γλώσσας* ή το *επίπεδο μετα-μοντέλου* (language or meta-model level), όπου χρησιμοποιούνται τα πρωταρχικά στοιχεία μιας γλώσσας, προκειμένου να οριστεί και να προσδιοριστεί μια οντολογία. Το δεύτερο επίπεδο είναι το *επίπεδο οντολογίας* ή το *επίπεδο μοντέλου* (ontology or model level), στο οποίο ενυπάρχει η πραγματική οντολογία ενός τομέα. Μία αναντιστοιχία σε αυτό το επίπεδο είναι η διαφορά στον τρόπο με τον οποίο μοντελοποιείται το εκάστοτε πεδίο ορισμού.

Στο πρώτο επίπεδο διακρίνονται τα εξής είδη αναντιστοιχιών:

- *Συντακτικό* (Syntax), οι διαφορετικές γλώσσες οντολογιών πολύ συχνά χρησιμοποιούν διαφορετικά συντακτικά..

- *Λογική αναπαράσταση* (Logical representation). ορισμένες γλώσσες οντολογιών δεν εμφανίζουν κοινά στιγμιότυπα .
- *Σημασιολογία των πρωταρχικών στοιχείων* (Semantics of primitives), ακόμα και όταν δύο οντολογίες φαίνεται να χρησιμοποιούν το ίδιο συντακτικό, η σημασιολογία μπορεί να διαφέρει.
- *Εκφραστικότητα της γλώσσας* (Language expressivity), αφορά στη διαφορά στην εκφραστικότητα ανάμεσα σε δύο γλώσσες και υποδηλώνει το γεγονός ότι ορισμένες γλώσσες είναι ικανές να εκφράσουν πράγματα που σε άλλες γλώσσες δεν είναι δυνατόν να εκφραστούν.

Αναντιστοιχίες σε επίπεδο οντολογίας προκύπτουν όταν συνδυάζονται δύο ή περισσότερες οντολογίες, που περιγράφουν (εν μέρει) αλληλεπικαλυπτόμενους τομείς. Αυτές οι αναντιστοιχίες δημιουργούνται, είτε όταν οι οντολογίες γράφονται στην ίδια γλώσσα, είτε όταν χρησιμοποιούν διαφορετικές γλώσσες. Σύμφωνα με τη βιβλιογραφία, διακρίνονται αρκετά είδη αναντιστοιχιών στο επίπεδο αυτό, ως προς τα εξής:

- *Εμβέλεια* (Scope), δύο κλάσεις φαίνονται να αναπαριστούν την ίδια έννοια, αλλά δεν έχουν ακριβώς τα ίδια στιγμιότυπα.
- *Κάλυψη μοντέλου και διαβάθμιση* (Model coverage and granularity), αφορά το θεματικό τομέα που καλύπτεται από την οντολογία.
- *Παράδειγμα* (Paradigm), χρησιμοποιούνται προκειμένου να αναπαραστήσουμε έννοιες όπως είναι ο χρόνος, η αιτιότητα, η κίνηση, οι διαδικασίες κλπ.
- *Περιγραφή έννοιας* (Concept description), αφορά στις συμβάσεις μοντελοποίησης οντολογιών (π.χ. μια διατριβή μπορεί να μοντελοποιηθεί ως εξής : Διατριβή < Επιστημονικό βιβλίο < Βιβλίο).
- *Αναντιστοιχία όρου* (term mismatch), μπορεί να διακριθεί σε: α) *Συνώνυμους όρους* (Synonym terms), οι έννοιες αναπαριστώνται με διάφορα ονόματα π.χ. (ο όρος “Car” σε μια οντολογία και ο όρος “automobile” σε

μια άλλη οντολογία). β) *Ομώνυμους όρους* (Homonym terms), η έννοια και η ερμηνεία ενός όρου διαφέρει από περιβάλλον σε περιβάλλον.

- *Κωδικοποίηση* (Encoding), οι τιμές στις οντολογίες μπορούν να κωδικοποιηθούν σε διάφορες διατάξεις, π.χ. (μια ημερομηνία μπορεί να αναπαρασταθεί ως “ημέρα/μήνας/χρόνος” ή ως “μήνας-μέρα-χρόνος”).

Συμπερασματικά σκοπός της δημιουργίας αυτών των οντολογιών είναι η καλύτερη συνεργασία μεταξύ των μελών της ερευνητικής κοινότητας ευθυγράμμισης οντολογίας. Με ένα κοινό αντιπροσωπευτικό σχέδιο για τη δήλωση και αξιολόγηση των ευθυγραμμίσεων, γίνεται σημαντικά ευκολότερη η σύγκριση αλγορίθμων ευθυγράμμισης. Επιπλέον, κανονικά σύνολα δεδομένων απαιτούνται για να εξασφαλιστούν ακριβείς συγκρίσεις. Η ευθυγράμμιση οντολογίας δεν είναι πολύτιμη για την υπηρεσία της ίδιας, αλλά είναι σημαντική μόνο στην υπηρεσία κάποιας άλλης λειτουργίας η οποία την απαιτεί. Στοχεύει στην αυτοματοποιημένη ευθυγράμμιση οντολογίας για τους πράκτορες οι οποίοι σημασιολογικά επικοινωνούν με ετερογενή συστήματα δεδομένων. Υπό αυτήν τη μορφή, οι οντολογίες για την ευθυγράμμιση οντολογίας πρέπει να αναπτυχθούν ώστε να καλύψουν τις περιπτώσεις χρήσης της σημασιολογικής λειτουργικότητας.

1.11

Συγχώνευση οντολογιών

Σημαντική είναι η συνεισφορά των οντολογιών στην προσπάθεια για δημιουργία επαναχρησιμοποιήσιμου και δια-μοιραζόμενου λογισμικού. Προς το παρόν, ένας από τους κύριους παράγοντες που εμποδίζουν την επαναχρησιμοποίηση των εφαρμογών, αλλά και την επικοινωνία μεταξύ διαφορετικών συστημάτων είναι το γεγονός της χρήσης διαφορετικών όρων και εννοιών για την περιγραφή του ίδιου πεδίου δραστηριότητας. Αναπτύσσοντας οντολογίες, δημιουργούμε τη βάση για την επικοινωνία πολύπλοκων

συστημάτων που θα χρησιμοποιούν ένα κοινό λεξιλόγιο, μια κοινή γλώσσα επικοινωνίας.

Οι οντολογίες όπως έχουμε ήδη αναφέρει, αποτελούν την τεχνολογία «κλειδί» για την διαμόρφωση και αξιοποίηση της πληροφορίας με σκοπό την αποτελεσματικότερη διαχείριση της γνώσης και την ανάπτυξη του Σημασιολογικού Ιστού και των εφαρμογών του. Μέσα σε ένα καταναμημένο περιβάλλον όπως είναι ο Σημασιολογικός Ιστός, οι οντολογίες αποτελούν μία κοινή ορολογία για τα μέλη μιας κοινωνίας που θέλουν να συνδέσουν, την γνώση που διαμορφώνεται μέσα από την πρακτική και την αλληλεπίδραση. Παρόλα αυτά, φαίνεται πως πάντα θα υπάρχει πάνω από μία οντολογίες για το ίδιο πεδίο ενδιαφέροντος, τα συστήματα αναζήτησης πληροφοριών (search engines) πρέπει να απαντούν σε ερωτήσεις άντλησης πληροφοριών αποτελεσματικά, γεφυρώνοντας τα κενά μεταξύ των διαφορετικών οντολογιών και των προτιθέμενων ερμηνειών των εννοιών που περιέχονται μέσα σε αυτές. Προς αυτήν την κατεύθυνση, δίκτυα εννοιολογικά συσχετιζόμενης πληροφορίας πρέπει να δημιουργούνται κατά απαίτηση. Γι' αυτό, ο συντονισμός (ευθυγράμμιση, αντιστοίχιση, συγχώνευση) των οντολογιών είναι μια μεγάλη πρόκληση για να γεφυρωθεί το κενό μεταξύ των πρακτόρων (προγραμμάτων και ανθρώπων) που αξιοποιούν την πληροφορία και που έχουν διαφορετικές εννοιολογικές μορφοποιήσεις για το ίδιο πεδίο ενδιαφέροντος.

Υπάρχουν σήμερα αρκετές δουλειές προς την κατεύθυνση της αντιστοίχισης/ συγχώνευσης οντολογιών. Αυτές οι δουλειές αξιοποιούν γνώση βασισμένη στη λεκτική ομοιότητα των εννοιών, στη δομή της οντολογίας, στο πεδίο ενδιαφέροντος, και σε εμπειρικούς κανόνες ταιριάσματος (Giunchiglia et al., 2004). Πρόσφατες προσεγγίσεις στοχεύουν στην αξιοποίηση όλων αυτών των τύπων γνώσης και επιπλέον στην καταγραφή των προτιθέμενων άτυπων ερμηνειών των όρων χρησιμοποιώντας εμπειρικούς κανόνες (Kalfoglou et al., 2003; Serafini et al., 2003). Απώτερος στόχος των πιο πρόσφατων προσεγγίσεων είναι ο αυτοματισμός της διαδικασίας συντονισμού

(coordination) των οντολογιών, ειδικά σε περιβάλλοντα όπως ο Σημασιολογικός Ιστός, όπου δεν υπάρχει δυνατότητα συνεχούς και άμεσου ελέγχου της εγκυρότητας των νέων παραγόμενων οντολογιών από μηχανικούς γνώσης, αλλά και που ο χρόνος για την παραγωγή των συγχωνευμένων οντολογιών πρέπει να είναι όσο το δυνατόν μικρότερος. Παρόλα αυτά, δεν υπάρχει κάποια προσέγγιση που να επιτυγχάνει τον απόλυτο αυτοματισμό της διαδικασίας αυτής. Πρόσφατες προσεγγίσεις υποστηρίζουν περιορισμένη (Noy, 2003) ή μηδαμινή (Giunchiglia et al., 2004) ανάμειξη του χρήστη στην διαδικασία της συγχώνευσης και αντιστοίχισης αντίστοιχα, χωρίς όμως να εγγυούνται την ορθότητα της παραγόμενης οντολογίας.

Η δημιουργία εργαλείων λογισμικού που θα βοηθούσαν στη συγχώνευση οντολογιών και στη μετάφραση των όρων μιας οντολογίας σε άλλη, θα έκανε δυνατή την επικοινωνία μεταξύ συστημάτων που βασίζονται σε διαφορετικές οντολογίες. Οι οντολογίες αναμένεται να αλλάξουν θεμελιωδώς τον τρόπο με τον οποίο κατασκευάζονται τα διάφορα συστήματα. Έως σήμερα, οι βάσεις γνώσεων παρείχαν μικρές δυνατότητες επαναχρησιμοποίησης και διαμοιρασμού. Κάθε σχεδιαστής θα έπρεπε να ξεκινήσει από το μηδέν. Μελλοντικά αναμένεται η δημιουργία βιβλιοθηκών από οντολογίες. Έτσι, θα διευκολύνεται η σύνθεση βάσεων γνώσης από στοιχεία που θα εξάγονται από τις βιβλιοθήκες αυτές.

Η συγχώνευση οντολογιών αναμένεται να αποτελέσει σημαντικό πεδίο έρευνας και για το χώρο της ηλεκτρονικής μάθησης. Οι έννοιες των διαφόρων μαθησιακών αντικειμένων μπορεί να περιέχονται σε διαφορετικές οντολογίες. Σε ότι προαναφέρθηκε, υπονοήθηκε η χρήση μιας και μοναδικής οντολογίας, στην οποία εντάσσονται οι έννοιες των μαθησιακών αντικειμένων. Είναι προφανές επομένως, ότι καθώς θα προτείνονται νέες οντολογίες και καθώς θα ωριμάζουν οι τεχνολογίες του Σημασιολογικού Ιστού, η συγχώνευση οντολογιών θα αποτελέσει πεδίο έρευνας.

1.12

Μερικές εφαρμογές των οντολογιών

Εφαρμογές που παρέχουν ολοκληρωμένο πλαίσιο εννοιών και ορολογίας μεταξύ ανθρώπων με διαφορετικές ανάγκες και οπτικές γωνίες στα πλαίσια ενός οργανισμού και που διευκολύνουν την επικοινωνία των ανθρώπων στα πλαίσια του οργανισμού.

1.12.1 Οι οντολογίες στο ηλεκτρονικό εμπόριο

Μια ειδική μορφή Ενοποίησης Εφαρμογών (Enterprise Application Integration) είναι το Ηλεκτρονικό Εμπόριο, όπου εφαρμογές σε διαφορετικές επιχειρήσεις, ή ακόμα και αυθαίρετες οντότητες στον Ιστό πρέπει να ενοποιηθούν. Το Ηλεκτρονικό Εμπόριο μπορεί να χωριστεί σε δύο περιοχές: στο B2C (Business-to-Consumer) and B2B (Business-to-Business).

Οι οντολογίες μπορούν να εφαρμοστούν και στις δύο περιοχές για να αυξήσουν την αποτελεσματικότητα των αναζητήσεων προϊόντων και να κάνουν ευκολότερη τη συνεργασία των ηλεκτρονικών επιχειρήσεων. Για παράδειγμα, στην περιοχή B2C, οι οντολογίες μπορούν να χρησιμοποιηθούν για την σύγκριση συγκεκριμένων ιδιοτήτων (π.χ. τιμή, έξοδα αποστολής) του ίδιου προϊόντος από πολλούς διαφορετικούς (ηλεκτρονικούς) πωλητές. Αυτή τη στιγμή χρησιμοποιεί ένας **wrapper** για κάθε πωλητή από όπου πρέπει να εξαχθούν πληροφορίες. Ο wrapper χρησιμοποιεί εμπειρικούς κανόνες (heuristics) για να εξάγει κείμενο από τις ιστοσελίδες του προϊόντος του πωλητή. Για κάθε πωλητή, ένας ξεχωριστός wrapper πρέπει να φτιαχτεί, και όταν ο πωλητής αλλάξει τη μορφή της ιστοσελίδας του προϊόντος ο wrapper πρέπει απαραίτητως να ενημερωθεί. Με την χρήση των οντολογιών από τους

πωλητές, ο πράκτορας (agent) απλά χρησιμοποιεί την οντολογία του πωλητή ή μια αντιστοίχιση από την δικιά του (του πράκτορα) οντολογία πάνω στην οντολογία του πωλητή, έτσι ώστε να ανακτά την πληροφορία για το προϊόν με τυπικό τρόπο. Έτσι, πληροφορίες για το προϊόν μπορούν να αναζητούνται από ένα αριθμό πωλητών και να ενοποιούνται από τον πράκτορα, πριν ακόμα εμφανιστεί το αποτέλεσμα στον καταναλωτή.

1.12.2 E-learning ontology (Ηλεκτρονική μάθηση οντολογιών)

Οι οντολογίες παίζουν όλο και πιο σημαντικό ρόλο στην ανάπτυξη συστημάτων ηλεκτρονικής μάθησης (e-learning systems). Στα σύγχρονα συστήματα ηλεκτρονικής μάθησης οντολογίες χρησιμοποιούνται για την αναπαράσταση του μαθησιακού υλικού (περιεχόμενο, μορφή, δομή), των μαθησιακών εννοιών και οντοτήτων (π.χ.. μάθημα, τάξη, μαθητής) και των μεταξύ τους σχέσεων. Πρότυπα μεταδεδομένων (Dublin Core, IEEE LOM) και ειδικές οντολογίες για την αναπαράσταση μαθησιακού υλικού έχουν προταθεί και χρησιμοποιηθεί σε ευφυή συστήματα e-learning, ενώ εκπαιδευτικές εφαρμογές όπως οι CIPHER, Connections, Conzilla, Edutella και άλλες βασίζουν τη λειτουργία τους σε οντολογίες.

Για την αξιολόγηση οντολογιών μαθησιακού υλικού προτείνουμε και υλοποιούμε με τη μορφή οντολογίας ένα μοντέλο αξιολόγησης λαμβάνοντας υπ' όψιν *συντακτικές* (syntactic), *σημασιολογικές* (semantic), *πραγματιστικές* (pragmatic) και *κοινωνικές* (social) μετρικές, καθώς και μετρικές *συνάφειας* (cohesion metrics).

Η *συντακτική* ανάλυση μιας οντολογίας αφορά τις συσχετίσεις μεταξύ των συμβόλων των (λέξεων, φράσεων, προτάσεων). Η *σημασιολογική* ανάλυση μιας οντολογίας μελετά τις σχέσεις ανάμεσα στα σύμβολα και τις έννοιες που αυτά αναπαριστούν. Η *πραγματιστική* ανάλυση αναφέρεται στη χρησιμότητα της

οντολογίας για τους χρήστες και τους πράκτορες λογισμικού που την προσπελαίνουν. Οι κοινωνικές μετρικές λαμβάνουν υπ' όψιν το γεγονός ότι πράκτορες λογισμικού (software agents) και οντολογίες συνυπάρχουν και επικοινωνούν μεταξύ τους σχηματίζοντας κοινότητες. Τέλος, οι μετρικές *συνάφειας* αναφέρονται στο βαθμό που οι κλάσεις μιας οντολογίας συσχετίζονται μεταξύ τους μέσω των ιδιοτήτων τους.

Αν και όλες οι μετρικές θεωρούνται εξίσου σημαντικές, και μπορούν να έχουν ίσα βάρη κατά τον υπολογισμό της ολικής ποιότητας, τα βάρη κάποιων μετρικών μπορούν να μεταβάλλονται ανάλογα με την εφαρμογή. Αφού καθοριστούν τα βάρη, το μοντέλο μπορεί να χρησιμοποιηθεί για την αξιολόγηση μιας ή περισσότερων οντολογιών μαθησιακού υλικού. Το μοντέλο περιλαμβάνει ένα τμήμα εξαγωγής συμπερασμάτων που περιέχει κανόνες που μετρούν τις επιμέρους ποιότητες μιας οντολογίας και μπορεί να υλοποιηθεί εύκολα σε OWL.

Ειδικές οντολογίες για την αναπαράσταση μαθησιακού υλικού έχουν προταθεί και χρησιμοποιηθεί σε ευφυή συστήματα e-learning (Τόγιας Κ.Ι., 2007). Κάποια από τα πιο γνωστά συστήματα e-learning που χρησιμοποιούν οντολογίες για την αναπαράσταση μαθησιακού υλικού είναι τα Edutella, Conzilla, SchollOnto και RichODL (Τόγιας Κ.Ι., 2007).

1.12.3 Ψηφιακές Βιβλιοθήκες

Οι Βιβλιοθήκες είναι κατά παράδοση ο φορέας οργανωμένης γνώσης. Στο νέο περιβάλλον του σημασιολογικού ιστού οφείλουν να προσφέρουν την εμπειρία που διαθέτουν για την οργάνωση της γνώσης από το συμβατικό περιβάλλον αλλά και να αξιοποιήσουν κατάλληλα όλα τα νέα εργαλεία, ώστε να συνεχίσουν να επιτελούν το κοινωνικό τους έργο διατηρώντας ίδια υψηλά επίπεδα ποιότητας. Έχουν το χρέος προς την κοινωνία να φέρουν σε επαφή το

συμβατικό περιβάλλον λειτουργίας τους με το ψηφιακό, με στόχο ποιοτικές υπηρεσίες για τον τελικό αποδέκτη, δηλαδή τον χρήστη.

Στις Ψηφιακές Βιβλιοθήκες η κύρια αλλαγή είναι ότι το τεκμήριο πλέον έχει αλλάξει υπόσταση, έχει περάσει δηλαδή από την υλική στην ψηφιακή. Ακόμα έχουν αλλάξει τα δεδομένα πρόσβασης στο υλικό τους, το περιεχόμενο των συλλογών τους και το οικονομικό τους περιβάλλον.

Οι συλλογές τους πλέον απαρτίζονται από ψηφιακά πολυμεσικά αντικείμενα σε διάφορα formats, ηλεκτρονικά βιβλία και περιοδικά, videos, φωτογραφίες κ.α. Το κοινό επίσης έχει αλλάξει. Πρόκειται για ένα παγκόσμιο κοινό που μπορεί να ζητά πρόσβαση στην πληροφορία από οποιοδήποτε μέρος του πλανήτη, γεγονός που θέτει και θέμα πολυγλωσσικότητας. Το οικονομικό περιβάλλον είναι διαφορετικό, πλέον η βιβλιοθήκη εκτός από κοινοφελής οργανισμός πρέπει να είναι και βιώσιμος, παρέχοντας υπηρεσίες προστιθέμενης αξίας.

Ο σημασιολογικός ιστός προσφέρει νέες δυνατότητες οργάνωσης, διαχείρισης και διάθεσης των συλλογών τους. Αυτό σημαίνει ότι με την χρήση των δυνατοτήτων που προσφέρουν τα σημασιολογικά εργαλεία, μπορούν να παρέχουν νέες υπηρεσίες, πιο γρήγορα και πιο αξιόπιστα, με την ελάχιστη ανθρώπινη προσπάθεια.

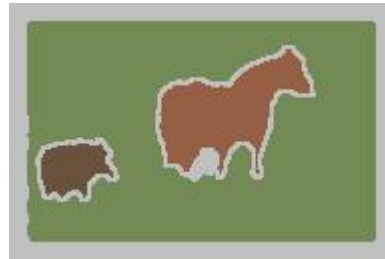
Όλα όσα αναφέραμε πιο πάνω για τον σημασιολογικό ιστό έχουν άμεση εφαρμογή στις ψηφιακές βιβλιοθήκες.

Το επόμενο βήμα για καλύτερη αναγνώριση είναι να συνδέσουμε τις λέξεις με συγκεκριμένες περιοχές εικόνων. Για να γίνει αυτό πρέπει να λυθεί ένα πρόβλημα ανταπόκρισης γιατί δεν είναι γνωστό σε ποια μέρη εικόνων αναφέρονται οι λέξεις από πριν. Γι' αυτό η αναγνώριση προσεγγίζεται ως η μετάφραση των οπτικών ερεθισμάτων σε σημασιολογικά συστατικά. Με αυτή τη λογική έχει αναπτυχθεί και η υπηρεσία Blobworld, που αναλύεται παρακάτω.

Το Blobworld είναι ένα σύστημα για ανάκτηση εικόνων βασισμένο στο περιεχόμενό τους. Αυτό το πετυχαίνει χωρίζοντας αυτόματα κάθε εικόνα σε

περιοχές που αντιπροσωπεύουν αντικείμενα η μέρη από αντικείμενα. Έτσι μπορούν οι χρήστες να διατυπώσουν ερωτήματα που βασίζονται στα αντικείμενα που περιέχουν οι φωτογραφίες.

Για παράδειγμα εδώ έχουμε μια φωτογραφία δύο αλόγων και την παρουσίασή τους σε Blobworld:



Όταν ο χρήστης πάει να διατυπώσει ένα ερώτημα επιλέγει ένα σχήμα από την blob εικόνα. Αλλιώς μπορεί να ψάξει με λέξεις κλειδιά. Με λέξεις κλειδιά η αναζήτηση θα γίνει σε Corel λέξεις-κλειδιά, σε τίτλους ή τίτλους από cd, και τότε η αναζήτηση γίνεται καθαρά ανάμεσα σε εικόνες που ταιριάζουν με τις λέξεις-κλειδιά. Αυτό βέβαια προσφέρει ένα ποσοστό ανακρίβειας με τους γνωστούς περιορισμούς που ισχύουν σε μια αναζήτηση με λέξη -κλειδί.

Αν διαλέξουμε να ψάξουμε με το σχήμα από το blob θα μας βγει μια σελίδα με επιπλέον επιλογές, όπως πόσο σημαντική είναι για εμάς η περιοχή που επιλέξαμε, πόσο σημαντικά είναι διάφορα στοιχεία της περιοχής όπως το χρώμα, το σχήμα, η τοποθεσία και το σχήμα/μέγεθος. Τέλος πόσο σημαντικό είναι το υπόβαθρο της εικόνας (ότι βρίσκεται έξω από την περιοχή που επιλέξαμε). Εκεί έχει ήδη κάποιες επιλογές που μπορούμε αν θέλουμε να τις προσαρμόσουμε ή να τις αφήσουμε όπως είναι. Ακολουθώντας πατάμε submit και μας φέρνει τα αποτελέσματα στα οποία αναγράφεται και ποσοστό ακρίβειας, όπως και πόσες εικόνες μας έφερε και πόσες είναι στο σύνολό τους οι εικόνες. Ακόμη αναφέρεται πόσος χρόνος χρειάστηκε για να γίνει το ευρετήριο, να επεξεργαστεί το σύστημα το ποσοστό ακρίβειας και να δημιουργήσει τις εικόνες blobworld. Από το συγκεκριμένο ερώτημα τα αποτελέσματα ήταν αρκετά

σχετικά αφού στα πρώτα 20 αποτελέσματα τα 13 περιείχαν άλογα, ενώ όλα τα λάθος αποτελέσματα περιείχαν παρόμοιο υπόβαθρο και προφανώς αν προσαρμόζαμε καλύτερα την αναζήτηση θα ήταν πιο αξιόπιστο το αποτέλεσμα. Επίσης ο χρόνος που χρειάστηκε ήταν πολύ μικρός, μόλις 5 δευτερόλεπτα. Βέβαια δεν ξέρουμε πως θα αντιδρούσε το σύστημα αν είχε να απαντήσει σε πολλούς χρήστες ταυτόχρονα, πράγμα που πιθανώς δεν συνέβη όταν κάναμε το ερώτημα αφού το πρόγραμμα είναι ακόμα σε επίπεδο έρευνας.

1.12.4 Ιατρική – βιολογία – βιοϊατρική

Η βιοϊατρική πληροφορία που υπάρχει σήμερα διαθέσιμη σε ψηφιακή μορφή, έχει ήδη χαρακτηριστεί «ζούγκλα πληροφορίας» (information jungle - G. Eysenbach), αφηγηματικής (narrative) μορφής με βασικό πλέον (αρνητικό) χαρακτηριστικό την ιδιαίτερη δυσκολία αξιοποίησής της. Ο στόχος της επιστημονικής – ερευνητικής προσπάθειας είναι η μετάβαση από την κατάσταση αυτή σε μια νέα όπου η πληροφορία θα είναι δομημένη με τρόπο ώστε να μπορεί να χαρακτηριστεί ως συλλογή – δεξαμενή γνώσης (knowledge repository) και η οποία θα επιτρέπει τη χρήση προηγμένων τεχνολογιών διαχείρισης γνώσης (knowledge management). Για να επιτευχθεί αυτό θα πρέπει σε πρώτη φάση να κατασκευαστούν εξελιγμένα εργαλεία για την αναζήτηση, τη συσχέτιση και το χαρακτηρισμό των βιοϊατρικών κειμένων. Η Οντολογία αποτελεί τον πλέον ενδεδειγμένο και αξιοποιήσιμο μηχανισμό για την κατασκευή τέτοιων εργαλείων ώστε να έχουμε άριστη αποτύπωση γνώσης μίας θεματικής περιοχής. Καθορίζει τον τρόπο αναπαράστασης με τον οποίο οι έννοιες τα αντικείμενα και οι μεταξύ τους σχέσεις αποτυπώνονται.

Για την ανάπτυξη μιας οντολογίας βιοϊατρικών όρων αρχικά θα πρέπει να καθοριστούν οι προδιαγραφές για την κατάρτιση σώματος (corpus) ιατρικών κειμένων και να σχεδιαστεί η οντολογία. Παράλληλα, θα πρέπει να σχεδιαστεί

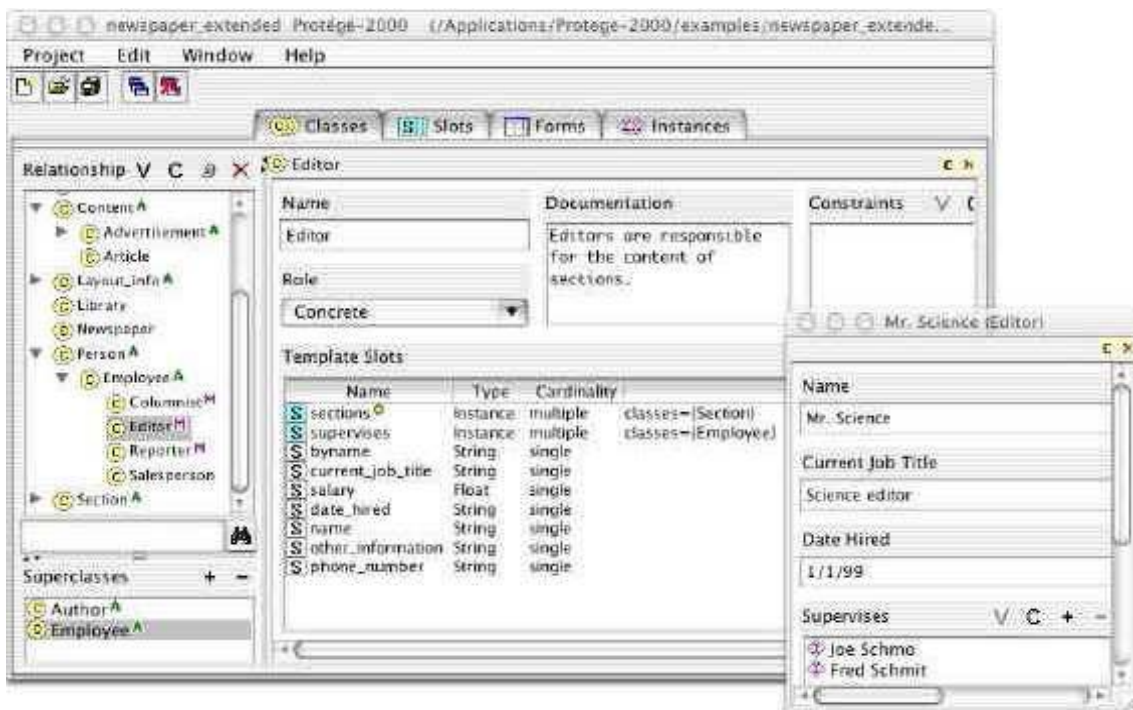
το μοντέλο αναπαράστασης της οντολογίας και να ξεκινήσει η ανάπτυξη πληροφοριακού συστήματος για τη διαχείριση και επεξεργασία των κειμένων του corpus. Στη συνέχεια θα πρέπει να υπάρξει συλλογή και επεξεργασία βιοϊατρικών κειμένων αλλά και συγκρότηση της πρώτης ταξινόμιας βιοϊατρικών εννοιών, πάνω στην οποία θα στηριχτεί η κατασκευή της οντολογίας.

Η ανάπτυξη μιας οντολογίας βιοϊατρικών όρων έχει ως σκοπό η επιστημονική κοινότητα του κλάδου της βιοϊατρικής να αποκτήσει μια σειρά εργαλείων που θα υποβοηθήσουν σε μεγάλο βαθμό το έργο της: από την καλύτερη αναζήτηση σε ιατρικές πηγές, έως την αυτόματη δεικτοδότηση, χαρακτηρισμό ή άλλη επεξεργασία των κειμένων. Παράλληλα ο απλός χρήστης μέσω μιας εξειδικευμένης πύλης αναζήτησης ιατρικών θεμάτων, θα μπορεί να εντοπίσει ευκολότερα πιο αξιόπιστη και χρήσιμη πληροφορία για θέματα υγείας που τον απασχολούν. Επιπλέον φορείς υγείας που παράγουν πληροφορία ιατρικού περιεχομένου (νοσοκομεία, δημόσια διοίκηση, κλπ) θα μπορούν να επεξεργαστούν τα δεδομένα τους με πληθώρα τρόπων με στόχο την ικανοποίηση των σκοπών λειτουργίας τους (π.χ. μελέτες συσχέτισης νόσων – ιατρικών εξετάσεων για τα νοσοκομεία, κλπ).

1.13 Το Protégé ως εργαλείο εφαρμογής οντολογιών

Το Protégé (<http://protege.stanford.edu>) (Noy et al., 2001) (Noy et al., 2000) είναι ένα εργαλείο για διαχείριση γνώσης που αναπτύσσεται από το πανεπιστήμιο του Stanford. Χιλιάδες χρήστες σε όλο τον κόσμο χρησιμοποιούν το Protégé σε έργα που ποικίλουν από τη μοντελοποίηση της νόσου του καρκίνου, έως την μοντελοποίηση πυρηνικών σταθμών παραγωγής ενέργειας. Το Protégé είναι ελεύθερο λογισμικό και διατίθεται κάτω από την άδεια χρήσης ελεύθερου κώδικα “Mozilla open-source license”.

Το Protégé παρέχει ένα γραφικό και διαδραστικό περιβάλλον για σχεδιασμό οντολογιών και ανάπτυξη βάσεων γνώσης. Επιτρέπει στους μηχανικούς και τους ειδικούς εμπειρογνώμονες να εκτελούν με ευκολία εργασίες διαχείρισης γνώσης. Όσοι αναπτύσσουν οντολογίες μπορούν να προσπελαίνουν τις σχετικές πληροφορίες γρήγορα όταν τις χρειάζονται και να διαμορφώνουν και να χειρίζονται οντολογίες. Η ιεραρχική δομή της οντολογίας αναπαριστάται με τη μορφή δένδρου δίνοντας στο χρήστη γρήγορη και εύκολη πρόσβαση στις κλάσεις και υπο-κλάσεις. Για την εισαγωγή τιμών στις ιδιότητες το Protégé χρησιμοποιεί ειδικές φόρμες (Εικόνα 1.5).



Εικόνα 1.5 : Στιγμιότυπο οθόνης του Protege

Το γνωστικό μοντέλο του Protégé είναι συμβατό με το πρωτόκολλο Open Knowledge Base Connectivity (OKBC). Υποστηρίζει κλάσεις και ιεραρχίες κλάσεων με πολλαπλή κληρονομικότητα, templates και slots, ορισμό προκαθορισμένων και αυθαίρετων facet για τα slots με σύνολα επιτρεπόμενων τιμών, περιορισμούς πλήθους, εξ ορισμού τιμές, και αντίστροφα slots, μετακλάσεις και ιεραρχίες μετα-κλάσεων. Εκτός από την ευχρηστία του, το

Protégé διακρίνεται από τα υπόλοιπα περιβάλλοντα διαχείρισης οντολογιών για τις δυνατότητες εξέλιξης και επέκτασης που προσφέρει.

Το Protégé έχει με επιτυχία χρησιμοποιηθεί για την κατασκευή και τη χρήση οντολογιών αποτελούμενων από 150.000 πλαίσια (frames). Η δυνατότητα υποστήριξης βάσεων γνώσης αποτελούμενων από εκατοντάδες χιλιάδες πλαίσια, απαιτεί την παρουσία δύο βασικών υποσυστημάτων: μιας βάσης δεδομένων για την αποθήκευση και προσπέλαση των δεδομένων και ενός μηχανισμού αποθηκευτικής μνήμης (caching) έτσι ώστε να είναι δυνατός ο χειρισμός νέων πλαισίων όταν ο αριθμός τους στη βασική μνήμη έχει ξεπεράσει το μέγιστο όριο.

Ένα από τα βασικότερα πλεονεκτήματα της αρχιτεκτονικής του Protégé είναι ότι το σύστημα είναι κατασκευασμένο με ανοικτό και αρθρωτό σχεδιασμό. Η βασισμένη σε υποσυστήματα αρχιτεκτονική του δίνει τη δυνατότητα προσθήκης νέων λειτουργιών με τη δημιουργία αντίστοιχων πρόσθετων υπομονάδων (plugins). Η βιβλιοθήκη πρόσθετων λειτουργιών του Protégé περιέχει συνδρομές από χιλιάδες προγραμματιστές από όλο τον κόσμο.

Κάποιες από τις διαθέσιμες προσθήκες παρέχουν λειτουργίες όπως :

- υποστήριξη για αποθήκευση και εισαγωγή οντολογιών γραμμένων σε RDF Schema, αρχεία XML με DTD, αρχεία XML Schema και DAML+OIL,
- εμφάνιση εικόνων μορφής GIF, _ίντεο και ήχου, καθώς και έναν επεξεργαστή διαγραμμάτων που επιτρέπει στον χρήστη να κατασκευάσει μια βάση γνώσης σχεδιάζοντας ένα διάγραμμα του οποίου οι κόμβοι και οι ακμές απεικονίζουν πλαίσια διαφορετικού είδους, ανάλογα με το σχήμα και το χρώμα,
- οπτικοποίηση οντολογιών (OntoViz plugin), διασύνδεση οντολογιών (Prompt plugin), μηχανές εξαγωγής συμπερασμάτων (Flora plugin, PAL plugin), εισαγωγή στοιχείων από πηγές του διαδικτύου (UMLS plugin, WordNet plugin) κλπ. (Τόγιας Κ.Ι., 2007)

ΚΕΦΑΛΑΙΟ 2: ΓΛΩΣΣΕΣ ΑΝΑΠΑΡΑΣΤΑΣΗΣ ΟΝΤΟΛΟΓΙΩΝ

Οι γλώσσες των οντολογιών προσφέρουν τη δυνατότητα στις μηχανές να μπορούν πλέον να τις επεξεργαστούν καλύτερα και να κατανοήσουν τα δεδομένα τα οποία μέχρι σήμερα περιορίζονται απλώς να παρουσιαστούν, αποδίδοντας έτσι πιο ακριβή και ολοκληρωμένα συμπεράσματα.

2.1 Που διακρίνονται οι γλώσσες αναπαράστασης οντολογιών

Το μεγάλο πρόβλημα με τον παγκόσμιο ιστό είναι ότι το μεγαλύτερο μέρος της διαθέσιμης πληροφορίας οργανώνεται με τρόπο που καθιστά δύσκολη την επεξεργασία της από μηχανή. Η πληροφορία αποθηκεύεται συνήθως σε μορφή αρχείων τα οποία προσφέρουν μόνο οπτική απεικόνιση και όχι σημασιολογική ταξινόμηση της πληροφορίας. Αυτό έχει ως αποτέλεσμα ένας άνθρωπος να μπορεί να διαχωρίζει σημασιολογικά την πληροφορία που περιέχει ένα τέτοιο αρχείο, αλλά αυτό είναι αδύνατο για μία εφαρμογή.

Το παραπάνω πρόβλημα που επιδιώκεται να λυθεί με το σημασιολογικό ιστό δεν περιορίζεται μόνο στο διαδίκτυο, αλλά αφορά και οποιονδήποτε άλλο τομέα στον οποίο προκύπτει η ανάγκη για αποθήκευση της πληροφορίας.

Για το λόγω αυτό δεν έχει δημιουργηθεί μια μόνο γλώσσα ανάπτυξης οντολογιών, αλλά ποικίλες γλώσσες που ολοένα εξελίσσονται και αναπτύσσονται προκειμένου να βελτιωθεί η αναζήτηση της πληροφορίας.

Τις γλώσσες αναπαράστασης οντολογιών θα μπορούσε κανείς να τις διακρίνει σε τρεις κατηγορίες :

§ Τις web – based γλώσσες : η σύνταξη των οποίων βασίζεται στη βασική μορφή της γλώσσας προγραμματισμού της xml όπως παράδειγμα η γλώσσα : Owl, Rdfs, Oil, Daml.

§ Τις γλώσσες που αναπτύχθηκαν για να αναπαραστήσουν συγκεκριμένες οντολογίες και να χρησιμοποιηθούν σε συγκεκριμένες εφαρμογές ,όπως παράδειγμα της CysL, Grail , Nkrl. Η CyCL για παράδειγμα είναι μια επίσημη γλώσσα της οποίας η σύνταξη προέρχεται από το first-order υπολογισμού κατηγορήματος (η γλώσσα της επίσημης λογικής) και από τον ψευδισμό, προκειμένου να εκφραστεί η γνώση κοινής λογικής.

§ Τις γλώσσες που αναπτύχθηκαν για αν αναπαραστήσουν τη λογική των πλαισίων περιγραφικά, βασισμένες στη κατηγορηματική λογική της πρώτης τάξης ,όπως παράδειγμα : η Loom, Carin, Flogic, Ontolingua. (Ντιούδης, 2008)

2.2 Που διαφέρουν οι γλώσσες αναπαράστασης οντολογιών

Όπως έχει ειπωθεί σε προηγούμενο κεφάλαιο, οντολογία είναι η περιγραφή με τη χρησιμοποίηση ενός συγκεκριμένου λεξιλογίου , ενός συνόλου από έννοιες, αντικείμενα και σχέσεις μεταξύ τους που αφορούν μια συγκεκριμένη περιοχή γνώσης. Ουσιαστικά μια οντολογία είναι μια ιεραρχία από κλάσεις, ιδιότητες και στιγμιότυπα κλάσεων που περιγράφουν ένα γνωστικό αντικείμενο. Παρόλο όμως που το γνωστικό αντικείμενο μιας πληροφορίας μπορεί να θεωρείτε ίδιο, οι γλώσσες αναπαράστασης των οντολογιών μπορούν να διαφέρουν σε καίρια σημεία όπως :

§ Σύνταξη : διαφορετικές εντολές μπορούν να δοθούν από κάθε γλώσσα και να εκφράσουν την ίδια πληροφορία

§ Σημασιολογία : η ίδια δήλωση μπορεί να σημαίνει διαφορετικά πράγματα σε διάφορες γλώσσες

§ Εκφραστικότητα: κάτι που μπορούμε να εκφράσουμε σε μια γλώσσα δεν μπορούμε σε μια άλλη

§ Ορολογία : διαφορετικοί όροι μπορεί να δοθούν για βασικές έννοιες όπως class-concept(τάξη – έννοια) , instance-object (παράδειγμα – αντικείμενο), slot-property (slot - ιδιοκτησίας).

2.3 Web – based γλώσσες

Σύμφωνα με την Wikipedia (WikiXml, 2008), η XML είναι μια γενικευμένη γλώσσα σήμανσης που προτείνεται για τη δημιουργία γλωσσών ειδικού σκοπού που είναι ικανές να περιγράφουν πολλά είδη δεδομένων. Ο κυριότερος σκοπός της είναι να διευκολύνει το διαμοιρασμό δεδομένων μεταξύ διαφορετικών συστημάτων, και ποιο συγκεκριμένα συστημάτων που είναι συνδεδεμένα στο διαδίκτυο. Οι γλώσσες που βασίζονται στην XML καθορίζονται με επίσημο τρόπο , τέτοιο ώστε να επιτρέπεται η ελεύθερη μετακίνηση των δεδομένων μέσω του διαδικτύου, τους δίνεται η δυνατότητα συνεργασίας με πολλές και διάφορες τεχνολογίες με σκοπό τη δημιουργία μεγάλων διασυνδέσεων δικτύων και τέλος έχουν την ικανότητα να προβάλουν τις υπηρεσίες ιστού (web services) στις οποίες ουσιαστικά επεκτείνουν το διαδίκτυο το οποίο προσφέρει υπηρεσίες σε έναν χρήστη σε ένα δίκτυο το οποίο με τη σειρά του προσφέρει υπηρεσίες σε ένα άλλο λογισμικό.

Χαρακτηριστικές γλώσσες που βασίζονται στη σύνταξη της XML είναι οι :

§ Ontology Web Language (OWL)

§ Ontology Interchange Language (OIL)

§ Ontology Exchange Language (XOL)

§ Simple HTML Ontology Extensions (SHOE)

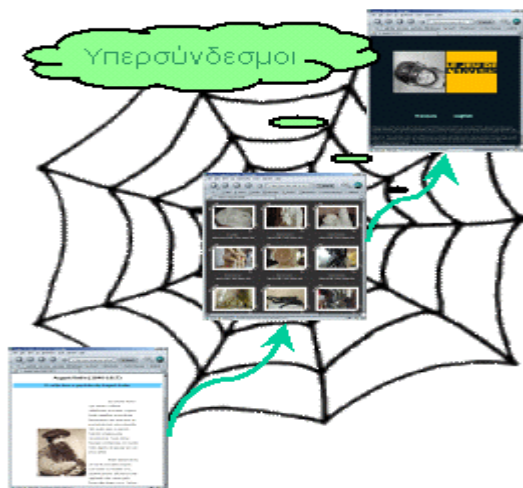
§ Resource Description Framework Schema Language (RDFS)

§ Darpa Agent Markup Language (DAML)

§ Ontology Markup Language (OML and KML) (Ντιούδης, 2008)

2.4 Λίγα λόγια για την γλώσσα HTML

Ο Παγκόσμιος Ιστός (World Wide Web) αποτελεί πλέον στην εποχή μας ένα βασικό μέσο επικοινωνίας και μεταφοράς γνώσεων και πληροφοριών. Αν το πρώτο βήμα προς την εξάπλωση του παγκόσμιου ιστού ήταν η τεχνολογική εξέλιξη των δικτύων που επέφερε την ταχύτατη και εύκολη μεταφορά οποιασδήποτε πληροφορίας από τον ένα υπολογιστή στον άλλο, η εδραίωση της γλώσσας υπομνηματισμού html (hypertext markup language) ως η γλώσσα αναπαράστασης της πληροφορίας στον παγκόσμιο ιστό, αποτέλεσε τον «κοινό κώδικα επικοινωνίας» όλης αυτής της διακινούμενης πληροφορίας. Η html είναι απλή, εύκολη στη χρήση της και ανθρωπίνως αναγνώσιμη, αφού μια σελίδα γραμμένη σε html είναι απλά ένα έγγραφο γραμμένο σε έναν οποιοδήποτε επεξεργαστή κειμένου. Ειδικές σημάνσεις (tags) ερμηνεύουν τον τρόπο παρουσίασης της πληροφορίας, «μαρκάρωντάς» την με την ανάλογη σήμανση. Κατάλληλες εφαρμογές με το όνομα φυλλομετρητές του παγκόσμιου ιστού (Web browsers), «αποκωδικοποιούν» την html πληροφορία και την αναπαριστούν με τον κατάλληλο τρόπο στην οθόνη του υπολογιστή κάθε χρήστη. Έτσι, ο φυλλομετρητής ενεργεί σαν μια ομοιόμορφη γραφική διεπαφή, μέσα από την οποία δίνεται πρόσβαση σε πληροφορίες που διακινούνται στον παγκόσμιο ιστό. Υπερσύνδεσμοι (hyperlinks) συνδέουν τις σελίδες μεταξύ τους, δημιουργώντας ένα γιγάντιο δίκτυο πληροφοριών (εικόνα 2.1). Ακολουθώντας τους διάφορους υπερσυνδέσμους, κάθε χρήστης έχει πρόσβαση σε κάθε είδους πληροφορία διακινείται στον παγκόσμιο ιστό, μέσα από τον φυλλομετρητή του.



Εικόνα 2.1 World Wide Web

2.5 Η γλώσσα XML

XML είναι τα αρχικά των λέξεων eXtensible Markup Language. Η XML είναι ένα σύνολο από κανόνες για τη δημιουργία ετικετών (tags) που περιγράφουν τα δεδομένα ενός εγγράφου ενώ προσδιορίζουν τα διάφορα μέρη από τα οποία αποτελείται ένα έγγραφο. Η XML είναι μια μεταγλώσσα αφού μπορεί να προσδιορίζει την σύνταξη άλλων γλωσσών.

Η XML δεν αποτελεί άλλη μια γλώσσα σήμανσης, αλλά είναι μια γλώσσα με την οποία μπορούμε να ορίσουμε άλλες γλώσσες σήμανσης. Οι γλώσσες σήμανσης όπως η HTML, καθορίζουν ένα σύνολο από ετικέτες οι οποίες περιγράφουν ένα σύνολο από στοιχεία. Σε περίπτωση που κάποιος χρήστης χρησιμοποιεί μια γλώσσα η οποία δεν περιέχει κάποια επιθυμητά στοιχεία δεν μπορεί να κάνει τίποτα άλλο από το να βρει μια νέα γλώσσα ή να περιγράψει για μια νέα έκδοση της γλώσσας που ήδη χρησιμοποιεί.

Εν αντίθεση η XML είναι μια μεταγλώσσα σήμανσης. Με την XML μπορεί κάποιος να δημιουργήσει ένα σύνολο από ετικέτες που επιθυμεί και να τις χρησιμοποιεί όπως αυτός ορίσει.

Οι ετικέτες, όπως είπαμε και προηγουμένως, μπορούν να ακολουθούν ένα σύνολο από κανόνες, ενώ η τεκμηρίωση μπορεί να υπάρχει σε ένα DTD⁹- Document Type Definition. Ένα DTD αρχείο περιέχει ένα λεξικό για συγκεκριμένα XML έγγραφα. Τέτοια έγγραφα υπάρχουν για ένα έγγραφο XML ενώ κάποιος μπορεί να δημιουργήσει ένα δικό του.

Η XML καθορίζει την δομή και την σύνταξη γλωσσών σήμανσης που πρέπει να ακολουθούν, όπως η MusicML, η MathML, η CML. Εάν μια εφαρμογή μπορεί και καταλαβαίνει τη σύνταξη αυτών των γλωσσών τότε μπορεί αμέσως να κατανοήσει και όλες τις γλώσσες που προέρχονται από αυτές τις γλώσσες. Ένας browser δεν είναι απαραίτητο να γνωρίζει όλες τις ετικέτες που υπάρχουν, αλλά διαβάζοντας το έγγραφο ή ένα DTD μπορεί να κατανοήσει και την δομή αλλά και την λειτουργικότητα του εγγράφου. Τέλος οι βασικές λεπτομέρειες για τον τρόπο απεικόνισης καθορίζονται σε ένα άλλο αρχείο που εμπεριέχεται μέσα στο έγγραφο.

Για παράδειγμα ας αναλογιστούμε τις μαθητικές εξισώσεις (οι οποίες υπάρχουν σχεδόν σε όλα τα επιστημονικά έγγραφα). Οι μαθητικές εξισώσεις για πάρα πολλά χρόνια δεν υποστηριζόταν από κάποιο γνωστό φυλλομετρητή. Ένας επιστήμονας θα έπρεπε, για να μεταδώσει ένα επιστημονικό έγγραφο στο διαδίκτυο. Με την XML θα μπορούσε να δημιουργήσει ετικέτες που θα περιγράφουν κάθε σύμβολο τι σημαίνει, ενώ με ένα άλλο έγγραφο θα μπορούσε να καθορίσει τον τρόπο με τον οποίο θα παρουσιαζόταν αυτό το έγγραφο στον φυλλομετρητή. Με την XML δεν είναι απαραίτητο να περιμένει κάποιος φυλλομετρητή να υποστηρίξει μια συγκεκριμένη λειτουργία σχετικά με τα έγγραφα του. Αντίθετα μπορεί να δημιουργήσει όποιος θέλει μια γλώσσα που να περιγράφει τις λειτουργίες που θέλει και κατόπιν να δηλώσει στον φυλλομετρητή τον τρόπο με τον οποίο θα τα εμφανίζει.

⁹ Ο καθορισμός τύπου εγγράφου DTD υπάρχει στην επίσημη προδιαγραφή της XML και αποτελεί μια γλώσσα δημιουργίας σχημάτων περιορισμένων δυνατοτήτων

2.5.1 Πλεονεκτήματα χρήσης της XML

Η XML έχει ένα μεγάλο αντίκτυπο σε πολλούς τομείς του κόσμου της πληροφορικής. Τα βασικά πλεονεκτήματα της είναι η απλότητα της και η δυνατότητα δημιουργίας γλωσσών. Πέραν όμως απ' αυτά υπάρχουν μερικά ακόμη γεγονότα στα οποία οφείλεται η επιτυχία της XML, αυτά στηρίζονται στο γεγονός ότι : τα αρχεία της XML είναι ευανάγνωστα, υποστηρίζονται από πολλές επιχειρήσεις και υπάρχει ένα σύνολο από τεχνολογίες οι οποίες υποστηρίζουν την επεξεργασία και τροποποίηση των XML εγγράφων έτσι ώστε να είναι δυνατή η προβολή τους στον παγκόσμιο ιστόν, ενώ είναι ιδανικά για τη δημιουργία αναφορών.

Γενικά η XML αποτελεί μεταγλώσσα σήμανσης η οποία περιγράφει την δομή και το περιεχόμενο ενός εγγράφου, ενώ δεν ασχολείται καθόλου με την μορφοποίηση των στοιχείων. Το αρχείο περιέχει μόνο ετικέτες οι οποίες καθορίζουν το είδος του εγγράφου και το περιεχόμενό του. Το πλεονέκτημα αυτό μεγάλο, καθώς το έγγραφο είναι πιο ευανάγνωστο από το ανθρώπινο μάτι αλλά και πιο εύκολο να αναγνωρίσει και να εντοπίσει συγκεκριμένα σημεία. Εν αντίθεση η HTML περιέχει ετικέτες σχετικές με την μορφή, τη δομή και την σημασία των περιεχομένων. Έτσι η XML σε σχέση με την HTML είναι περισσότερο ευέλικτη για να χρησιμοποιηθεί από διαφορετικούς χρήστες για διαφορετικές χρήσεις και αυτό χάρη στις απεριόριστες ετικέτες που μπορούν να δημιουργηθούν. Οι ετικέτες οροθετούν και μαρκάρουν κομμάτια του εγγράφου καθώς επίσης προσθέτουν επιπλέον πληροφορία που βοηθά στον ορισμό της δομής. Το κείμενο μεταξύ των ετικετών είναι το περιεχόμενο του εγγράφου, το οποίο μπορεί να είναι απλό κείμενο ή οτιδήποτε άλλο. Η ανατομία ενός εγγράφου XML ξεκινά πάντα με την ετικέτα : `<?xml version="1.0" ?>`, ενώ το περιεχόμενο μπορεί να περιέχεται μέσα στην ετικέτα `<message>` όπου δίνεται

ουσιαστικά η εντολή πως ακολουθεί κείμενο, το οποίο μόλις ολοκληρώνεται κλείνει με την ετικέτα </message>. (Αυτομ, 2006)

2.5.2 Αδυναμίες χρήσης της XML

Παρόλα τα δυνατά της σημεία που την καθιστούν ικανή στη δημιουργία μεγάλων και πολύπλοκων εγγράφων καθώς τα δεδομένα της είναι δομημένα, και παρόλο που έχει υιοθετηθεί από την κοινότητα των προγραμματιστών καθώς είναι μια γλώσσα εύκολα επεκτάσιμη, η XML δεν παύει να βασίζεται σε μοντέλα δεδομένων χαμηλού επιπέδου. Το γεγονός αυτό την καθιστά ανίκανη να μπορεί να χρησιμοποιηθεί στη δημιουργία οντολογιών εξειδικευμένων πεδίων ή οντολογικών λεξικών καθώς δεν μπορεί να χρησιμοποιήσει βασικές οντολογικές αρχές μοντελοποίησης. Επιπλέον δεν θεωρείται κατάλληλη για διαμοιραζόμενες πηγές στον παγκόσμιο ιστό καθώς δεν διαθέτει μηχανή συμπεράσματος.

Όλα τα παραπάνω οδήγησαν στη δημιουργία νέων γλωσσών που βασίζονται στο συντακτικό της XML, ωστόσο είναι πιο επεκτάσιμες και ανεπτυγμένες.

2.6 Η γλώσσα RDF/S- (Resource Description Framework- Πλαίσιο Περιγραφής Πόρων)

Η RDF (Resource Description Framework) συστήνεται από τη W3C (World Wide Web Consortium) ως μια ολοκληρωμένη γλώσσα δημιουργίας, ανταλλαγής, επεξεργασίας και επαναχρησιμοποίησης μεταδεδομένων πληροφοριών μέσα στο διαδίκτυο. Προσφέρει ένα τυποποιημένο μοντέλο

δεδομένων για την περιγραφή πληροφοριών ώστε να είναι δυνατή η ανάγνωση και η κατανόηση τους από τους υπολογιστές.

Το βασικό μοντέλο του RDF αποτελείται από τρεις βασικούς τύπους αντικειμένων, τους Πόρους (Resources), τις Ιδιότητες (Properties) και τις Προτάσεις (Statements).

§ Resources : οτιδήποτε μπορεί να περιγραφεί με RDF εκφράσεις ονομάζεται πόρος. Σε κάθε πόρο, ο οποίος μπορεί να είναι ένα οποιοδήποτε αντικείμενο προσπελάσιμο ή μη από τον παγκόσμιο ιστό αποδίδεται ένα μοναδικό αναγνωριστικό URI¹⁰ (Uniform Resource Identifier), το οποίο μπορεί να είναι ένα URL (Unified Resource Locator) ή οτιδήποτε άλλα μπορεί να προσδιορίσει μοναδικά ένα resource.

§ Properties : με τη βοήθεια του RDF μοντέλου ορίζονται ιδιότητες και σχέσεις με τις οποίες περιγράφονται οι ιδιότητες ενός πόρου (τα resources) σχηματίζοντας τριάδες της μορφής <υποκείμενο, κατηγορημα, αντικείμενο> τις οποίες αποκαλούμε δηλώσεις.

§ Statements : κάθε περιγραφή σε RDF δηλώνει ότι ορισμένα αντικείμενα resources (δικτυακοί τόποι, πρόσωπα, αντικείμενα, συγκεκριμένα έγγραφα ή οτιδήποτε) έχουν κάποιες ιδιότητες με ορισμένες τιμές property, με κάθε μέρος αυτής της δήλωσης να είναι μοναδικά ορισμένο.

Πέρα από τις RDF περιγραφές των πληροφοριακών πόρων, το μοντέλο δεδομένων RDF αποτελείται και από μια γλώσσα ορισμού σχημάτων (συνόλων κλάσεων και ιδιοτήτων), την RDF Schema Specification Language [RDFS¹¹]. Με τη βοήθεια του RDFS, είναι δυνατός ο προσδιορισμός μηχανισμών καθορισμού κλάσεων πόρων, καθώς και ο περιορισμός των πιθανών συνδυασμών κλάσεων μεταξύ τους, χρησιμοποιώντας κατάλληλες συσχετίσεις.

¹⁰ Εκτός από την ονομασία URI/IRI, οι προσδιοριστές πόρων είναι γνωστοί και ως URIs όταν χρησιμοποιούν αποκλειστικά URI για τον προσδιορισμό πόρων

¹¹ Resource Description Framework Schema, δεν παρέχει ένα λεξιλόγιο με έτοιμες κλάσεις εφαρμογών και ιδιοτήτων, αλλά παρέχει υποστήριξη για την περιγραφή τέτοιων κλάσεων και ιδιοτήτων και για την ένδειξη των κλάσεων και ιδιοτήτων που συνδυάζουν

Ένα RDF σχήμα αποτελείται από τις δηλώσεις κλάσεων, γνωρισμάτων και των σχέσεων μεταξύ των κλάσεων. Όμοιοι πόροι είναι ομαδοποιημένοι κάτω από την ίδια κλάση.

Με βάση τα παραπάνω, μπορούμε να διακρίνουμε τρία διαφορετικά επίπεδα

αφαίρεσης στο μοντέλο δεδομένων RDF/S: Στο κατώτερο επίπεδο υπάρχουν οι ίδιοι οι πόροι (έγγραφα, δικτυακοί τόποι, πρόσωπα ή οτιδήποτε άλλο). Το επόμενο επίπεδο αφαίρεσης είναι το *επίπεδο δεδομένων*, όπου γίνεται η περιγραφή των πληροφοριακών πόρων με τη χρήση λεξιλογίων, τα οποία περιγράφονται στο τελευταίο επίπεδο, το *επίπεδο σχήματος*. Το επίπεδο σχήματος είναι το επίπεδο αφαίρεσης όπου αναπτύσσονται RDF σχήματα για να διευκολύνουν τη σημασιολογική περιγραφή των πόρων. Σε αυτό το επίπεδο, οι *κλάσεις* αναπαριστούν αφηρημένες οντότητες και αναφέρονται συλλογικά σε σύνολα παρομοίων αντικειμένων.(Carroll et al., 2004), (W3C, 2004)

2.6.1 Γραφική απεικόνιση RDF/ S

Για την γραφική απεικόνιση των RDF/S περιγραφών και σχημάτων χρησιμοποιείται ένα μοντέλο κατευθυνόμενων γράφων με ετικέτες τόσο στις ακμές όσο και στους κόμβους που μπορεί εύκολα να συνδυάσει πολλά διαφορετικά λεξιλόγια και να επεκταθεί προσθέτοντας απλώς περισσότερες ακμές. Οι κόμβοι αναπαριστούν αντικείμενα (πόρους ή κλάσεις) και οι ακμές συμβολίζουν σχέσεις μεταξύ των κόμβων (ιδιότητες). Ως κόμβους αναπαριστούμε επίσης και τύπους *Literal*, δηλαδή αλφαριθμητικά και άλλους βασικούς τύπους δεδομένων, όπως αυτοί ορίζονται στο πρότυπο XML schema. Οι κόμβοι συμβολίζονται ως ελλείψεις και οι ατομικές τιμές ως παραλληλόγραμμα. Οι ακμές μπορεί να είναι τριών ειδών: *απόδοσης γνωρισμάτων (attributes)*, *δημιουργίας στιγμιότυπων (instances)* και *υπαλληλίας*.

Οι ακμές απόδοσης γνωρισμάτων αναπαριστούν γνωρίσματα κόμβων και σχέσεις μεταξύ τους, ενώ οι ακμές υπαλληλίας χρησιμοποιούνται για να δηλώσουν σε επίπεδο σχήματος ότι ένας κόμβος (κλάση) ή ιδιότητα είναι υποκατηγορία ενός σημασιολογικά ευρύτερου κόμβου ή ιδιότητας αντίστοιχα. Τέλος, οι ακμές δημιουργίας στιγμιότυπων αποτελούν τη σύνδεση ανάμεσα στα πρότυπα RDF και RDFS¹², επιτρέποντας τη δημιουργία στιγμιότυπων μίας κλάσης και την απόδοση τύπων σε πληροφοριακούς πόρους που περιγράφονται.

Έτσι η αναπαράσταση RDF σχημάτων περιγραφής σε σύνταξη RDF/XML, ξεκινάει πάντα με την ετικέτα: `<?xml version="1.0"?>` και την ετικέτα: `<rdf:RDF xml:lang="ex" xmlns:rdf="http://www. _ _ _ -rdf-syntax-ns >"` ενώ ακολουθεί εντολή που ορίζει το έγγραφο μέσω `xmlns:rdfs` και ετικέτες σε γλώσσα `<rdf: =""/>`.

Γενικά, το RDF φέρνει ένα Πλαίσιο Δεδομένων για τον παγκόσμιο ιστό. Το RDF – το Πλαίσιο Περιγραφής Πόρων – αποτελεί πρότυπο για να γίνονται απλές περιγραφές. Αυτό που είναι η XML για το συντακτικό, το RDF είναι για τη σημασιολογία, ένα ξεκάθαρο σύνολο κανόνων για την παροχή απλής περιγραφικής πληροφορίας. Το RDF Schema τότε, παρέχει ένα τρόπο ώστε αυτές οι περιγραφές να μπορούν να συνδυαστούν σε ένα μοναδικό λεξιλόγιο.

Ενσωματώνεται σε ποικιλία εφαρμογών συμπεριλαμβανομένων των :

§ Καταλόγων Βιβλιοθηκών

§ Παγκοσμίων Καταλογών

§ Δημοσιεύσεων και συλλογής νέων, λογισμικού και περιεχομένου

§ Προσωπικών συλλογών μουσικής, φωτογραφιών και εκδηλώσεων

Το RDF είναι μέρος μια μεγάλης προόδου στη δύναμη του Παγκοσμίου Ιστού. Ουσιαστικά, χρήστες και εφαρμογές μπορούν να συνδυάσουν τις πληροφορίες που αναπαριστώνται με RDF από διάφορες πηγές στον παγκόσμιο ιστό, με τρόπους που μέχρι τώρα δεν είχαν συλληφθεί. Η ομάδα εργασίας RDF

¹² Το σύστημα των τύπων της RDFS είναι κατά κάποιον τρόπο παρόμοιο με τα συστήματα των τύπων των αντικειμενοστραφών γλωσσών προγραμματισμού JAVA, SMALLTALK.

Core έχει μετατρέψει τις προδιαγραφές του RDF σε πρακτικό αλλά και μαθησιακά ακριβές υπόβαθρο, πάνω στο οποίο η OWL (Ontology Web Language- Γλώσσα Οντολογίας του Παγκόσμιου Ιστού) και ο υπόλοιπος σημασιολογικός ιστός μπορούν να χτιστούν ,σύμφωνα με τα όσα ανέφερε ο πρόεδρος της ομάδας εργασίας της RDF Core το 2004 στη κοινοπραξία του παγκόσμιου ιστού και στην τελική έγκριση των δύο τεχνολογιών του σημασιολογικού ιστού, αυτή του αναθεωρημένου Πλαισίου Περιγραφής Πόρων (RDF) και της Γλώσσας Οντολογίας του Παγκόσμιου Ιστού (OWL). (Carroll et al., 2004), (W3C, 2004)

2.7 Πότε εμφανίζεται η OWL – Γλώσσα Οντολογίας του Παγκόσμιου Ιστού

Τα RDF και OWL είναι πρότυπα του Σημασιολογικού Ιστού που παρέχουν ένα πλαίσιο διαχείρισης του δυναμικού, επιχειρησιακή ένταξη, καθώς και κοινή χρήση και επαναχρησιμοποίηση δεδομένων στον Παγκόσμιο Ιστό. Αυτές οι πρότυπες μορφές επεκτείνονται σε εύρος εφαρμογών, επιχειρήσεις, καθώς και σε κοινότητες. Όλες οι διαφορετικές κοινότητες ‘χρηστών’ μπορούν να μοιραστούν την ίδια πληροφορία, ακόμη και αν δε χρησιμοποιούν το ίδιο λογισμικό. Αυτό που χρειάζεται ωστόσο είναι ένας τρόπος ανάπτυξης θεματικών ή σχετικών με την περιοχή λεξιλογίων. Αυτός είναι ο ρόλος μιας οντολογίας. Μια οντολογία ορίζει τους όρους που χρησιμοποιούνται για να περιγράψουν και να αναπαραστήσουν μια περιοχή γνώσης. Οι οντολογίες χρησιμοποιούνται από ανθρώπους, βάσεις δεδομένων και εφαρμογές που χρειάζεται να μοιράζονται πληροφορίες σχετικές με το θέμα(περιοχή) – όπως η ιατρική, κατασκευή εργαλείων, κτηματαγορά, επισκευές αυτοκινήτων, οικονομική διαχείριση, κλπ. Οι οντολογίες περιλαμβάνουν ορισμούς βασικών εννοιών της περιοχής καθώς και τις σχέσεις μεταξύ τους, που είναι

χρησιμοποιήσιμοι από υπολογιστές. Κωδικοποιούν τη γνώση σε μια περιοχή καθώς και τη γνώση που διαπερνά περιοχές. Με αυτό τον τρόπο κάνουν τη γνώση επαναχρησιμοποιήσιμη.

Η OWL- η Γλώσσα του Παγκόσμιου Ιστού- παρέχει μια γλώσσα για τον ορισμό δομημένων οντολογιών που βασίζονται στον παγκόσμιο ιστό, η οποία φέρνει πλουσιότερο συνδυασμό και διαλειτουργικότητα δεδομένων μεταξύ περιγραφικών κοινοτήτων. Εκεί όπου προγενέστερες γλώσσες είχαν χρησιμοποιηθεί για την ανάπτυξη εργαλείων και οντολογιών για συγκεκριμένες κοινότητες χρηστών (ειδικά σε επιστήμες και ειδικά για κάθε επιχείρηση εφαρμογές ηλεκτρονικού εμπορίου), δεν είχαν ορίσει να είναι συμβατές με την αρχιτεκτονική του παγκόσμιου ιστού γενικότερα και του σημασιολογικού ιστού συγκεκριμένα.

2.8Η γλώσσα OWL (Ontology Web Language- Γλώσσα Οντολογίας του Παγκόσμιου Ιστού)

Η OWL (Web Ontology Language) είναι μια πολύ δημοφιλή γλώσσα περιγραφής οντολογιών, πιο εκφραστική ως προς την αναπαράσταση του νοήματος και της σημασιολογίας κάποιων όρων και σχέσεων από τις RDF και RDF(S). Αποτελεί την πιο πρόσφατα ανεπτυγμένη γλώσσα για οντολογίες από το W3C.

Το πιο συνηθισμένο συντακτικό που ακολουθεί η OWL βασίζεται στην RDF/XML και έχει σχεδιαστεί για χρήση από εφαρμογές που χρειάζεται να επεξεργάζονται το πληροφοριακό περιεχόμενο των εγγράφων, σε αντίθεση με τις περιπτώσεις που απαιτείται απλά παρουσίαση του περιεχομένου στους ανθρώπους.

Η OWL μπορεί να χρησιμοποιηθεί για την περιγραφή της σημασίας των όρων που τοποθετούνται σε λεξιλόγια, και των σχέσεων μεταξύ αυτών των όρων. Αυτή η αναπαράσταση των όρων και των σχέσεων που τους διέπουν ονομάζεται οντολογία. Η OWL έχει περισσότερες δυνατότητες για την έκφραση των σημασιών και των σημασιολογιών από τις τεχνολογίες XML, RDF και RDFS, και επομένως μπορεί να περιγράψει ακόμη καλύτερα το Διαδικτυακό περιεχόμενο που διερμηνεύουν οι μηχανές (machine interpretable).

2.8.1 Οι τρεις υπογλώσσες της OWL

Η OWL παρέχει τρεις υπογλώσσες με αυξανόμενες δυνατότητες έκφρασης (increasingly expressive sublanguages)¹³ οι οποίες έχουν σχεδιαστεί για χρήση από συγκεκριμένες κοινότητες χρηστών (Mc Guinness et al., 2004):

- **OWL-LITE** : Είναι η πιο απλή συντακτικά και η λιγότερο εκφραστική υπογλώσσα. Χρησιμοποιείται για οντολογίες που αποτελούνται από μια απλή ιεραρχία κλάσεων και απλούς περιορισμούς. Παράδειγμα, αν η OWL-Lite υποστηρίζει τους περιορισμούς εύρους τιμών, οι επιτρεπτές τιμές του εύρους είναι μόνο το μηδέν ή το ένα. Επίσης έχει μικρότερη πολυπλοκότητα από την OWL-DL.
- **OWL-DL**: Είναι πιο εκφραστική από την OWL-Lite και βασίζεται στα Description Logics¹⁴. Μια οντολογία εκφρασμένη σε OWL-DL μπορεί να χρησιμοποιηθεί από κάποιο μηχανισμό εξαγωγής συμπερασμάτων καθώς

¹³ Αυτό σημαίνει ότι η OWL Lite παρέχει μικρότερες δυνατότητες έκφρασης από την OWL DL, και αντίστοιχα η OWL DL παρέχει μικρότερες δυνατότητες έκφρασης από την OWL Full.

¹⁴ Η περιγραφική λογική αποτελεί το σημαντικότερο επίσημο τρόπο αναπαράστασης γνώσης στην τεχνική νοημοσύνη, ο οποίος αποτελεί τη λογική βάση για την ανάπτυξη των σημασιολογικών δικτύων (Semantic Networks), των συστημάτων που βασίζονται σε πλαίσια (Frame-based systems), κτλ. [LUTZ07].

και να ελεγχθεί αυτόματα αν περιέχει ασυνέπειες. Όλα τα συμπεράσματα είναι σίγουρα υπολογίσιμα και όλοι οι υπολογισμοί ολοκληρώνονται σε πεπερασμένο χρονικό διάστημα. Παράδειγμα, ενώ μια κλάση μπορεί να είναι υποκλάση άλλων κλάσεων, δεν μπορεί να είναι στιγμιότυπο μιας άλλης κλάσης.

- **OWL-Full:** Είναι η πιο εκφραστική υπο-γλώσσα της OWL. Χρησιμοποιείται σε περιπτώσεις που απαιτείται μεγάλη εκφραστικότητα και ελευθερία σύνταξης του RDF, δίχως να εξασφαλίζονται υπολογιστικές εγγυήσεις, ενώ δεν είναι δυνατόν να γίνει αυτόματος συμπερασμός σε μια OWL-Full οντολογία. Παράδειγμα, μια κλάση στη OWL-Full μπορεί να αντιμετωπιστεί ταυτόχρονα ως μια συλλογή ατόμων και ως ένα αυτόνομο άτομο.

Καθεμιά από τις παραπάνω υπογλώσσες αποτελεί επέκταση του απλούστερου προκατόχου της, για το λόγο αυτό οι κατασκευαστές οντολογιών που υιοθετούν την OWL θα πρέπει να εξετάζουν ποια υπογλώσσα ικανοποιεί τις ανάγκες τους. Για παράδειγμα η επιλογή μεταξύ των OWL Lite και OWL DL εξαρτάται από το πόσο αναγκαίες είναι οι πιο εκφραστικές δομές που χρησιμοποιεί η OWL DL (π.χ. εύρος τιμών μεγαλύτερο του ένα). Ενώ η επιλογή ανάμεσα στην OWL DL και OWL Full εξαρτάται από την ανάγκη για χρήση των διευκολύνσεων όσον αφορά τη μοντελοποίηση που προσφέρει η OWL Full (π.χ. η δημιουργία κλάσεων από κλάσεις, ενσωμάτωση ιδιοτήτων σε κλάσεις).

Για τις τρεις αυτές υπογλώσσες της OWL ισχύουν οι παρακάτω προτάσεις:

- Κάθε νόμιμη OWL Lite οντολογία είναι και νόμιμη OWL DL οντολογία.
- Κάθε νόμιμη OWL DL οντολογία είναι και νόμιμη OWL Full οντολογία.
- Κάθε έγκυρο OWL Lite συμπέρασμα είναι και έγκυρο OWL DL συμπέρασμα.

- Κάθε έγκυρο OWL DL συμπέρασμα είναι και έγκυρο OWL Full συμπέρασμα.

Η OWL Full μπορούμε να πούμε ότι είναι μία επέκταση της RDF, ενώ οι OWL Lite και OWL DL επεκτάσεις μίας περιορισμένης όψης της RDF.

Για αυτό, ισχύουν επιπλέον τα παρακάτω:

- Κάθε OWL (Lite, DL, Full) έγγραφο είναι και RDF έγγραφο.
- Κάθε RDF έγγραφο είναι και OWL Full έγγραφο.
- Μόνο ορισμένα RDF έγγραφα μπορούν να είναι και OWL Lite ή OWL DL έγγραφα. (Mc Guinness et al., 2004) .

2.8.2 Βασικά δομικά στοιχεία της OWL

Πως ορίζεται η δομή της owl με βάση τις εκάστοτε ετικέτες που μπορεί να χρησιμοποιήσει ο κατασκευαστής οντολογιών ανάλογα με τη μηχανή αναζήτησης που θέλει να δημιουργήσει:

1. Στοιχεία της OWL Lite που προέρχονται από την RDFS (Mc Guinness et al., 2004) , (Smith et al., 2004) :
 - owl: Class: Μία κλάση (Class) ορίζει μία ομάδα στιγμιότυπων (αντικειμένων) που μοιράζονται κάποιες κοινές ιδιότητες. Οι κλάσεις μπορούν να οργανωθούν σε μία ιεραρχία χρησιμοποιώντας την ιδιότητα rdfs:subClassOf. Υπάρχει μία προκαθορισμένη κλάση, η owl:Thing, που περιέχει όλα τα στιγμιότυπα και είναι υπερκλάση όλων των κλάσεων και η κλάση owl:Nothing, που δεν περιέχει κανένα στιγμιότυπο και είναι υποκλάση όλων των κλάσεων της OWL.
 - rdfs:subClassOf: Ιεραρχίες κλάσεων μπορούν να δημιουργηθούν με τη δήλωση ότι μία κλάση είναι υποκλάση (subClassOf) μίας άλλης κλάσης. Η ίδια ακριβώς ιδιότητα λειτουργεί και στην RDFS.

- `rdfs:Property`: Μία ιδιότητα (`Property`) δηλώνει μία σχέση μεταξύ στιγμιότυπων ή μία σχέση ανάμεσα σε στιγμιότυπο και μία τιμή δεδομένου. Οι ιδιότητες της πρώτης κατηγορίας ονομάζονται ιδιότητες αντικειμένων (`owl:ObjectProperty`), ενώ της δεύτερης κατηγορίας ιδιότητες τύπων δεδομένων (`owl:DatatypeProperty`). Και οι δύο είναι υποκλάσεις της `rdf:Property`
 - `rdfs:subPropertyOf`: Ιεραρχίες ιδιοτήτων μπορούν να δημιουργηθούν με τη δήλωση ότι μία ιδιότητα είναι υποιδιότητα (`subPropertyOf`) μίας άλλης ιδιότητας. Η ιδιότητα αυτή λειτουργεί όπως ακριβώς λειτουργεί και στην RDFS.
 - `rdfs:domain`: Το πεδίο ορισμού (`domain`) μίας ιδιότητας περιορίζει τα αντικείμενα στα οποία μπορεί να εφαρμοστεί η ιδιότητα. Αν μία ιδιότητα δηλώνει μία σχέση ενός στιγμιότυπου με ένα άλλο κα έχει ως πεδίο ορισμού μία κλάση, τότε το πρώτο στιγμιότυπο πρέπει να ανήκει σε αυτήν την κλάση. Το πεδίο ορισμού είναι ένας γενικευμένος περιορισμός, αφού χαρακτηρίζει μία ιδιότητα γενικά και όχι μία ιδιότητα που συνδέεται με μία συγκεκριμένη κλάση. Η ιδιότητα αυτή λειτουργεί όπως ακριβώς λειτουργεί και στην RDFS.
 - `rdfs:range`: Το πεδίο τιμών (`range`) μίας ιδιότητας περιορίζει τα αντικείμενα τα οποία μπορούν να αποτελούν τιμή της ιδιότητας. Αν μία ιδιότητα δηλώνει μία σχέση ενός στιγμιότυπου με ένα άλλο κα έχει ως πεδίο τιμών μία κλάση, τότε το δεύτερο στιγμιότυπο πρέπει να ανήκει σε αυτήν την κλάση. Το πεδίο τιμών είναι επίσης γενικευμένος περιορισμός.
 - `Owl:Individual`: Το άτομο (`Individual`) είναι ένα στιγμιότυπο μίας κλάσης. Οι ιδιότητες μπορούν να χρησιμοποιηθούν για να συσχετίσουν ένα άτομο με κάποιο άλλο.
2. Στοιχεία της OWL Lite που σχετίζονται με ισότητα και ανισότητα (Mc Guinness et al., 2004) , (Smith et al., 2004):

- `equivalentClass`: Ισοδύναμες (`equivalentClass`) μπορούν να δηλωθούν δύο κλάσεις που έχουν τα ίδια στιγμιότυπα. Με αυτόν τον τρόπο δημιουργούνται συνώνυμες κλάσεις. π.χ. μπορούν να οριστούν ως ισοδύναμες κλάσεις `car` και `automobile`, αυτό δίνει τη δυνατότητα σ' ένα μηχανισμό συλλογής να συμπεράνει ότι κάθε άτομο που είναι στιγμιότυπο της κλάσης `car` να είναι και στιγμιότυπο της κλάσης `automobile` και το αντίστροφο.
- `equivalentProperty`: Ισοδύναμες (`equivalentProperty`) μπορούν να δηλωθούν δύο ιδιότητες που συνδέουν ένα στιγμιότυπο με την ίδια ομάδα στιγμιότυπων. Με αυτόν τον τρόπο δημιουργούνται συνώνυμες ιδιότητες. π.χ. μπορούν να οριστούν ως ισοδύναμες οι ιδιότητες `studiesC++` και `isProgrammer`. Από αυτή τη σχέση ο μηχανισμός μπορεί να συμπεράνει ότι αν το `X` σχετίζεται με το `Y` από την ιδιότητα `studiesC++`, τότε το `X` σχετίζεται με το `Y` και από την ιδιότητα `isProgrammer`, και αντίστροφα.
- `sameAs`: Όταν υπάρχουν διαφορετικά ονόματα που αναφέρονται ουσιαστικά στο ίδιο άτομο, πρέπει να χρησιμοποιείται η συγκεκριμένη ιδιότητα, με σκοπό να δημιουργηθούν περισσότερα από ένα όνομα αναφοράς στο ίδιο στιγμιότυπο. π.χ. το άτομο `trash` μπορεί να οριστεί ως όμοιο με το άτομο `wastebasket`.
- `differentFrom`: Ένα στιγμιότυπο μπορεί να χαρακτηριστεί διαφορετικό (`differentFrom`) από ένα άλλο στιγμιότυπο για να δηλωθεί ρητά ότι τα δύο στιγμιότυπα δεν είναι ίδια. Αυτό είναι πολύ σημαντικό σε γλώσσες όπως η OWL, οι οποίες δεν υποθέτουν ότι κάθε στιγμιότυπο έχει ένα και μόνο όνομα. Ουσιαστικά το αποτέλεσμα που προκύπτει είναι το αντίθετο της ιδιότητας `sameAs`. π.χ. το άτομο `Μαρίνα` μπορεί να δηλωθεί ως διαφορετικό από τα άτομα `Μαρία` και `Φυλιώ`. Αν σε αυτή τη περίπτωση τα άτομα `Μαρίνα` και `Μαρία` δηλωθούν ως τιμές μιας ιδιότητας υπάρχει αντίφαση καθώς έχει δηλωθεί ρητά ότι τα δύο άτομα διαφέρουν.
- `AllDifferent`: Ένα σύνολο από στιγμιότυπα μπορεί να δηλωθεί `AllDifferent`, δηλαδή ότι κάθε στιγμιότυπο είναι διαφορετικό από τα υπόλοιπα. Αυτό είναι

πολύ σημαντικό όταν απαιτείται από ένα σύνολο στιγμιότυπων να έχει το καθένα ένα και μόνο όνομα. π.χ. έστω ότι τα άτομα, Μαρίνα, Μαρία, Φυλιό δηλώνονται ως ξεχωριστά άτομα, σε αυτή τη περίπτωση κανένα από τα άτομα δεν μπορεί να χρησιμοποιηθεί μαζί με κάποιο άλλο άτομο ως ίδια ιδιότητα σε αντίθεση με την ιδιότητα `differentFrom`.

3. Στοιχεία της OWL Lite που σχετίζονται με τις ιδιότητες και τις τιμές τους (Mc Guinness et al., 2004) , (Smith et al., 2004):

- `inverseOf`: Μία ιδιότητα μπορεί δηλωθεί ως η αντίστροφη (`inverseOf`) μίας άλλης ιδιότητας. Σε αυτήν την περίπτωση αν το στιγμιότυπο X συνδέεται μέσω μίας ιδιότητας με το στιγμιότυπο Y , τότε το Y συνδέεται με το X μέσω της αντίστροφης ιδιότητας. Παράδειγμα αντίστροφης σχέσης είναι η σχέση “`hasChild`” με τη σχέση “`hasParent`”.
- `TransitiveProperty`: Μία ιδιότητα μπορεί δηλωθεί ως η μεταβατική (`TransitiveProperty`). Σε αυτήν την περίπτωση αν τα ζεύγη (X, Y) και (Y, Z) είναι στιγμιότυπα μίας μεταβατικής ιδιότητας, τότε και το ζεύγος (X, Z) είναι στιγμιότυπο της ίδιας ιδιότητας. Οι OWL Lite και OWL DL επιβάλλουν τον περιορισμό ότι οι μεταβατικές ιδιότητες (και οι υπεριδιότητές τους) δεν μπορούν να έχουν περιορισμό μέγιστης πληθικότητας 1. Σε διαφορετική περίπτωση δεν θα ίσχυε η αποκρισιμότητα της γλώσσας. Παράδειγμα μεταβατικής ιδιότητας μπορεί να είναι η σχέση “`has Ancestor`” αφού αν ο A είναι πρόγονος του B και ο B είναι πρόγονος του Γ , τότε ο A είναι πρόγονος του Γ .
- `SymmetricProperty`: Μία ιδιότητα μπορεί δηλωθεί ως η συμμετρική (`SymmetricProperty`). Σε αυτήν την περίπτωση αν το ζεύγος (X, Y) είναι στιγμιότυπο μίας συμμετρικής ιδιότητας, τότε και το ζεύγος (Z, X) είναι στιγμιότυπο της ίδιας ιδιότητας. Παράδειγμα συμμετρικής σχέσης μπορεί να θεωρηθεί η σχέση “`hasCousin`” αφού αν ο A έχει ξάδερφο τον B , τότε και ο B έχει ξάδερφο τον A .

- **FunctionalProperty:** Μία ιδιότητα μπορεί δηλωθεί ως συναρτησιακή (FunctionalProperty) στην περίπτωση που δεν έχει παραπάνω από μία τιμή για κάθε στιγμιότυπο, στην ιδιότητα αυτή υπάγονται και τα άτομα τα οποία δεν ορίζονται με καμία τιμή. Στην περίπτωση αυτή μια σχέση είναι συναρτησιακή για ένα στιγμιότυπο y , τότε το πολύ ένα στιγμιότυπο μπορεί να σχετίζεται με το στιγμιότυπο y μέσω της συγκεκριμένης ιδιότητας. π.χ. συναρτησιακή σχέση αποτελεί η σχέση “hasMother”, αφού οποιοσδήποτε άνθρωπος έχει μια μόνο μητέρα.
 - **InverseFunctionalProperty:** Μία ιδιότητα μπορεί δηλωθεί ως η αντιστρόφως συναρτησιακή (InverseFunctionalProperty) στην περίπτωση που η αντίστροφη της είναι συναρτησιακή. Ένα παράδειγμα αντίστροφα συναρτησιακής σχέσης είναι η σχέση “isMotherOf” αφού η αντίστροφη σχέση είναι συναρτησιακή “hasMother”.
4. Περιορισμοί ιδιοτήτων της OWL Lite που καθορίζουν πως μια ιδιότητα μπορεί να χρησιμοποιηθεί από τα στιγμιότυπα μιας κλάσης (Mc Guinness et al., 2004) , (Smith et al., 2004) , (Horridge et al., 2004):
- **allValuesFrom:** Ο περιορισμός allValuesFrom (όλες οι τιμές από) τίθεται σε μία ιδιότητα, αναφερόμενος σε μία κλάση. Αυτό σημαίνει ότι η ιδιότητα, για τη συγκεκριμένη κλάση, έχει ένα περιορισμένο πεδίο τιμών. Έτσι, αν ένα στιγμιότυπο X της κλάσης συνδέεται με τη συγκεκριμένη ιδιότητα με ένα στιγμιότυπο Y , τότε το Y ανήκει στην κλάση που προσδιορίζεται από τον αντίστοιχο περιορισμό του πεδίου τιμών.
 - **someValuesFrom:** Ο περιορισμός allValuesFrom (όλες οι τιμές από) τίθεται σε μία ιδιότητα, αναφερόμενος σε μία κλάση. Αυτό σημαίνει ότι η ιδιότητα, για τη συγκεκριμένη κλάση, για μία τουλάχιστον από τις τιμές της, έχει ένα περιορισμένο πεδίο τιμών.

5. Περιορισμοί εύρους τιμών της OWL Lite (Mc Guinness et al., 2004) , (Smith et al., 2004):

- **minCardinality**: Ο περιορισμός ελάχιστης πληθικότητας (**minCardinality**) τίθεται σε μία ιδιότητα, αναφερόμενος σε μία κλάση. Αν για μία ιδιότητα έχουμε **minCardinality 1**, τότε κάθε στιγμιότυπο της κλάσης θα πρέπει να συνδέεται τουλάχιστον με ένα άλλο στιγμιότυπο, μέσω αυτής της ιδιότητας. Με άλλα λόγια, η ιδιότητα θα πρέπει να έχει τουλάχιστον μία τιμή για κάθε στιγμιότυπο της κλάσης. Στην OWL Lite οι τιμές ελάχιστης πληθικότητας που επιτρέπονται είναι 0 και 1. Τιμή 0 σημαίνει ότι η ιδιότητα είναι προαιρετική για τη συγκεκριμένη κλάση. π.χ. η κλάση **person** δεν πρέπει να έχει περιορισμούς επειδή δεν έχουν υποχρεωτικά όλοι οι άνθρωποι απογόνους, ενώ η κλάση **parent** πρέπει να έχει ελάχιστο εύρος ίσο με ένα καθώς κάποιος γονέας έχει τουλάχιστον ένα απόγονο.
- **maxCardinality**: Ο περιορισμός μέγιστης πληθικότητας (**maxCardinality**) τίθεται σε μία ιδιότητα, αναφερόμενος σε μία κλάση. Αν για μία ιδιότητα έχουμε **maxCardinality 1**, τότε κάθε στιγμιότυπο της κλάσης θα πρέπει να συνδέεται το πολύ με ένα άλλο στιγμιότυπο, μέσω αυτής της ιδιότητας. Αν έχουμε **maxCardinality 0**, τότε υποδηλώνουμε ότι κάποια κλάση δεν έχει τιμές για τη συγκεκριμένη ιδιότητα. π.χ. όταν η ιδιότητα **ψήφος** σχετίζεται με την κλάση **Ευρωπαίος πολίτης** είναι σωστό να έχει μέγιστη τιμή εύρους ίση με ένα, καθώς ένας Ευρωπαίος πολίτης έχει δικαίωμα να ψηφίζει μόνο σε μια χώρα.
- **cardinality**: Ο περιορισμός πληθικότητας (**maxCardinality**) τίθεται σε μία ιδιότητα, αναφερόμενος σε μία κλάση και καθορίζει την ακριβή πληθικότητα της ιδιότητας για αυτήν την κλάση έχει δηλαδή ορίσει και στους δύο περιορισμούς εύρους τιμών την τιμή μηδέν ή την τιμή ένα (**minCardinality** και **maxCardinality**). π.χ. η κλάση **person** είναι σωστό να δέχεται μία τιμή στην ιδιότητα **hasBirthMother** καθώς κάθε άνθρωπος έχει μόνο μια βιολογική μητέρα.

Επίσης στοιχείο της OWL Lite είναι η τομή (intersectionOf) της οποίας η χρήση περιορίζεται σε επώνυμες κλάσεις.

2.8.3 Στοιχεία που επεκτείνουν την OWL Lite σε OWL DL και OWL Full

Οι γλώσσες OWL DL και OWL Full χρησιμοποιούν το ίδιο λεξιλόγιο, αν και η OWL DL υπόκειται σε κάποιους περιορισμούς. π.χ. η OWL DL απαιτεί διαχωρισμό τύπων (μία κλάση δεν μπορεί να είναι επιπλέον και άτομο ή ιδιότητα, και μία ιδιότητα δεν μπορεί να είναι επιπλέον και άτομο ή κλάση). Επιπρόσθετα, η OWL DL απαιτεί οι ιδιότητες να είναι είτε ιδιότητες τύπων δεδομένων (datatype properties) είτε ιδιότητες αντικειμένων (object properties), κάτι το οποίο δεν ισχύει στην OWL Full. Οι διαφορές μεταξύ των λεξιλογίων των δύο γλωσσών δημιουργούν ένα κοινό λεξιλόγιο που επεκτείνει τα χαρακτηριστικά της OWL Lite προσθέτοντας τα παρακάτω στοιχεία (Mc Guinness et al., 2004) , (Smith et al., 2004):

- oneOf: Οι κλάσεις μπορούν να περιγραφούν ως απαρίθμηση των στιγμιότυπων από τα οποία αποτελούνται. Τα μέλη της κλάσης είναι ακριβώς τα απαριθμημένα στιγμιότυπα και μόνο αυτά. π.χ. η κλάση DaysOfTheWeek μπορεί να περιγραφεί απαριθμώντας τα άτομα Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday απ' αυτό το χαρακτηριστικό ο μηχανισμός μπορεί να συμπεράνει πως το μέγιστο εύρος τιμών είναι επτά.
- hasValue: Μία ιδιότητα ενδέχεται να απαιτείται να έχει ως τιμή ένα συγκεκριμένο στιγμιότυπο. π.χ. τα στιγμιότυπα της κλάσης DutchCitizens μπορούν να χαρακτηριστούν ως άνθρωποι που έχουν εθνικότητα το στιγμιότυπο theNetherlands, όπου theNetherlands είναι μία σειρά του στιγμιότυπου της κλάσης Nationalities.

- **disjointWith**: Κάποιες κλάσεις μπορούν να δηλωθούν ότι είναι ξένες μεταξύ τους (**disjointWith**), δηλαδή ότι δεν έχουν ούτε ένα κοινό στοιχείο. Έτσι, αν κάποιο στιγμιότυπο ανήκει σε μία κλάση, αποκλείεται να ανήκει και σε μία ξένη προς αυτήν κλάση. π.χ. έστω οι κλάσεις **Man** και **Woman** έχουν δηλωθεί ως διαχωριστές δηλαδή ξένες μεταξύ τους, εφόσον γίνει αυτή η δήλωση ο μηχανισμός αυτόματα μπορεί να συμπεράνει ότι υπάρχει ασυνέπεια όταν ένα άτομο δηλωθεί ως στιγμιότυπο και των δύο κλάσεων. Επίσης ο μηχανισμός μπορεί να συμπεράνει ότι αν το **A** είναι στιγμιότυπο της κλάσης **Man**, τότε το **A** δεν είναι στιγμιότυπο της κλάσης **Woman**.
- **unionOf**, **complementOf**, **intersectionOf** : Σε **OWL DL** και **OWL Full** η ένωση, το συμπλήρωμα και η τομή αντίστοιχα επιτρέπουν διάφορους συνδυασμούς κλάσεων και περιορισμών. π.χ. η χρήση του τελεστή **unionOf** δηλώνει ότι ένα φρούτο(**Fruit**) μπορεί να έχει είτε γλυκιά γεύση(**SweetFruit**) είτε πικρή γεύση(**NonSweetFruit**). Ο τελεστής **complementOf** διαχωρίζει τα αγαθά που χρειάζονται συχνά ανανέωση(**ConsumableThing**), από όλα τα υπόλοιπα, τα οποία προκύπτουν και αποτελούν το συμπλήρωμα της κλάσης **ConsumableThing**.
- **minCardinality**, **maxCardinality**, **cardinality** : Οι περιορισμοί πληθικότητας επιτρέπεται να παίρνουν τιμές, όχι μόνο 0 και 1, αλλά οποιονδήποτε φυσικό αριθμό. π.χ. η ιδιότητα **team** θα έχει ελάχιστο εύρος ίσο με τρία ,καθώς μια ομάδα αποτελείται από τουλάχιστον τρία μέλη, αφού τα δύο άτομα θεωρούνται ζεύγος.
- **complex classes** : Σε πολλές δομές, η **OWL Lite** περιορίζει τη σύνταξη σε μοναδικά ονόματα κλάσεων. Η **OWL Full** επεκτείνει αυτόν τον περιορισμό στο να επιτρέπει σύνθετες περιγραφές κλάσεων, όπως απαριθμημένες κλάσεις, περιορισμούς ιδιοτήτων και συνδυασμούς λογικών εκφράσεων. Επίσης, η **OWL Full** επιτρέπει σε κλάσεις να χρησιμοποιούνται ως στιγμιότυπα.

2.9 Η γλώσσα OIL (Ontology Interchange Language- Στρώμα Διεπαφής Οντολογίας)

Άλλη μια γλώσσα περιγραφής οντολογιών η οποία είναι εκφραστική αλλά όχι τόσο δημοφιλής όπως η OWL είναι η Ontology Interchange Language(OIL) η οποία αναπτύχθηκε με στόχο τη σημασιολογική διαλειτουργικότητα μεταξύ των διαδικτυακών πόρων.

Η OIL ανήκει στις web – based γλώσσες που σημαίνει ότι το συντακτικό και η σημασιολογία της γλώσσας βασίζονται κυρίως στη βασική μορφή της γλώσσας προγραμματισμού της XML καθώς και σε υπάρχουσες προτάσεις όπως η OKBC¹⁵, XOL¹⁶, RDF/S¹⁷.

Η τυπική σημασιολογία της OIL και η δυνατότητα διεξαγωγής με αυτήν προέρχονται από τις Λογικές Περιγραφές¹⁸. Η OIL παρουσιάζει μια στρωματική προσέγγιση για μια τυποποιημένη γλώσσα οντολογίας, αυτό μπορεί να γίνει άμεσα αντιληπτό στο άκουσμα και μόνο του ονόματος της OIL, Ontology Interchange Language που ερμηνεύεται ως Στρώμα Διεπαφής Οντολογίας. Έτσι κάθε πρόσθετο στρώμα προσθέτει λειτουργικότητα και πολυπλοκότητα στο προηγούμενο βασισμένο στη λογική της διαστρωμάτωσης.

2.9.1 Βασική δομή διαστρωμάτωσης της OIL

¹⁵ Open Knowledge Base Connectivity

¹⁶ XML- based Ontology exchange Language

¹⁷ Resource Description Framework Schema

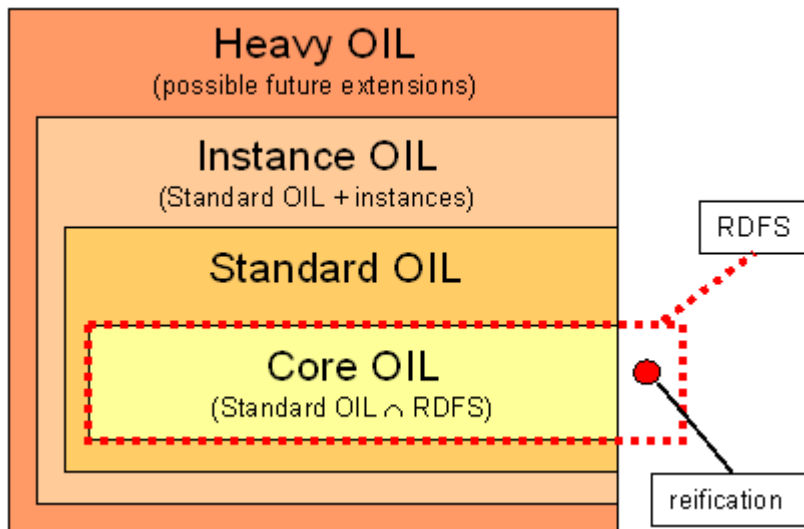
¹⁸ Ο όρος αφορά σε πιο περίπλοκα μοντέλα τα οποία περιγράφονται από εκφραστικότερες γλώσσες όπως OIL, DAML+ OIL, OWL κ.λ.π

Στην ενότητα αυτή παρουσιάζεται το μοντέλο διαστρωμάτωσης της OIL που ελέγχει την πολυπλοκότητα των διαδικτυακών πόρων, έτσι:

Στο πρώτο επίπεδο του μοντέλου βρίσκεται η Καθορισμένη γλώσσα OIL(Standard OIL) που περιλαμβάνει την τομή της Απλής γλώσσας της OIL(Core OIL) με τις ιδιότητες της RDFS. Δηλαδή το RDF Schema καθορίζει τις κλάσεις και τις ιδιότητες που μπορούν να χρησιμοποιηθούν για να περιγράψουν τις κλάσεις, τις ιδιότητες και άλλες πηγές έτσι ώστε να αποδώσουν την λειτουργικότητα αυτών των πόρων σχηματίζοντας την Καθορισμένη γλώσσα της OIL.

Στο δεύτερο επίπεδο αναπτύσσεται η Πρόταση της OIL(Instance OIL) η οποία καθορίζεται από την Standard OIL και παραδείγματα αυτής (instances) προσθέτοντας λειτουργικότητα στους πόρους που την χρησιμοποιούν και ελέγχοντας παράλληλα και την πολυπλοκότητα τους.

Τέλος στο τρίτο επίπεδο και στην κορυφή του μοντέλου βρίσκεται η Σκληρή γλώσσα της OIL (Heavy OIL), η οποία δεν περιορίζετε και πιθανόν στο μέλλον να επεκταθεί (possible future extensions). Η Heavy OIL είναι αρκετά εκφραστική, καθώς έχει σχηματιστεί και αναπτυχθεί βασισμένη στα στρώματα της OIL γλώσσας που βρίσκονται κάτω απ' αυτή.



Εικόνα 2.2 Μοντέλο της OIL που ελέγχει την πολυπλοκότητα

2.10 Η DAML-OIL γλώσσα (Darpa Agent Markup Language)

Η DAML ουσιαστικά αποτελεί προσχέδιο γλώσσας γνωστό ως προτάσεις Οντολογίας της Γλώσσας Σήμανσης Πρακτόρων DARPA(DAML-ONT) και αναπτύχθηκε από Ευρωπαίους ερευνητές με την υποστήριξη της Ευρωπαϊκής Επιτροπής ως ερευνητική προσπάθεια ανάπτυξης γλωσσών που αποδίδουν οντολογικές εκφράσεις στον Παγκόσμιο ιστό.

Έκτοτε όμως ειδική ομάδα ερευνητών σχημάτισε την κοινή επιτροπή για γλώσσες σήμανσης πρακτόρων όπως είναι η Οντολογία της Γλώσσας Σήμανσης Πρακτόρων (DARPA) και αυτή του Στρώματος Διεπαφής της Οντολογίας (OIL) δίνοντας σαν αποτέλεσμα μια νέα έκδοση της γλώσσας Σήμανσης Πρακτόρων που συγχωνεύει την DAML και OIL.

Η υβριδική ακόμη DAML+OIL αποτελεί έκφραση μιας νέας γλώσσας που επιτρέπει την αναπαράσταση οντολογιών βασιζόμενη όμως στα χαρακτηριστικά ήδη υπάρχουσών γλωσσών όπως αυτή της RDF και άλλων

γλωσσών περιγραφής τεχνητής νοημοσύνης και όπως κάθε οντολογική γλώσσα σχεδιάστηκε για να περιγράψει τη δομή (structure) ενός πεδίου γνώσης διατηρώντας αντικειμενική προσέγγιση. Συνεπώς η δομή της περιγράφεται με τάξεις και ιδιότητες. Μια DAML+OIL οντολογία περιέχει ένα σύνολο από αξιώματα , τα οποία υποστηρίζουν χαρακτηριστικά των τάξεων και των ιδιοτήτων αυτών.

Ο δικτυακός τύπος για την γλώσσα DAML (DARPA Agent Markup Language) περιέχει πληθώρα εγγράφων, εργαλείων, παραδειγμάτων και οντολογιών για το Σημασιολογικό Διαδίκτυο και αποτελεί πρόδρομό της OWL. Ουσιαστικά η γλώσσα DAML έχει αναπτυχθεί για την περιγραφή των οντολογιών, ενώ για την υποδομή των οντολογιών για το σημασιολογικό ιστό έχει αναπτυχθεί η OIL γλώσσα. Συνεπώς η συγχώνευση των DAML+OIL γλωσσών προσφέρει εκφραστικότητα, καλή καθορισμένη σημασιολογία καθώς επίσης και την ικανότητα αντιστοίχισης, επαναχρησιμοποίησης και υποστήριξης των οντολογιών. Τέλος προσφέρει την δυνατότητα αυτομάτων web services και χρήση από agent καθώς επίσης και την δυνατότητα παραγωγής συμπερασμάτων σε web services. (W3C, 2004)

2.10.1 Που συναντάμε την DAML+OIL

Η συγχώνευση των DAML+OIL γλωσσών όπως έχει αναφερθεί ήδη έχει ως στόχο την ανάπτυξη μιας γλώσσας και ορισμένων εργαλείων που διευκολύνουν την έννοια του Σημασιολογικού Ιστού και βασισμένο στο ότι είναι μέρος του DARPA Agent Markup Language(DAML) Program ,έχει αναπτυχθεί η βιβλιοθήκη οντολογιών DAML. Η βιβλιοθήκη οντολογιών περιλαμβάνει έναν κατάλογο ο οποίος είναι διαθέσιμος σε XML, HTML και DAML μορφές. Η HTML μορφή διευκολύνει τους χρήστες στην αναζήτηση και δημιουργία ευρετηρίων οντολογιών σύμφωνα με το URI, την ημερομηνία

καταχώρησης κλπ. Οι άλλες δύο μορφές υποστηρίζουν απλό φυλλομέτρημα και έτσι οι χρήστες έχουν την δυνατότητα να υποβάλουν νέες οντολογίες μέσω της βιβλιοθήκης.

Η βιβλιοθήκη οντολογιών DAML βασίζεται στην αρχιτεκτονική πελάτη – εξυπηρετητή και έτσι οι οντολογίες ταξινομούνται σύμφωνα με το Open Directory Category (www.dmoz.org) και προσδιορίζονται μοναδικά σύμφωνα με τους αναγνωριστές του και το URI. Ωστόσο το σύστημα δε διαθέτει λειτουργίες υποστήριξης πολλαπλών εκδόσεων, επεξεργασίας και συλλογικής υποστήριξης καθώς και δεν υπάρχουν δυνατότητες αναζήτησης.

2.11 γλώσσα XOL (Ontology eXchange Language)

Άλλη μια γλώσσα web based που βασίζεται στη σύνταξη της XML είναι η Ontology eXchange Language(XOL), Οντολογία Ανταλλαγής Γλώσσας.

Η XOL δημιουργήθηκε με σκοπό την γρήγορη ανταλλαγή δεδομένων στον τομέα της Βιοπληροφορικής, ωστόσο σήμερα μπορεί να χρησιμοποιηθεί σε οντολογίες για κάθε τομέα. Σχεδιάστηκε ως απάντηση σε μια μελέτη των γλωσσών των νόσων που εκτελούνται από την BioOntology Core Group. Η μελέτη διαπίστωσε ότι μέχρι σήμερα δεν υπήρχε μια γλώσσα ανταλλαγής οντολογιών και η XOL πληρούσε τις απαιτήσεις της κοινότητας της Βιοπληροφορικής. Επιπλέον η ομάδα τόνισε την ανάγκη για δημιουργία μιας γλώσσας του σημασιολογικού ιστού που ως αντικείμενο γνώσης θα λειτουργεί με τον προσανατολισμό και την εκπροσώπηση συστημάτων βασισμένη στη σύνταξη XML γλώσσας.

Η XOL γλώσσα προσφέρει ένα μηχανισμό για την κωδικοποίηση των οντολογιών μέσα σ' ένα απλό αρχείο που μπορεί να εύκολα να συγγράψει κείμενο για το WWW για ανταλλαγή εφαρμογών μεταξύ ενός συνόλου

ανάπτυξης. Η γλώσσα είναι σχεδιασμένη να είναι αναγνώσιμη από τον άνθρωπο(σε σχέση με την XML γλώσσα η οποία δεν είναι πλήρη κατανοητή από τον άνθρωπο) και να είναι εύκολα προσπελάσιμη από προγράμματα μέτριας πολυπλοκότητας.

Η XOL γλώσσα είναι αρκετά εκφραστική και έχοντας αυτό σαν βασικό της πλεονέκτημα συλλαμβάνει μια πλούσια ποικιλία οντολογιών. Για το λόγο τούτο η XOL προορίζεται να χρησιμοποιηθεί ως ενδιάμεση γλώσσα για οντολογίες μεταφοράς δεδομένων μεταξύ των διαφορετικών συστημάτων, οντολογία- εργαλεία ανάπτυξης ή εφαρμογής προγραμμάτων, καθώς αποτελεί γλώσσα ανταλλαγής των οντολογιών. (XOL, 2008)

2.12 Η SHOE γλώσσα (Simple Html Ontology Extensions)

Η SHOE αναπτύχθηκε από το πανεπιστήμιο του Maryland(USA) και είναι η πρώτη γλώσσα για τη σημασιολογία του Ιστού που αναπτύχθηκε ως γλώσσα σήμανσης και είναι αποτέλεσμα των προσπαθειών τυποποίησης που βασίζεται στην W3C(Word Wide Web Consortium).

Η SHOE είναι μια μικρή επέκταση σε HTML ιστοσελίδα, η οποία επιτρέπει τους δημιουργούς να εκφράζουν μια ιστοσελίδα με έγγραφα από μηχανήματα γνώσης(machine readable knowledge), ουσιαστικά η SHOE δημιουργεί ευφυής πράκτορες λογισμικού μέσα στο διαδίκτυο.

Βασιζόμενη στην HTML, που η κύρια λειτουργία της είναι η απεικόνιση δεδομένων αναγνώσιμα από άνθρωπο, δηλαδή δημιουργία ιστοσελίδας που καθορίζεται με διάφορα στοιχεία όπως πίνακες, γραφικά και πλαίσια που μπορούν μόνο οι άνθρωποι οπτικά να κατανοήσουν. Ωστόσο οι ευφυείς πράκτορες λογισμικού δεν μπορούν να τα αντιληφθούν. Για το λόγο τούτο κρίθηκε αναγκαίο η δημιουργία μιας νέας γλώσσας(SHOE) επεκτάσιμης της

HTML που παρέχει την δυνατότητα στους ευφυείς πράκτορες να διαβάσουν και να κατανοήσουν web έγγραφα. (SHOE, 2008)

2.12.1 Που βρίσκει εφαρμογές η SHOE γλώσσα

Έχοντας σαν βασικό της όπλο τη δυνατότητα ευφύων πρακτόρων λογισμικού μέσα στο διαδίκτυο, η SHOE δημιούργησε την δική της βιβλιοθήκη οντολογιών.

Η βιβλιοθήκη της SHOE περιλαμβάνει λίστες οντολογιών οι οποίες βρίσκονται στο ευρετήριο αλφαβητικά και ταξινομούνται ως προς την εξάρτηση της οντολογίας, καθώς κάθε οντολογία αναγνωρίζεται μοναδικά από το όνομα της. Δεν περιέχει περιβάλλον αναζήτησης και επεξεργασίας όπως άλλες βιβλιοθήκες καθώς δίνει την δυνατότητα στους χρήστες να επεξεργαστούν σε κάποιο άλλο περιβάλλον την οντολογία και έπειτα να την καταχωρήσουν στη SHOE. Εν συνεχεία, μπορεί να αναγνωρίσει ή να αναζητήσει οντολογίες μέσω του αλφαβητικού ευρετηρίου. Τέλος η SHOE παρέχει στο σύστημα υποστήριξη για διαδικασίες εξαγωγής συμπερασμάτων για το χειρισμό των προβλημάτων των αναθεωρήσεων.

2.13 Άλλες γλώσσες βιβλιοθηκών οντολογιών

- KIF: Το πρότυπο KIF (Knowledge Interchange format) αποτελεί μία γλώσσα περιγραφής που βασίζεται στη γλώσσα προγραμματισμού LISP (Genesereth and Fikes 1992). Στον ορισμό της γλώσσας περιλαμβάνονται τόσο ο προσδιορισμός του συντακτικού όσο και της σημασιολογίας των δεσμευμένων λέξεων της. Επιτρέπει την περιγραφή απλών ή περιπλοκών

δεδομένων, λογικών εκφράσεων, γνώσης, μεταγνώσης και διαδικασιών. Το πρότυπο KIF χρησιμοποιείται σε συνδυασμό με την γλώσσα KQML. Η γλώσσα KIF κωδικοποιεί τις οντολογίες στη βιβλιοθήκη οντολογιών OntoLingua, που αναπτύχθηκε στις αρχές της δεκαετίας του '90 στο Knowledge Systems Laboratory του Stanford University. Αποτελείται από έναν εξυπηρετητή και μια γλώσσα αναπαράστασης. Ο εξυπηρετητής παρέχει ένα αποθετήριο οντολογιών (βιβλιοθήκη οντολογιών), προκειμένου να βοηθήσει τους χρήστες να δημιουργούν νέες οντολογίες και να τροποποιούν και να βελτιώνουν τις ήδη υπάρχουσες συνεργατικά. Η οντολογία που αποθηκεύεται στον εξυπηρετητή είναι δυνατόν να μετατραπεί σε διαφορετικές διατάξεις. Βασίζεται στην αρχιτεκτονική πελάτη-εξυπηρετητή. Οι οντολογίες που αποθηκεύονται στο OntoLingua δεν ταξινομούνται σύμφωνα με κάποιες πρότυπες κατηγορίες αλλά οργανώνονται με βάση τη θεωρία του πλέγματος (lattice theory), κατά την οποία κάθε οντολογία συμπεριλαμβάνει τις οντολογίες εκείνες, κάτω από τις οποίες βρίσκεται αυτή τοποθετημένη με εσοχή. Κάθε οντολογία έχει ένα μοναδικό όνομα που τη διακρίνει από τις άλλες οντολογίες. Το Ontolingua δεν παρέχει τη δυνατότητα υποστήριξης πολλαπλών εκδόσεων. Το σύστημα παρέχει ανάγνωση οντολογιών και υποστηρίζει γρήγορες μεταβάσεις από τον ένα όρο μιας οντολογίας σε έναν άλλο όρο χρησιμοποιώντας υπερσυνδέσεις. Η OntoLingua παρέχει τέσσερα βασικά είδη σελίδων για την απλή διεπαφή: τον πίνακα περιεχομένων, τις σελίδες σύνοψης της οντολογίας, τις σελίδες πλαισίου (για κλάσεις, σχέσεις ή στιγμιότυπα) και το φυλλομετρητή κλάσεων.

- OCML: αποτελεί γλώσσα αναπαράστασης η οποία υποστηρίζει εξαγωγή συμπερασμάτων βάσει κανόνων και είναι υπεύθυνη για τη δημιουργία της WebOnto βιβλιοθήκης οντολογιών. Το WebOnto είναι μια βιβλιοθήκη οντολογιών που αναπτύχθηκε από το Knowledge Media Institute του

ανοικτού πανεπιστημίου της Αγγλίας. Υποστηρίζει συνεργατική περιήγηση, δημιουργία και επεξεργασία οντολογιών με χρήση της γλώσσας μοντελοποίησης γνώσης OCML και βασίζεται σε μια αρχιτεκτονική πελάτη-εξυπηρετητή. Οι εξυπηρετητές είναι υπεύθυνοι για την αποθήκευση και διατήρηση των οντολογιών και των διαλόγων του χρήστη. Τα βασικά χαρακτηριστικά του είναι : διαχείριση οντολογιών μέσω γραφικού περιβάλλοντος, αυτόματη δημιουργία στιγμιοτύπων μέσω φορμών επεξεργασίας, έλεγχος συνέπειας και υποστήριξη συνεργατικής εργασίας. Ο WebOnto Server είναι μια ελεύθερα διαθέσιμη υπηρεσία και μια βιβλιοθήκη με πάνω από 100 οντολογίες είναι προσπελάσιμη χωρίς περιορισμούς μέσω του WebOnto.

2.14 Αξιολόγηση γλωσσών αναπαράστασης οντολογιών

Έχοντας ως βάση τον Παγκόσμιο Ιστό (WWW) και τη δημιουργία του Σημασιολογικού που ολοένα και περισσότερο εξελίσσονται, το διαδίκτυο σήμερα κινείται προς πιο ευφυείς και αυτοματοποιημένους τρόπους αναζήτησης και εξαγωγής της πληροφορίας. Για το λόγο τούτο οι υπολογιστές θα πρέπει να έχουν πρόσβαση σε οργανωμένες συλλογές πληροφοριών που καλούνται οντολογίες. Οι οντολογίες παρέχουν μια μέθοδο αναπαράστασης της γνώσης και μπορούν να χρησιμοποιηθούν πλέον από τα υπολογιστικά συστήματα για τη διεξαγωγή ακριβή συμπερασμάτων. Η περιγραφή και αναπαράσταση των οντολογιών αναφορικά με τα παραπάνω γίνεται μέσω γλωσσών αναγνώσιμων από τη μηχανή.

Οι γλώσσες των οντολογιών προσφέρουν τη δυνατότητα στις μηχανές να μπορούν πλέον να τις επεξεργαστούν καλύτερα και να κατανοήσουν τα

δεδομένα τα οποία μέχρι σήμερα περιορίζονται απλώς να παρουσιαστούν, αποδίδοντας έτσι πιο ακριβή και ολοκληρωμένα συμπεράσματα.

Συνεπώς, έχοντας ως πυρήνα την OWL και η οποία αναπτύχθηκε με βάση πρωτοπόρων γλωσσών όπως XML, RDF/S, για να επιδείξει τη δυνατότητα εκφραστικών συμπερασμάτων βασισμένων στη λογική περιγραφή και τις άλλες γλώσσες που βρίσκονται ακόμη υπό εξέλιξη όπως η XOL, DAML+OIL, SHOE και άλλες, εξάγονται συμπεράσματα τα οποία καθιστούν δυνατή την αναπαράσταση γνώσης στο Σημασιολογικού Ιστού.

ΚΕΦΑΛΑΙΟ 3: ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ – ΠΑΓΚΟΣΜΙΟΣ ΙΣΤΟΣ

3.1 Το Διαδίκτυο (Ιστός) σήμερα

Η ευρεία χρήση του Διαδικτύου¹⁹ (Ιστού) στις μέρες μας έχει αλλάξει ριζικά τον τρόπο επικοινωνίας των ανθρώπων, τον τρόπο διάθεσης και ανάπτυξης της πληροφορίας καθώς και τον τρόπο διεξαγωγής των επιχειρηματικών δραστηριοτήτων. (Berners-Lee et al., 2000) Ασκεί τρομερή επιρροή στην γνώση και τη διαμόρφωση απόψεων. Η ανάπτυξη του Διαδικτύου δεν θα ήταν τόσο ραγδαία αν δεν υπήρχαν οι μηχανές αναζήτησης (π.χ. [Google](#), [Yahoo!](#), [AltaVista](#)) που βασίζονται σε λέξεις-κλειδιά και μέσω αυτών εκατομμύρια άνθρωποι έχουν εύκολη και άμεση πρόσβαση σε ένα τεράστιο, παγκόσμιο και ποικίλο όγκο πληροφοριών. Συγκρινόμενο με τις έντυπες εγκυκλοπαίδειες και τις παραδοσιακές βιβλιοθήκες, το Διαδίκτυο αντιπροσωπεύει μία ξαφνική και απότομη αποκέντρωση των πληροφοριών και των δεδομένων.

Παρόλα αυτά υπάρχουν κάποια προβλήματα που σχετίζονται με τη χρήση των μηχανών αναζήτησης. Ένα σημαντικό πρόβλημα είναι η μεγάλη «ευαισθησία» τους στο λεξιλόγιο. Πολύ συχνά δεν επιστρέφονται έγγραφα που σχετίζονται με τις λέξεις-κλειδιά γιατί χρησιμοποιούν διαφορετική ορολογία. Αυτό δεν είναι αποτελεσματικό γιατί οι σημασιολογικά όμοιες αναζητήσεις θα έπρεπε να επιστρέφουν και ίδια αποτελέσματα. Έτσι, υπηρεσίες όπως η αναζήτηση πληροφοριών, έχουν γίνει πλέον φοβερά κουραστικές για τους χρήστες. Ακόμη και οι καλύτερες μηχανές αναζήτησης αδυνατούν να

¹⁹ Ο όρος *Διαδίκτυο* (με κεφαλαίο δέλτα) αναφέρεται στο παγκόσμιο δίκτυο Internet, δηλαδή στο παγκόσμιο πλέγμα διασυνδεδεμένων υπολογιστών και των υπηρεσιών και πληροφοριών που παρέχει στους χρήστες.

Η λέξη *διαδίκτυο* (με πεζό δέλτα) αναφέρεται σε ιδιωτικά δίκτυα που χρησιμοποιούν την τεχνολογία TCP/IP.

ανταποκριθούν στις απαιτήσεις τους, λόγω του μεγάλου όγκου των αρχείων καθώς και των πολλαπλών ερμηνειών που μπορεί να έχει μία λέξη.

Το μεγαλύτερο ποσοστό του περιεχομένου που υπάρχει σήμερα στο Διαδίκτυο είναι σχεδιασμένο ώστε να είναι κατανοητό μόνο από τους ανθρώπους, και όχι από τις εφαρμογές των Η/Υ. Για παράδειγμα μπορούν να καταλάβουν ότι σε μία ιστοσελίδα υπάρχει ένας σύνδεσμος (link) για μία άλλη ιστοσελίδα. Σε καμία περίπτωση όμως δεν μπορούν να «κατανοήσουν» τη σημασιολογία, λόγω χάριν ότι η ιστοσελίδα <http://bpis.teipat.gr> είναι η ιστοσελίδα του τμήματος Επιχειρηματικού Σχεδιασμού και Πληροφοριακών Συστημάτων του ΑΤΕΙ Πατρών, και ότι ο σύνδεσμος «βιβλιοθήκη» εντός της οδηγεί στη κεντρική σελίδα της βιβλιοθήκης του ΑΤΕΙ Πατρών. Βέβαια οι Η/Υ έχουν τη δυνατότητα να σαρώνουν ιστοσελίδες και να αναγνωρίζουν ορισμένα πεδία, τις τιμές τους, κτλ.

Σε αυτό το σημείο τη λύση δίνει ο Σημασιολογικός Ιστός, όραμα του Tim-Berners Lee, ο οποίος σκέφτηκε τη δημιουργία ενός Ιστού, στον οποίο τα δεδομένα θα είναι σε μορφή όπου οι μηχανές θα μπορούν αμοιβαία να κατανοήσουν, θα είναι δηλαδή έξυπνα δεδομένα. Θα μπορούσαμε να πούμε ότι ο Tim-Berners Lee προτείνει την εισαγωγή της Τεχνητής Νοημοσύνης στις τεχνολογίες Διαδικτύου.

3.2 Ορισμός σημασιολογικού ιστού (Semantic Web)

Ο Σημασιολογικός Ιστός (Semantic Web, SW) (Antoniou et al., 2004) είναι ένα όραμα και μια πρόταση για την μετεξέλιξη Παγκόσμιου Ιστού. Ο στόχος του Σημασιολογικού Ιστού είναι να εξελίξει το σημερινό Διαδίκτυο έτσι ώστε οι πληροφορίες που υπάρχουν και διακινούνται σε αυτό να είναι κατανοητές, και κατ' επέκταση αυτόματα επεξεργάσιμες από τους υπολογιστές.

Ο Σημασιολογικός Ιστός δεν είναι ένας νέος Παγκόσμιος Ιστός. Είναι μια επέκταση και βελτίωση του σημερινού ιστού στην κατεύθυνση, κυρίως, της δόμησης της πληροφορίας έτσι ώστε να είναι προσπελάσιμη από προγράμματα υπολογιστών. Η σημερινή αναπαράσταση των κειμένων στις σελίδες του Ιστού που προορίζεται για χρήση από ανθρώπους θα αντικατασταθεί από αναπαράσταση κατανοητή στους υπολογιστές.

Ο Tim Berners-Lee, που επινόησε τον Παγκόσμιο Ιστό το 1989, είχε το όραμα ενός ιστού δεδομένων που μπορούν να επεξεργαστούν από μηχανές και έδωσε τον ακόλουθο ορισμό για τον Σημασιολογικό Ιστό.

“The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation.” (Berners-Lee et al., 2001)

Ο Σημασιολογικός Ιστός, που αποτελεί μια πρωτοβουλία της Κοινοπραξίας του Παγκοσμίου Ιστού (World Wide Web Consortium – W3C), παρέχει μια διεθνώς προσβάσιμη πλατφόρμα που επιτρέπει σε αυτοματοποιημένα εργαλεία αλλά και σε ανθρώπους να επεξεργάζονται και να μοιράζονται δεδομένα. Το κλειδί για την επίτευξη του παραπάνω στόχου είναι τα μεταδομένα²⁰ (metadata) ή, αλλιώς, η μεταπληροφορία.

²⁰ Τα μεταδεδομένα είναι δομημένα και κωδικοποιημένα δεδομένα τα οποία περιγράφουν χαρακτηριστικά πληροφοριακών οντοτήτων, αποσκοπώντας στην ταύτιση, αναγνώριση, ανακάλυψη, αξιολόγηση και διαχείριση των οντοτήτων που περιγράφονται.

Τα μεταδεδομένα κάνουν σαφή την πληροφορία που είναι αόριστη και την εκθέτουν προς αναζήτηση, επεξεργασία και ενοποίηση (integration). Τα μεταδεδομένα είναι δεδομένα που αναφέρονται σε άλλα δεδομένα (data about data).

Συγκεκριμένα περιέχουν μέρος της σημασίας των δεδομένων, γεγονός που δικαιολογεί τον όρο «σημασιολογικός» στον Σημασιολογικό Ιστό.

3.3 Σχέση σημασιολογικού με παγκόσμιο ιστό

Οι τεχνολογίες που επιτρέπουν την κατανόηση δεδομένων του WWW από υπολογιστές έχουν αρχίσει να γίνονται ολοένα και πιο δημοφιλείς. (ISSEL, 2006) Ο παγκόσμιος ιστός μπορεί να φτάσει το μέγιστο των δυνατοτήτων του, μόνο όταν ο τεράστιος όγκος των δεδομένων που υπάρχουν σε αυτόν γίνουν επεξεργάσιμα τόσο από αυτοματοποιημένα εργαλεία όσο και από ανθρώπους. Για να παραμείνει ο παγκόσμιος ιστός επεκτάσιμος, οι αυριανές εφαρμογές λογισμικού θα πρέπει να μπορούν να μοιράζονται και να επεξεργάζονται δεδομένα, από διάφορες πηγές, ακόμα κι αν οι ίδιες οι εφαρμογές έχουν σχεδιαστεί εντελώς ανεξάρτητα μεταξύ τους αλλά και από τα δεδομένα του WWW. Αυτός είναι και ο στόχος του Σημασιολογικού Ιστού. Το να υπάρχουν δεδομένα καθορισμένα και συνδεδεμένα σημασιολογικά έτσι ώστε διάφορες εφαρμογές να μπορούν να τα επεξεργαστούν, να τα συνδυάζουν και να τα επαναχρησιμοποιούν.

Ο Παγκόσμιος Ιστός βασίζεται κυρίως σε έγγραφα γραμμένα σε HTML (Hypertext Markup Language), μια γλώσσα η οποία περιγράφει το σώμα ενός δομημένου κειμένου δίνοντας έμφαση στην οπτική παρουσίαση, διανθίζοντάς το με αντικείμενα πολυμέσων όπως εικόνες και φόρμες διαλόγου.

Ο Σημασιολογικός Ιστός αντιμετωπίζει την αδυναμία αυτή χρησιμοποιώντας γλώσσες που περιγράφουν δεδομένα και τη σχέση που έχουν αυτά μεταξύ τους. Δύο από αυτές τις γλώσσες είναι η RDF (Resource Description Framework) και η OWL (Web Ontology Language) οι οποίες περιγράφονται στο τρίτο κεφάλαιο και είναι κατανοητές από τους υπολογιστές.

Το μεγαλύτερο ποσοστό του περιεχομένου που υπάρχει σήμερα στο Διαδίκτυο (Internet) είναι σχεδιασμένο ώστε να είναι κατανοητό μόνο από τους ανθρώπους, και όχι από τις εφαρμογές των Η/Υ. Βέβαια οι Η/Υ έχουν τη δυνατότητα να σαρώνουν ιστοσελίδες (websites) και να αναγνωρίζουν ορισμένα πεδία (tags), τις τιμές τους (values), κτλ.

1. Ο σημασιολογικός ιστός είναι εξέλιξη του παγκοσμίου ιστού πληροφοριών και όχι επανάσταση.
2. Η προσθήκη μεταδεδομένων δίνει τη δυνατότητα στους πράκτορες λογισμικού²¹ να κατανοήσουν καλύτερα την πληροφορία και να αναγνωρίσουν με τι σχετίζεται.
3. Ενώ στο παρελθόν, το περιεχόμενο ήταν κατανοητό μόνο από τους ανθρώπους, πλέον είναι κατανοητό και από τις μηχανές με χρήση των πρακτόρων.
4. Οι οντολογίες (στις οποίες θα γίνει αναλυτική αναφορά στο επόμενο κεφάλαιο), ως λεξιλόγια με τα οποία δεσμεύονται οι πράκτορες, δίνουν το κατάλληλο νόημα στα μεταδεδομένα, ώστε να αποφεύγονται τυχόν συγχύσεις στην επικοινωνία και ανταλλαγή πληροφοριών μεταξύ των πρακτόρων.

²¹ Προγράμματα που σαρώνουν τον παγκόσμιο ιστό για να βρουν πληροφορίες σύμφωνα με τις ανάγκες μας και τις ερωτήσεις μας.

5. Η χρήση του σημασιολογικού ιστού θα γίνεται κυρίως με χρήση των πρακτόρων, οι οποίοι θα αναλαμβάνουν τη διεκπεραίωση σύνθετων εργασιών, αλλά και μέσω εφαρμογών συγκεκριμένου τύπου.

3.4 Οι πράκτορες Λογισμικού

Οι ευφυείς Πράκτορες Λογισμικού μπορούν να θεωρηθούν ως οντότητες λογισμικού, οι οποίες επιδεικνύουν ένα σύνολο από χαρακτηριστικά, όπως αυτονομία, επικοινωνία, κινητικότητα και προνοητικότητα. Είναι επίσης ικανοί να επιδεικνύουν ευφυείς συμπεριφορές, μέσα στο πλαίσιο κοινωνιών πρακτόρων.

Στην πληροφορική, (WikiSoft, 2008) ένας **πράκτορας λογισμικού** είναι ένα κομμάτι του λογισμικού που ενεργεί για έναν χρήστη ή άλλο πρόγραμμα σε μια σχέση αντιπροσωπείας. Τέτοια «δράση εξ ονόματος» υπονοεί την αρχή για να αποφασίσει ποια (και εάν) δράση είναι κατάλληλη. Η ιδέα είναι ότι οι πράκτορες δεν επικαλούνται αυστηρά για έναν στόχο, αλλά ενεργοποιούνται για αυτόν.

Ένας πράκτορας λογισμικού είναι μία οντότητα που αντιλαμβάνεται το περιβάλλον μέσα στο οποίο βρίσκεται με τη βοήθεια αισθητήρων (sensors), είναι μέρος του περιβάλλοντος αυτού, κάνει συλλογισμούς για το περιβάλλον και δρα πάνω σε αυτό με τη βοήθεια μηχανισμών δράσης (effectors), για την επίτευξη κάποιων στόχων. Η έννοια ενός πράκτορα παρέχει έναν κατάλληλο και ισχυρό τρόπο να περιγραφεί μια σύνθετη οντότητα λογισμικού που είναι ικανή να λειτουργεί με έναν ορισμένο βαθμό αυτονομίας προκειμένου να ολοκληρωθούν οι στόχοι εξ ονόματος του χρήστη της. Αλλά αντίθετα από το

σκοπό, που καθορίζεται από τις μεθόδους και τις ιδιότητες του , ένας πράκτορας προσδιορίζεται από την συμπεριφορά του. (Βλαχάβας et al, 2005)

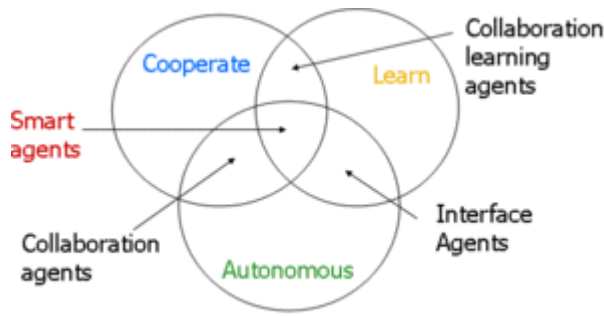
Συντάκτες, έχουν προτείνει πολλούς διαφορετικούς ορισμούς για τους πράκτορες. Αυτοί περιλαμβάνουν συνήθως έννοιες όπως

- *εμμονή* (ο κώδικας δεν εκτελείται κατόπιν παραγγελίας αλλά τρέχει συνεχώς και αποφασίζει για το πότε πρέπει να εκτελέσει κάποια δραστηριότητα)
- *αυτονομία* (οι πράκτορες έχουν την ικανότητα επιλογής στόχου, καθορισμού προτεραιοτήτων, κατευθυνόμενης από το στόχο συμπεριφοράς, και λήψης αποφάσεων χωρίς ανθρώπινη επέμβαση)
- *κοινωνική δυνατότητα* (οι πράκτορες είναι σε θέση να δεσμεύσουν άλλα συστατικά μέσω κάποιου είδους επικοινωνίας και συντονισμού, και να συνεργαστούν για έναν στόχο)
- *ικανότητα αμέσου αντιδράσεως* (οι πράκτορες αντιλαμβάνονται το πλαίσιο στο οποίο λειτουργούν και αντιδρούν σε αυτό κατάλληλα).

Η έννοια *πρακτόρων* είναι πιο χρήσιμη ως εργαλείο που αναλύει τα συστήματα

Οι διάφοροι ερευνητές, (WikiSoft, 2008) δίνοντας βάρος σε διαφορετικές πτυχές της έννοιας ανάλογα με το πρόβλημα που προσπαθούν να επιλύσουν, κάνουν λόγο για "ευφυείς πράκτορες", για "αυτόνομους πράκτορες", γενικότερα για "υπολογιστικούς πράκτορες" ή για συνδυασμούς των όρων αυτών. Άλλοι πράκτορες σχεδιάζονται να δουλεύουν μόνοι τους, άλλοι να συνεργάζονται μέσα σε ομάδες ή κοινωνίες, άλλοι έχουν τη δυνατότητα να μαθαίνουν, άλλοι είναι κινητοί και άλλοι όχι.

Η εικόνα 3.1 αναπαριστά γραφικά τις σχέσεις μεταξύ πρακτόρων.



Εικόνα 3.1 Based on Nwana's primary attribute dimension

Αναλυτικότερα:

- Οι **ευφυείς πράκτορες**, που συγκεκριμένα εκθέτουν κάποια πτυχή της τεχνητής νοημοσύνης, όπως η εκμάθηση και ο συλλογισμός. Οι ικανότητες των ευφών πρακτόρων περιλαμβάνουν:

Την δυνατότητα να *προσαρμοστούν*: Ανάλογα με την αίσθηση του περιβάλλοντος, αναπροσαρμόζουν την απάντηση. Αυτό μπορεί να επιτευχθεί μέσω της επιλογής των εναλλακτικών προβλημάτων ή αλγορίθμων, ή μέσω της ανακάλυψης των στρατηγικών επίλυσης προβλήματος. Η προσαρμογή μπορεί επίσης να περιλάβει άλλες πτυχές της εσωτερικής κατασκευής ενός πράκτορα, όπως η στρατολόγηση των πόρων επεξεργαστών ή αποθήκευσης.

Την δυνατότητα να *μάθει*: Η εκμάθηση μπορεί να προχωρήσει δοκιμάζοντας και λάθη, περιέχει την ικανότητα της ενδοσκόπησης και της ανάλυσης της συμπεριφοράς και να επιτύχει. Εναλλακτικά, η εκμάθηση μπορεί να προχωρήσει από παραδείγματα και γενικεύσεις, έχοντας την ικανότητα να αφαιρεί και να γενικεύει.

- Οι **αυτόνομοι πράκτορες** είναι ικανοί να μεταβάλουν τον τρόπο με τον οποίο επιτυγχάνουν τους στόχους τους. Όντας ανεξάρτητοι και με ανεξάρτητες αποφάσεις, είναι σε θέση να ικανοποιήσουν τους εσωτερικούς στόχους που

βασίζονται στο αντιληπτό περιβάλλον τους. Όλοι οι πράκτορες λογισμικού στις σημαντικές εφαρμογές τους, εποπτεύονται από τους ανθρώπους που τις ξεκινούν, οι οποίοι παρακολουθούν και τροποποιούν συνεχώς τη συμπεριφορά τους.

- **οι διανεμημένοι πράκτορες** εκτελούνται στους φυσικά ευδιάκριτους υπολογιστές. Δεδομένου ότι οι πράκτορες είναι καλά ταιριασμένοι έτσι ώστε να περιλάβουν τους απαραίτητους πόρους στην περιγραφή τους, γίνεται πιο εύκολο να εκτελεστούν ως ανεξάρτητα τμήματα στους διανεμημένους επεξεργαστές. Ο κώδικας πρακτόρων είναι ιδιαίτερα εύκολο να εφαρμοστεί σε μια διανεμημένη μόδα.
- **τα συστήματα πολύ-πρακτόρων** διανεμημένοι πράκτορες που δεν έχουν τις ικανότητες να επιτύχουν έναν στόχο από μόνοι τους και πρέπει να επικοινωνήσουν. Όταν διάφοροι πράκτορες διενεργούν μπορούν να διαμορφώσουν ένα σύστημα πολλαπλών-πρακτόρων ή αλλιώς πολύ-πρακτόρων. Τέτοιοι πράκτορες δεν θα έχουν όλα τα στοιχεία ή όλες τις μεθόδους διαθέσιμα για να επιτύχουν έναν στόχο και έτσι θα πρέπει να συνεργαστούν με άλλους πράκτορες. Όπως και με τους διανεμημένους πράκτορες, το στοιχείο αποκεντρώνεται και η εκτέλεση είναι ασύγχρονη. Σχετικοί τομείς περιλαμβάνουν τη διανεμημένη τεχνητή νοημοσύνη (Distributed Artificial Intelligence) και τη διανεμημένη επίλυση προβλήματος (distributed problem solving).
- **οι κινητοί πράκτορες** (WikiMob, 2008) αναφέρονται σε μια διαδικασία που μπορεί να μεταφέρει το κράτος της από ένα περιβάλλον σε άλλο, με τα στοιχεία της άθικτα, και όντας σε θέση ακόμα να αποδώσει κατάλληλα στο νέο περιβάλλον. Οι κινητοί πράκτορες αποφασίζουν πότε και πού να κινηθούν

3.4.1 Χαρακτηριστικά πρακτόρων

Αν και υπάρχουν πολλά είδη πρακτόρων, υπάρχει ένα κοινό χαρακτηριστικό που πρέπει να υπάρχει σε όλα. Το χαρακτηριστικό αυτό είναι η **αυτονομία**, δηλαδή η αυτενέργεια ενός πράκτορα για την επίτευξη των στόχων του. Για να υπάρχει αυτονομία σε ένα πράκτορα, πρέπει να υπάρχει σε αυτόν ενσωματωμένη “νοημοσύνη”, τουλάχιστον μέχρι κάποιο βαθμό (έστω και μικρό). Για το λόγο αυτό αναφέρεται ότι όλοι οι πράκτορες ανήκουν στην ευρύτερη κατηγορία των ευφύων πρακτόρων.

• **Αυτονομία (autonomy)**, δηλαδή το να έχει η ίδια οντότητα έλεγχο επάνω στις ενέργειες της και στην εσωτερική της κατάσταση.

• **Προνοητικότητα ή προενεργητικότητα (pro-activeness)**, η ικανότητα δηλαδή της οντότητας να λαμβάνει πρωτοβουλίες για ενέργειες τέτοιες που να βοηθούν στην πραγμάτωση των σκοπών της.

• **Αντιδραστικότητα (reactivity) ή Αποκρισιμότητα (responsiveness)**, δηλαδή η ικανότητα γρήγορης και «σωστής» απόκρισης στις αλλαγές του περιβάλλοντός της, όπως το αντιλαμβάνεται η ίδια η οντότητα («αντίληψη περιβάλλοντος»).

• **Κοινωνική ικανότητα (social ability)**, δηλαδή η ικανότητα κοινωνικής συμπεριφοράς κατά την αλληλεπίδραση με άλλους πράκτορες και ανθρώπους. Η επικοινωνία μεταξύ των πρακτόρων μπορεί να γίνεται με χρήση ειδικής γλώσσας επικοινωνίας πρακτόρων.

Πέρα από αυτά τα κύρια χαρακτηριστικά, δευτερεύουσες ιδιότητες που συχνά χαρακτηρίζουν τους πράκτορες είναι η **ειλικρίνεια**, δηλαδή η μη "συνειδητή" παραπληροφόρηση άλλων, η **καλή προαίρεση**, δηλαδή η μη ύπαρξη αλληλοσυγκρουόμενων στόχων, και η **λογική**, δηλαδή η εκτέλεση μόνο ενεργειών που φαίνεται ότι θα συμβάλλουν στην επίτευξη των στόχων (ή τουλάχιστον δε φαίνεται να την αποτρέπουν). Επίσης, άλλη ιδιότητα που συχνά

χαρακτηρίζει πράκτορες είναι η **κινητικότητα**, δηλαδή η δυνατότητα να μεταφέρονται μεταξύ συστημάτων (να "μετοικούν"), ώστε να εκμεταλλεύονται απομακρυσμένους πόρους ή να συνεργάζονται με άλλους πράκτορες. (Βλαχάβας et al, 2005)

3.4.2 Σύγκριση ειδών πρακτόρων

Αν γίνει μία σύγκριση των σημερινών πρακτόρων που υπάρχουν στο Διαδίκτυο, με τους πράκτορες του σημασιολογικού ιστού, θα εντοπιστούν σαφή πλεονεκτήματα στην δεύτερη περίπτωση.

Ένας πράκτορας στο τωρινό διαδίκτυο:

- Πρέπει να γνωρίζει λεπτομέρειες μορφοποίησης κάθε ιστοσελίδας που αντλεί πληροφορίες, για να ανακτήσει το τμήμα που τον ενδιαφέρει.
- Πρέπει να χρησιμοποιεί διαφορετική μέθοδο ανάκτησης, για κάθε νέα ιστοσελίδα που ανακτά πληροφορίες. Επιπλέον, η παραμικρή αλλαγή στη δομή-μορφή της ιστοσελίδας συνεπάγεται επαναπρογραμματισμό της μεθόδου εύρεσης πληροφοριών.
- Έχει μεγάλο κόστος επεκτασιμότητας και συντήρησης.

Αντίθετα, με τους πράκτορες του σημασιολογικού ιστού:

- Υπάρχει πρόσβαση σε πληροφορίες ανεξάρτητες από τη μορφή της παρουσίασης τους.
- Οι πληροφορίες είναι συνδεδεμένες με κοινόχρηστο πλαίσιο αναφοράς (οντολογία).
- Το νόημα των πληροφοριών είναι προσβάσιμο και κατανοητό από όλους τους πράκτορες, ανεξάρτητα από το σκοπό για τον οποίο οι πράκτορες αυτοί αναπτύχθηκαν.

Η λειτουργία των πρακτόρων αναβαθμίζεται συνεχώς. Μέχρι σήμερα οι πράκτορες χρησιμοποιούνταν κυρίως:

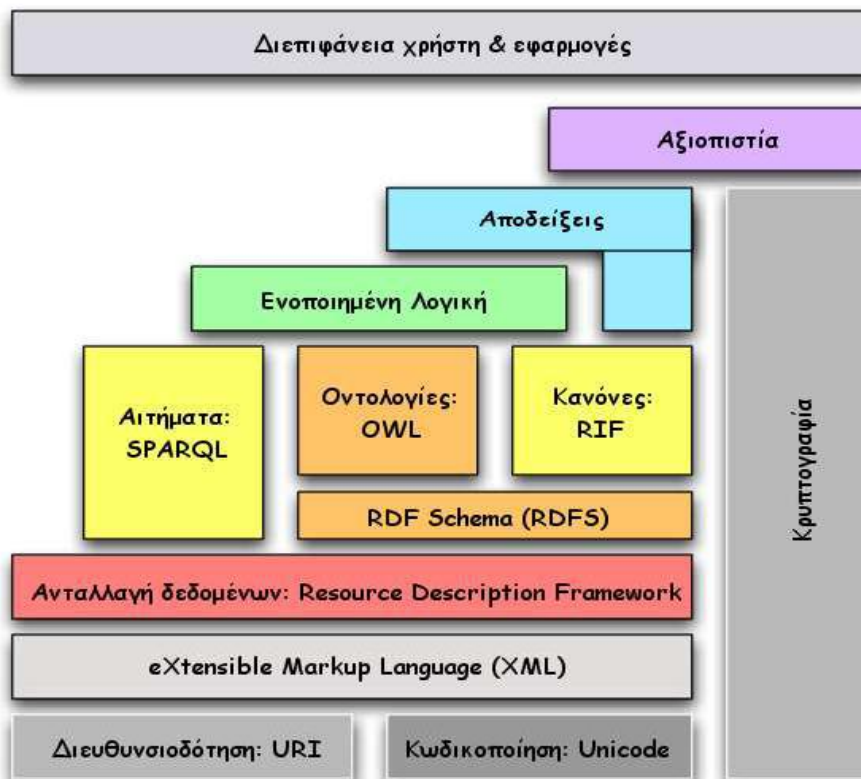
- για την εξαγωγή πληροφοριών από ιστοσελίδες με τη χρήση ευριστικών κανόνων (wrappers, crawlers, κτλ),
- ως διαμεσολαβητές (mediators) μεταξύ ετερογενών πηγών πληροφόρησης με αυστηρά καθορισμένη σύνταξη και σημασιολογία ανταλλαγής πληροφοριών.

Στον σημασιολογικό ιστό οι πράκτορες αναμένεται να:

- δέχονται αιτήματα εξυπηρέτησης από τους χρήστες τους,
- βρίσκουν τις πιθανές απαντήσεις στις απαιτήσεις τους,

παρουσιάζουν τις εναλλακτικές επιλογές δράσης, ώστε οι χρήστες να παίρνουν την τελική απόφαση. (Βλαχάβας et al, 2005)

3.5Η αρχιτεκτονική του σημασιολογικού ιστού



Εικόνα 3.2 Αρχιτεκτονική σημασιολογικού ιστού

3.5.1 Κωδικοποίηση χαρακτήρων

Το πρώτο επίπεδο ξεκινά με τμήμα της κωδικοποίησης καθώς και οι καλύτερες μηχανές αναζήτησης αδυνατούν να ανταποκριθούν στις απαιτήσεις τους, λόγω του μεγάλου όγκου των αρχείων καθώς και των πολλαπλών ερμηνειών που μπορεί να έχει μία λέξη. Με την αναζήτηση αλφαριθμητικών (strings) που χρησιμοποιούν οι μηχανές, είναι αδύνατο να κατανοήσουν λόγω χάριν αν η λέξη “μοντέλο” αναφέρεται σε κάποιο μοντέλο μίας μηχανής αυτοκινήτου, σε κάποιο συλλεκτικό μίνι συναρμολογούμενο μοντέλο

πολεμικών αεροπλάνων, ή σε κάποιον άνθρωπο που εργάζεται επαγγελματικά με την ιδιότητα του μοντέλου.

Η κωδικοποίηση²² των εγγράφων γίνεται με χρήση του διεθνώς καθιερωμένου προτύπου κωδικοποίησης Unicode. Το πρότυπο Unicode μπορεί και υποστηρίζει όλες τις γλώσσες, ακόμη και τις πιο πλούσιες και σύνθετες, όπως είναι για παράδειγμα τα παραδοσιακά Κινέζικα (Βλαχάβας et al, 2005), (WikiChar, 2008)

3.5.2 Διευθυνσιοδότηση εγγράφων

Το αριστερό τμήμα του πρώτου επιπέδου αφορά τη διευθυνσιοδότηση²³. Η διευθυνσιοδότηση των εγγράφων στο διαδίκτυο γίνεται με τη βοήθεια URI, ο πιο γνωστός τύπος του είναι ο ενιαίος εντοπιστής πόρων (Uniform Resource Locator - URL), μέσω του οποίου δηλώνονται οι συμβολοσειρές που χαρακτηρίζουν μοναδικά οποιοδήποτε δικτυακό πόρο. Π.χ. ιστοσελίδα ή άλλου τύπου ηλεκτρονικό αρχείο, διεύθυνση ηλεκτρονικού ταχυδρομείου, διαδικτυακή υπηρεσία κτλ.

Με τη χρήση των URI/IRI είναι πολύ εύκολο για τον καθένα να προσθέτει νέες έννοιες (concepts) στη σημασιολογία που κατασκευάζει. Άλλα πλεονεκτήματα των URI/IRI είναι ότι είναι ανεξάρτητοι περιεχομένου και πρωτοκόλλου επικοινωνίας, καθώς και το ότι η τεχνολογία τους είναι επεκτάσιμη. Δεν περιορίζεται σε μια μόνο μέθοδο ονοματολογίας και διευθυνσιοδότησης πόρων, ούτε μόνο σε υπαρκτές και προσβάσιμες

²² Κωδικοποίηση είναι η διαδικασία μετατροπής των πληροφοριών (χαρακτήρες, λέξεις, φράσεις, κτλ) σε δεδομένα (δυαδικές τιμές).

²³ Διευθυνσιοδότηση είναι η χρήση αλφαριθμητικών που χαρακτηρίζουν μοναδικά οποιοδήποτε Διαδικτυακό πόρο.

διαδικτυακές πληροφορίες. Π.χ. Προτυποποίηση διευθυνσιοδότησης ανθρώπων χωρίς να έχουν οι αναπαριστώμενοι άνθρωποι υποχρεωτικά υπόσταση στο διαδίκτυο (e-mail, home page).

Π.χ. Ένα βιβλίο **urn : isbn : 960 - 7013 - 28 -X**

3.5.3 Extensible Markup Language (XML)

Το δεύτερο επίπεδο αφορά τη χρήση της τεχνολογίας XML η οποία αποτελεί τη βάση για το Σημασιολογικό Ιστό. Με τη βοήθεια της XML και των βοηθητικών τεχνολογιών που τη συνοδεύουν (DTD, XML Schema, XLL) επιτυγχάνεται η αναπαράσταση του περιεχομένου των εγγράφων. Η XML ανήκει στην κατηγορία των μετά-γλωσσών, διαθέτει ευέλικτο συντακτικό για τη δημιουργία οποιωνδήποτε στοιχείων και την περιγραφή οποιουδήποτε είδους εγγράφου. Η XML είναι κατασκευασμένη έτσι ώστε να επιτρέπει στον καθένα να δημιουργήσει μία δική του δομή για την περιγραφή ενός εγγράφου, αλλά δεν παρέχει καμία ερμηνεία για τη σημασία αυτής της δομής. (Βλαχάβας et al., 2005) (Berners-Lee et al., 2001)

Π.χ. εάν μια εταιρεία έχει το στοιχείο `<price>€12,00</price>` στο τιμολόγιο της ενώ η άλλη το `<cost>€12,00</cost>` δεν υπάρχει τρόπος ένας υπολογιστής να γνωρίζει ότι τα δύο στοιχεία σημαίνουν το ίδιο εκτός εάν εφαρμοστούν οι τεχνικές του Σημασιολογικού Ιστού.

3.5.4 Ανταλλαγή δεδομένων (RDF – Resource Description Framework)

Το τρίτο επίπεδο της αρχιτεκτονικής του Σημασιολογικού Ιστού αποτελεί η τεχνολογία RDF, η οποία είναι ένας σκελετός περιγραφής πόρων. Με τον

σκελετό RDF εκφράζεται η σημασία (meaning) των εννοιών που περιγράφονται σε ένα έγγραφο. Ο σκελετός είναι κοινός και χρησιμοποιείται από όλους τους δημιουργούς εγγράφων. Αυτό σημαίνει ότι με χρήση του RDF η ανταλλαγή δεδομένων είναι εύκολη, αφού όλα τα έγγραφα RDF έχουν κοινή μορφή. Ο RDF (Βλαχάβας et al., 2005) αποτελείται από δηλώσεις ή ισχυρισμούς (statements) οι οποίες αφορούν μεταδεδομένα, δεν εξαρτάται καθόλου από το δεντρικό μοντέλο της XML ωστόσο χρησιμοποιεί τη σύνταξη της για την αναπαράσταση δηλώσεων. Μια δήλωση συνδέει δύο πόρους μεταξύ τους και έχει τη μορφή «υποκείμενο- ρήμα - αντικείμενο» όπως π.χ. «Ο Γιώργος μένει στην Αθήνα», αυτό είναι στην ουσία το μοντέλο δεδομένων της RDF .

3.5.5 Resource Description Framework Schema (RDFS)

Το τέταρτο επίπεδο περιέχει την γλώσσα περιγραφής λεξιλογίων Resource Description Framework Schema (RDFS). Η RDF Schema (RDFS) η οποία ορίζει ένα λεξικό – εκφρασμένο επίσης σε RDF – για να εκφράζονται κατηγορίες (κλάσεις) των πόρων, οι ιδιότητες των πόρων και οι μεταξύ τους σχέσεις. Η RDF Schema θεωρείται ότι είναι η Σημασιολογική επέκταση της RDF, καθώς παρέχει μηχανισμούς για την περιγραφή συνόλων από σχετικούς μεταξύ τους πόρους και τις σχέσεις που τους συνδέουν. Στην RDFS ορίζονται τα περιγραφικά χαρακτηριστικά της RDF τα οποία καθορίζουν διάφορα χαρακτηριστικά των πόρων όπως το πεδίο τιμών και το πεδίο ορισμού των ιδιοτήτων τους. Η γλώσσα RDFS είναι παρόμοια με τις γλώσσες αντικειμενοστραφούς προγραμματισμού με τη διαφορά ότι η RDFS, αντί να ορίζει μια κλάση με βάση τις ιδιότητες που μπορεί να έχουν τα στιγμιότυπα της,

περιγράφει τις ιδιότητες αυτές με βάση τις κλάσεις πόρων τις οποίες αφορούν. Το μεγάλο πλεονέκτημα που απορρέει από αυτή την προσανατολισμένη στις ιδιότητες προσέγγιση είναι ότι επιτρέπει στον καθένα να επεκτείνει την περιγραφή ενός πόρου χωρίς να χρειάζεται να ορίσει ξανά την περιγραφή της κλάσης του πόρου αυτού.

3.5.6 Οντολογίες : OWL

Το πέμπτο επίπεδο αποτελείται από τις οντολογίες , η συνεισφορά των οποίων κρίνεται ιδιαίτερος σημαντική λόγω της δυνατότητας τους για δημιουργία επαναχρησιμοποιήσιμου και δια- μοιραζόμενου λογισμικού. Προς το παρόν, ένας από τους κύριους παράγοντες που εμποδίζουν την επαναχρησιμοποίηση των εφαρμογών, αλλά και την επικοινωνία μεταξύ διαφορετικών συστημάτων είναι το γεγονός της χρήσης διαφορετικών όρων και εννοιών για την περιγραφή του ίδιου πεδίου δραστηριότητας. Αναπτύσσοντας οντολογίες, δημιουργούμε τη βάση για την επικοινωνία πολύπλοκων συστημάτων που θα χρησιμοποιούν ένα κοινό λεξιλόγιο, μια κοινή γλώσσα επικοινωνίας. (Γιουβανάκης Α., 2006)

3.5.7 Κανόνες

Στο πέμπτο επίπεδο υπάρχει επίσης και το τμήμα που αφορά τους κανόνες. Δηλωτικός κανόνας (rule of inference) ενός συστήματος είναι ένα σχήμα δημιουργίας έγκυρων δηλώσεων, τέτοιου είδους σχήματα επιτρέπουν τη δημιουργία συντακτικών σχέσεων μεταξύ ενός συνόλου τύπων (formulas) που ονομάζονται προϋποθέσεις (premises) και ενός ισχυρισμού (assertion) που

ονομάζεται συμπέρασμα (conclusion). Αυτές οι συντακτικές σχέσεις χρησιμοποιούνται στην διαδικασία της δήλωσης (process of inference), και με τη βοήθεια τους μπορούν να δημιουργούνται νέοι ισχυρισμοί από αυτούς που είναι ήδη γνωστοί. (WikiRule, 2008) (Berners-Lee et al., 2001).

Συγκεκριμένα έχει δημιουργηθεί μία ομάδα που ασχολείται με την ανάπτυξη της μορφής ανταλλαγής κανόνων (Rule Interchange Format - RIF). Στόχος της είναι η δημιουργία μίας γλώσσας κανόνων η οποία θα επιτρέπει τη διερμηνεία και τη μεταφορά κανόνων μεταξύ πολλαπλών γλωσσών και συστημάτων (Cover, 2005) .

3.5.8 Αιτήματα

Το τμήμα των αιτημάτων (queries) παρέχει υπηρεσίες στα επίπεδα τέσσερα και πέντε, αυτά χρειάζονται ώστε να είναι δυνατή η εξαγωγή και ο συνδυασμός των πληροφοριών που περιέχουν τα διάφορα έγγραφα που κατανέμονται σε όλο το εύρος του σημασιολογικού ιστού (RDF, RDFS, OWL). Μάλιστα έχει κατασκευαστεί μία γλώσσα υποβολής αιτημάτων η οποία ονομάζεται SPARQL η οποία είναι μία γλώσσα υποβολής αιτημάτων για την εξαγωγή πληροφορίας από γραφήματα RDF.

3.5.9 Ενοποιημένη Λογική

Το έκτο επίπεδο το της στοίβας του σημασιολογικού ιστού αποτελείται από το τμήμα της ενοποιημένης λογικής. Το επίπεδο αυτό χρειάζεται ώστε να γίνονται κατανοητές διάφορες ενέργειες όπως οι κανόνες μετατροπής εγγράφων από έναν τύπο σε κάποιο άλλο, ο έλεγχος ενός εγγράφου ως προς τη συνέπεια του (π.χ. έλεγχος αν ένα έγγραφο ικανοποιεί όλους τους περιορισμούς που θέτει

μία ομάδα οντολογιών), το αποτέλεσμα μίας ερώτησης με σκοπό την ανάκληση όρων ή την αναγνώριση άγνωστων όρων.

Ακόμα μέσω αυτού του επιπέδου παρέχεται η δυνατότητα αναπαράστασης της κατηγορηματικής λογικής (π.χ. τελεστές AND, OR, NOT) και της συσχέτισης των δεδομένων (π.χ. αν ισχύει το X, τότε ισχύουν όλα τα Y). Τέλος το λογικό επίπεδο επιτρέπει την επεξεργασία των δεδομένων των κατώτερων επιπέδων με τη χρήση σημασιολογικών σχέσεων υπό μορφή κανόνων. (Βλαχάβας et al., 2005)

3.5.10 Αποδείξεις

Το έβδομο επίπεδο αποτελεί το τμήμα των αποδείξεων, το οποίο περιλαμβάνει μηχανισμούς και διαδικασίες εξαγωγής συμπερασμάτων που χρησιμοποιούν τη γνώση του λογικού επιπέδου καθώς και τρόπους αναπαράστασης διαδικασιών απόδειξης σε XML ή RDF. Η χρήση των αποδείξεων στην αρχιτεκτονική του σημασιολογικού ιστού παρέχει εύκολη δημιουργία εξηγήσεων προς το χρήστη σχετικά με τα συμπεράσματα στα οποία κατέληξε. Τέλος το έβδομο επίπεδο αυξάνει την εμπιστοσύνη των χρηστών στις παρεχόμενες υπηρεσίες. (Βλαχάβας et al., 2005)

3.5.11 Αξιοπιστία

Το όγδοο και υψηλότερο επίπεδο είναι αυτό της αξιοπιστίας. Η αξιοπιστία είναι ένα βασικό χαρακτηριστικό που πρέπει να παρέχει ο σημασιολογικός ιστός. Υλοποιείται κατά ένα μεγάλο βαθμό χρησιμοποιώντας το τμήμα της

κρυπτογραφίας²⁴ (κρυπτογράφηση, ψηφιακές υπογραφές, κτλ), που παρέχει τις υπηρεσίες του στο επίπεδο της αξιοπιστίας. Στον σημασιολογικό ιστό, όλοι οι χρήστες θα πρέπει να υπογράφουν ψηφιακά τα έγγραφα τους, για να υπάρχει μία μορφή εξακρίβωσης ότι ο χρήστης που ζητάει ή παρέχει μία υπηρεσία είναι όντως αυτός που παριστάνει ότι είναι.

Η αναζήτηση πληροφορίας θα πρέπει να περιλαμβάνει την απόδειξη ότι ο χρήστης έχει το δικαίωμα να λάβει απάντηση. Λόγου χάριν για να επιτραπεί σε κάποιον η πρόσβαση σε ένα δικτυακό τόπο πρέπει να παραδώσει έγγραφο (ηλεκτρονικό) το οποίο να αποδεικνύει το λόγο για τον οποίο του επιτράπηκε η πρόσβαση. Οι ηλεκτρονικές συναλλαγές απαιτούν εχεμύθεια. Τα έγγραφα που διακινούνται θα πρέπει να περιλαμβάνουν επιπλέον πληροφορίες που αφορούν την αξιοπιστία τους. Επίσης, αν τεθεί θέμα, η απάντηση - απόδειξη θα είναι εύκολο να κατασκευαστεί υπολογιστικά. (Βλαχάβας et al., 2005)

3.6 Τα κέρδη που θα επιφέρει η σημασιολογία

Με τη βοήθεια του σημασιολογικού ιστού το Διαδίκτυο θα πάψει να είναι χαώδες κατά ένα μεγάλο βαθμό. Σε συνεργασία με τους πράκτορες, οι οποίοι θα είναι εύκολο να χρησιμοποιηθούν από οποιονδήποτε χρήστη δίχως να απαιτείται προηγούμενη εμπειρία, διευκολύνονται και κυρίως βελτιώνονται διαδικασίες που σχετίζονται με την πληροφορία, όπως (Βλαχάβας et al., 2005) :

- η αναζήτηση,
- η αυτοματοποίηση,
- η ομαδοποίηση,
- η επαναχρησιμοποίηση της από τελείως διαφορετικές διαδικτυακές εφαρμογές.

²⁴ Κρυπτογράφηση είναι η εξασφάλιση αξιόπιστων πηγών πληροφοριών

Πρέπει να γίνει κατανοητό ότι ο σημασιολογικός ιστός δεν βασίζεται στην κωδικοποίηση ιστοσελίδων με σημασιολογία (αν και παρέχει υποστήριξη για κάτι τέτοιο). Βασίζεται κυρίως στην εξαγωγή δεδομένων τα οποία είναι αναγνώσιμα από τις μηχανές - Η/Υ (machine readable) σε πολύ μεγαλύτερη κλίμακα από τη σημερινή. Αυτό σημαίνει ότι για την ανάπτυξη του σημασιολογικού ιστού θα πρέπει ολοένα και περισσότερα εργαλεία δημιουργίας εγγράφων (authoring tools), τεχνολογίες βάσεων δεδομένων, εφαρμογές μεγάλης κλίμακας (enterprise level applications), κτλ να χρησιμοποιούν τα πρότυπα του σημασιολογικού ιστού που έχει ορίσει το W3C για τη δημιουργία περιγραφών (RDF) και οντολογιών (OWL) (Berners-Lee, 2006).

ΚΕΦΑΛΑΙΟ 4: ΑΝΑΠΤΥΞΗ ΤΟΥ ΟΝΤΟΛΟΓΙΚΟΥ PORTAL

4.1 Σκοπός εφαρμογής

Στόχος της ανάπτυξης του οντολογικού πόρταλ είναι η σχεδίαση και υλοποίηση ενός δικτυακού τόπου ο οποίος να αποτελεί σημείο αναφοράς για τις οντολογίες, τα επιμέρους πεδία έρευνας, τις εφαρμογές των οντολογιών και γενικά οτιδήποτε αφορά τις οντολογίες. Ο χρήστης του οντολογικού πόρταλ θα μπορεί εύκολα να πλοηγηθεί πολυμεσικά σε κατάλληλα σχεδιασμένους κόμβους έτσι ώστε να ενημερωθεί για τρέχοντα θέματα οντολογιών, όπως η έρευνα και η εφαρμογή.

Βέβαια για να γίνει αυτό έπρεπε να προηγηθεί η μελέτη και κατηγοριοποίηση όλων των ερευνητικών και εφαρμοσμένων θεμάτων που αφορούν τις οντολογίες όπως αναλύσαμε στα προηγούμενα κεφάλαια.

4.2 Τεχνολογίες – πλατφόρμες που επιλέχθηκαν

Το περιβάλλον ανάπτυξης που επιλέχθηκε έχει σαν βάση του έναν “web server” (εξυπηρετητής-διακομιστής δικτύου) που θα μας παρέχει το πρόγραμμα Apache web server με PHP²⁵ και μια βάση δεδομένων MySQL.

Με τον όρο web server εννοούμε τον υπολογιστή που είναι συνδεδεμένος με το Internet ο οποίος με τη χρήση κατάλληλου λογισμικού επιτρέπει σε άλλες ηλεκτρονικές συσκευές (H/Y, Mobile Phones) να έχουν πρόσβαση στις πληροφορίες ή υπηρεσίες που αυτός παρέχει και με αυτή την έννοια να εξυπηρετεί αυτές τις συσκευές. Ουσιαστικά είναι ένα πρόγραμμα υπολογιστών


²⁵ Γλώσσα προγραμματισμού για τη δημιουργία σελίδων web με δυναμικό περιεχόμενο. Μια σελίδα PHP περνά από επεξεργασία από ένα συμβατό διακομιστή του Παγκόσμιου Ιστού, ώστε να παραχθεί σε πραγματικό χρόνο το τελικό περιεχόμενο, που θα σταλεί στο πρόγραμμα περιήγησης των επισκεπτών σε μορφή κώδικα HTML.
<http://el.wikipedia.org/wiki/PHP>

που είναι αρμόδιο για την αποδοχή των αιτημάτων HTTP²⁶ και δίνει απαντήσεις HTTP μαζί με το περιεχόμενο στοιχείων, το οποίο είναι συνήθως ιστοσελίδας όπως τα έγγραφα HTML και τα συνδεδεμένα αντικείμενα (εικόνες, κ.λπ.). (WikiServer, 2008)

Ο **Apache HTTP** γνωστός και απλά ως Apache είναι ένας εξυπηρετητής του παγκόσμιου ιστού (web). Όποτε επισκέπτεστε ένας ιστότοπο ο πλοηγός σας επικοινωνεί με έναν διακομιστή HTTP. Ο Apache είναι ένας από τους δημοφιλέστερους, εν μέρει γιατί λειτουργεί σε διάφορες πλατφόρμες σαν τα Windows, το Linux, το Unix, και το Mac OS X. (WikiApache, 2008)

Το χώρο αυτό νοικιάσαμε από τον ιστότοπο <http://www.pair.com>.

Ο έλεγχος και η ενοικίαση ονόματος χώρου (domain name²⁷) για να προσδιορίσουμε το portal έγινε στον ιστότοπο <http://www.godaddy.com>.

Στη συνέχεια χρειαζόμασταν ένα Σύστημα Διαχείρισης Περιεχομένου (Content Management System (CMS)) και διαλέξαμε το “Joomla! ”. (Άλλα CMS: Mambo, Drupal).

Ένα CMS μπορεί να υποστηρίξει τα ακόλουθα χαρακτηριστικά γνωρίσματα:

²⁶ Το πρωτόκολλο μεταφοράς υπερκειμένων (HTTP) είναι ένα πρωτόκολλο επικοινωνιών. Μια υπηρεσία που εκτελείται στους web servers και αναλαμβάνει τη διεκπεραίωση των αιτημάτων HTTP, για λήψη σελίδων και άλλων αρχείων από τους επισκέπτες. <http://en.wikipedia.org/wiki/HTTP>

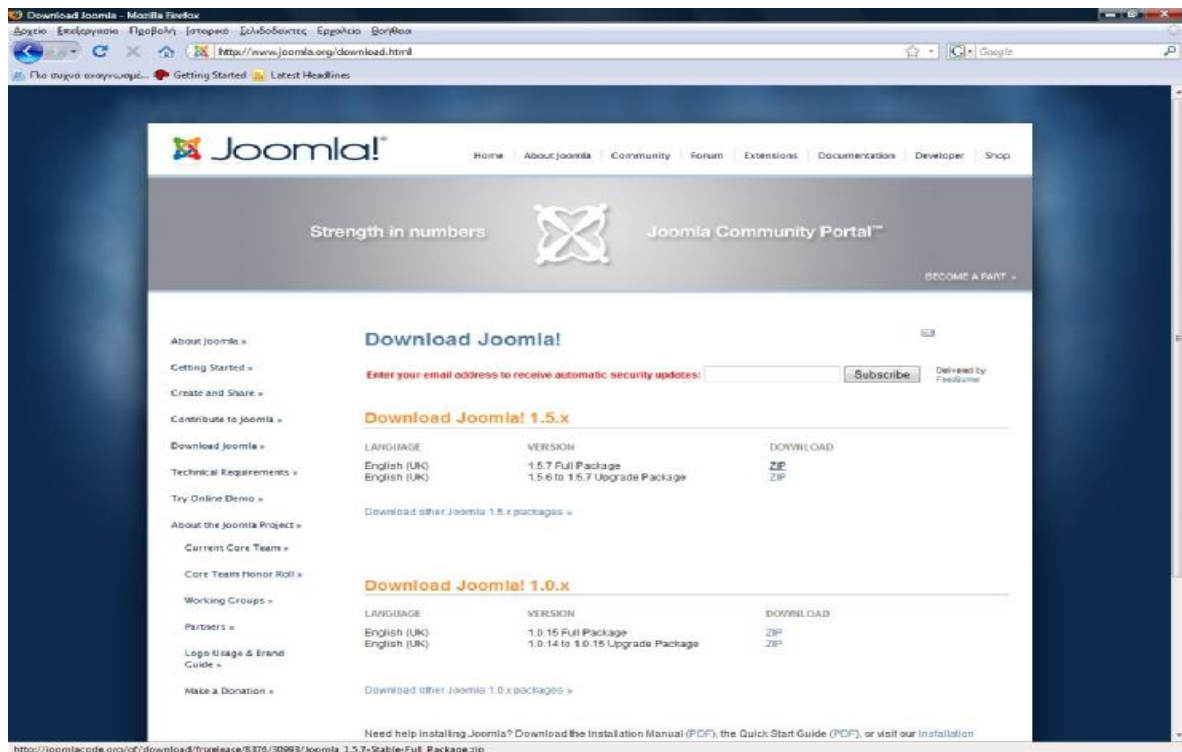
²⁷ Σειρά λέξεων ή συντομογραφίες ή φράσεων που προσδιορίζει ένα συγκεκριμένο υπολογιστή συνδέεται με το Διαδίκτυο και χρησιμεύει ως διεύθυνση, αλφαβητική μορφή για διεύθυνση Διαδικτύου.

- προσδιορισμός όλων των βασικών χρηστών και των διοικητικών ρόλων τους
- η δυνατότητα να δοθούν οι ρόλοι σε διαφορετικές κατηγορίες ή τύπους
- καθορισμός των στόχων ροής της δουλειάς για τη συνεργάσιμη δημιουργία, έτσι ώστε οι διαχειριστές προειδοποιούνται για τις αλλαγές στο περιεχόμενο (παραδείγματος χάριν, ένας ικανοποιημένος δημιουργός υποβάλλει μια ιστορία, η οποία δημοσιεύεται μόνο αφού την αναθεωρεί ο συντάκτης αντιγράφων editor-in-chief και την εγκρίνει.)
- η δυνατότητα να ανιχνευθεί το περιεχόμενο (π.χ. αναζήτηση)
- χωρισμός του σημασιολογικού στρώματος του περιεχομένου από το σχεδιάγραμμά του (παραδείγματος χάριν, το CMS μπορεί αυτόματα να θέσει το χρώμα, τις πηγές, ή την έμφαση του κειμένου.).

Το Joomla! είναι ένα αρθρωτό σύστημα διαχείρισης περιεχομένου (CMS) ανοικτού/ελεύθερου λογισμικού για το περιεχόμενο στο World Wide Web. Είναι γραμμένο στη γλώσσα προγραμματισμού PHP και χρησιμοποιεί τη βάση δεδομένων MySQL εξ' ορισμού. Τελειώνει επίσημα με ένα σημάδι θαυμαστικού, αλλά αυτό παραλείπεται συνήθως.

Το Joomla!, όπως πολλά σύγχρονα CMS, επιτρέπει στο διαχειριστή συστήματος να οργανώνει το περιεχόμενο, να προσαρμόζει την παρουσίαση, να αυτοματοποιεί διαχειριστικές εργασίες και να διαχειρίζεται τους επισκέπτες του ιστοτόπου και αυτούς που συνεισφέρουν. Το σύστημα περιλαμβάνει τα χαρακτηριστικά γνωρίσματα όπως η εναποθήκευση σελίδων για να βελτιώσει την απόδοση, τις εκτυπώσιμες εκδόσεις των σελίδων, τις λάμπσεις ειδήσεων, blogs, τις ψηφοφορίες, την έρευνα ιστοχώρου. Παρόλο που υπάρχει μια πολύπλοκη προγραμματιστική διεπαφή, οι περισσότερες εργασίες μπορούν να γίνουν με λίγο ή και καθόλου προγραμματισμό.

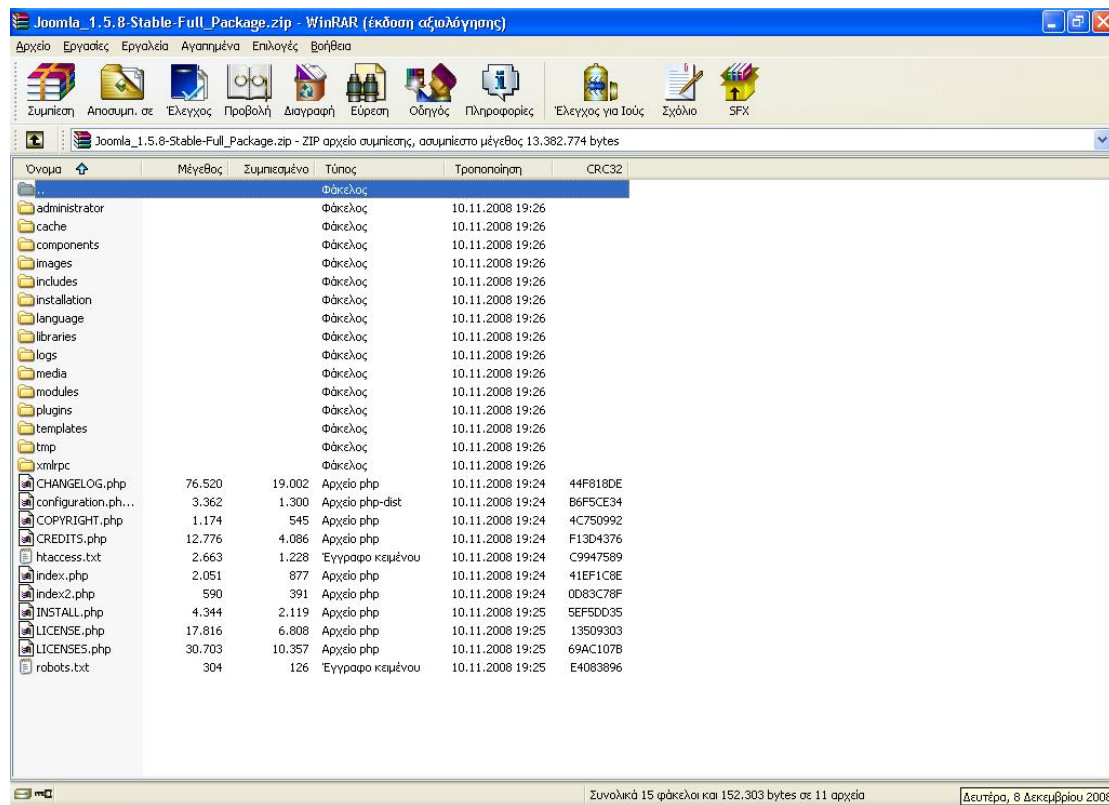
Στο επόμενο στάδιο κατεβάζουμε τοπικά το συμπιεσμένο αρχείο Joomla! 1.5.8 από τη διεύθυνση <http://www.joomla.org>.



Εικόνα 4.1 Πρώτο στιγμιότυπο κατά την διαδικασία εγκατάστασης του Joomla!

Στη συνέχεια αποσυμπιέζουμε το αρχείο μας Joomla_1.5.8-Stable-Full_Package, χρησιμοποιώντας το πρόγραμμα Winrar²⁸.

²⁸ Το WinRAR είναι ένα αρχείο διανεμούμενων λογισμικών archiver και χρησιμοποιείται για την συμπίεση-αποσυμπίεση δεδομένων. Είναι μια από τις λίγες εφαρμογές που είναι σε θέση να δημιουργήσει RAR αρχεία. <http://www.liutilities.com/products/campaigns/vendors/rarlab/featured1/>

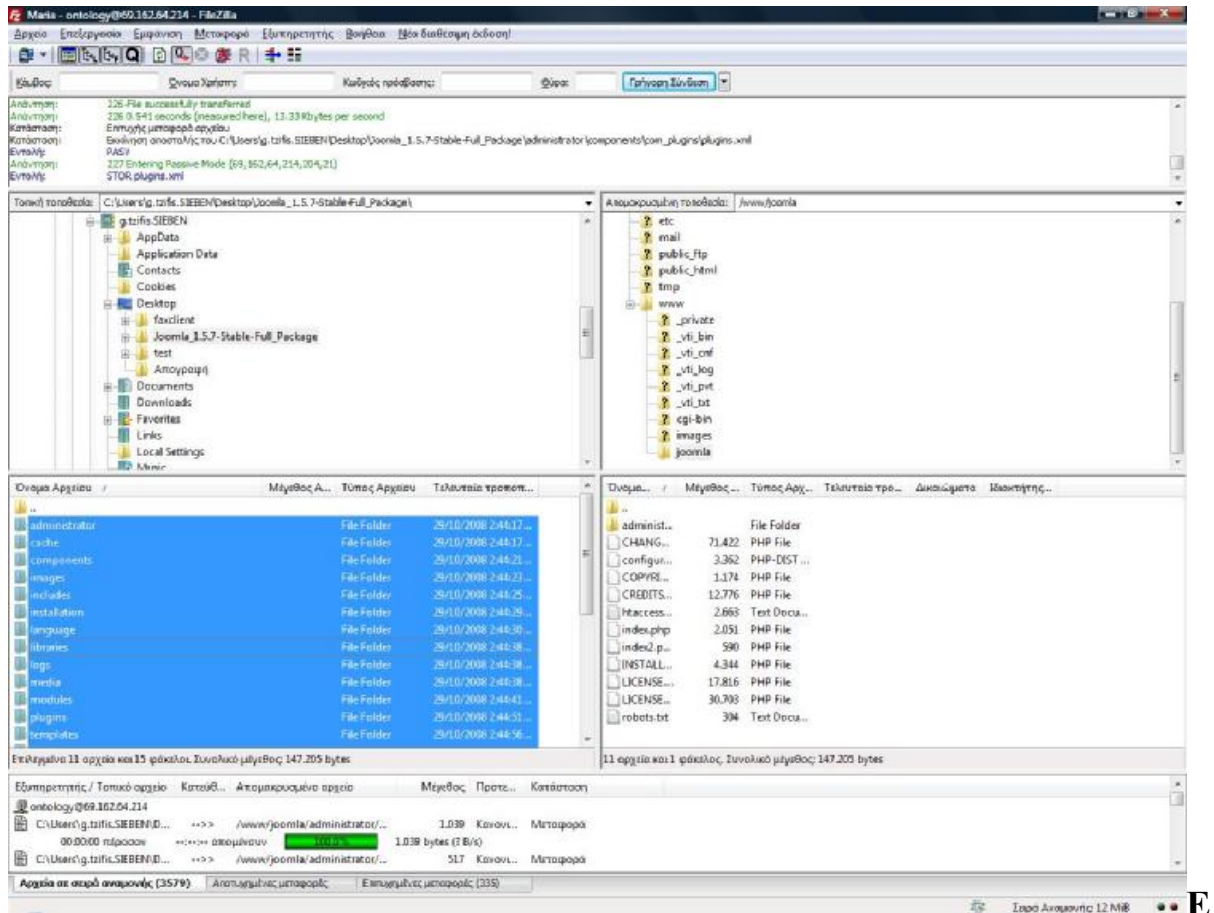


Εικόνα 4.2 Αποσυμπίεση δεδομένων μέσω winrar

Με ένα πρόγραμμα FTP²⁹ (χρησιμοποιήθηκε το Filezilla³⁰ - ftp client) μεταφέρουμε στο server τα αρχεία, ώστε όταν ένας χρήστης ζητάει την εμφάνιση της ιστοσελίδας μας από τον φυλλομετράτε του, ο server να ανταποκρίνεται με την εφαρμογή joomla που έχουμε παραμετροποιήσει για τις ανάγκες μας.

²⁹ Το File Transfer Protocol (FTP), (ελληνικά: *Πρωτόκολλο Μεταφοράς Αρχείων*) είναι ένα ευρέως χρησιμοποιούμενο πρωτόκολλο σε δίκτυα τα οποία υποστηρίζουν το πρωτόκολλο TCP/IP (δίκτυα όπως internet ή intranet). Ο υπολογιστής που τρέχει εφαρμογή FTP client μόλις συνδεθεί με τον server μπορεί να εκτελέσει ένα πλήθος διεργασιών όπως ανέβασμα αρχείων στον server, κατέβασμα αρχείων από τον server, μετονομασία ή διαγραφή αρχείων από τον server κ.ο.κ.

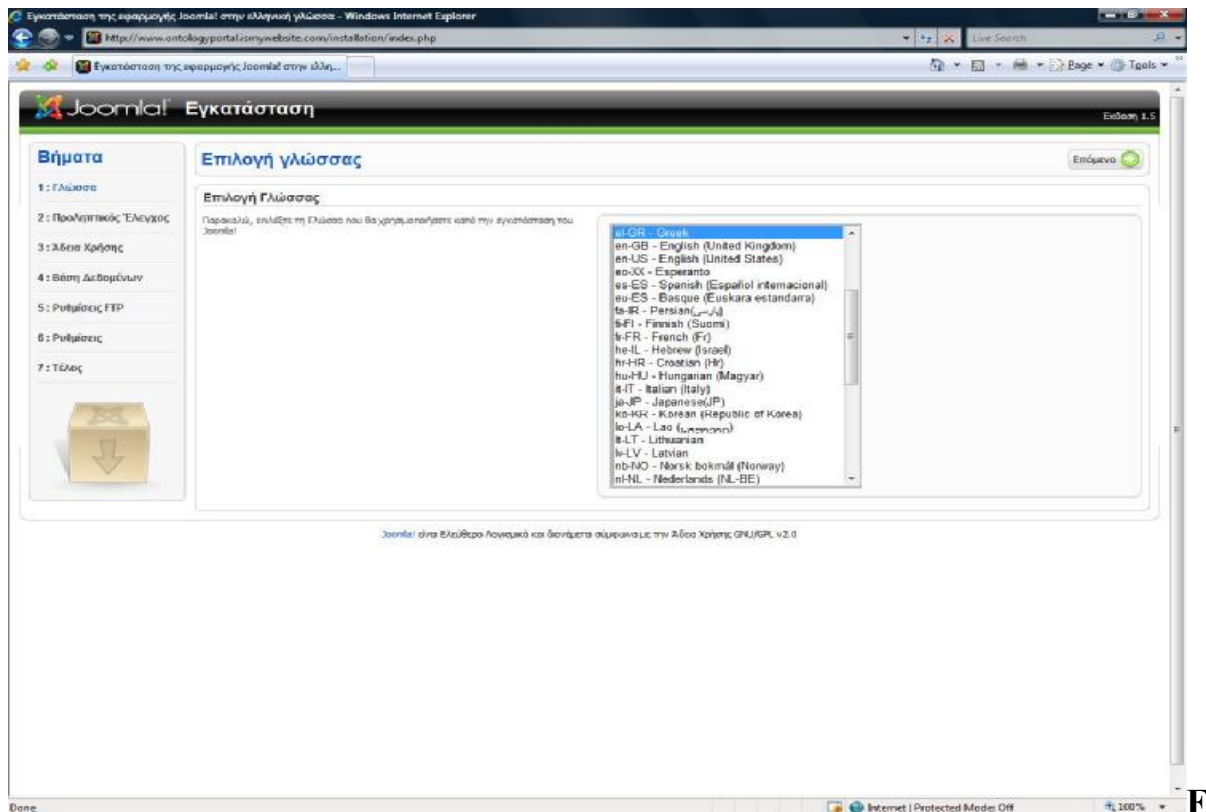
³⁰ Από <http://filezilla-project.org/>.



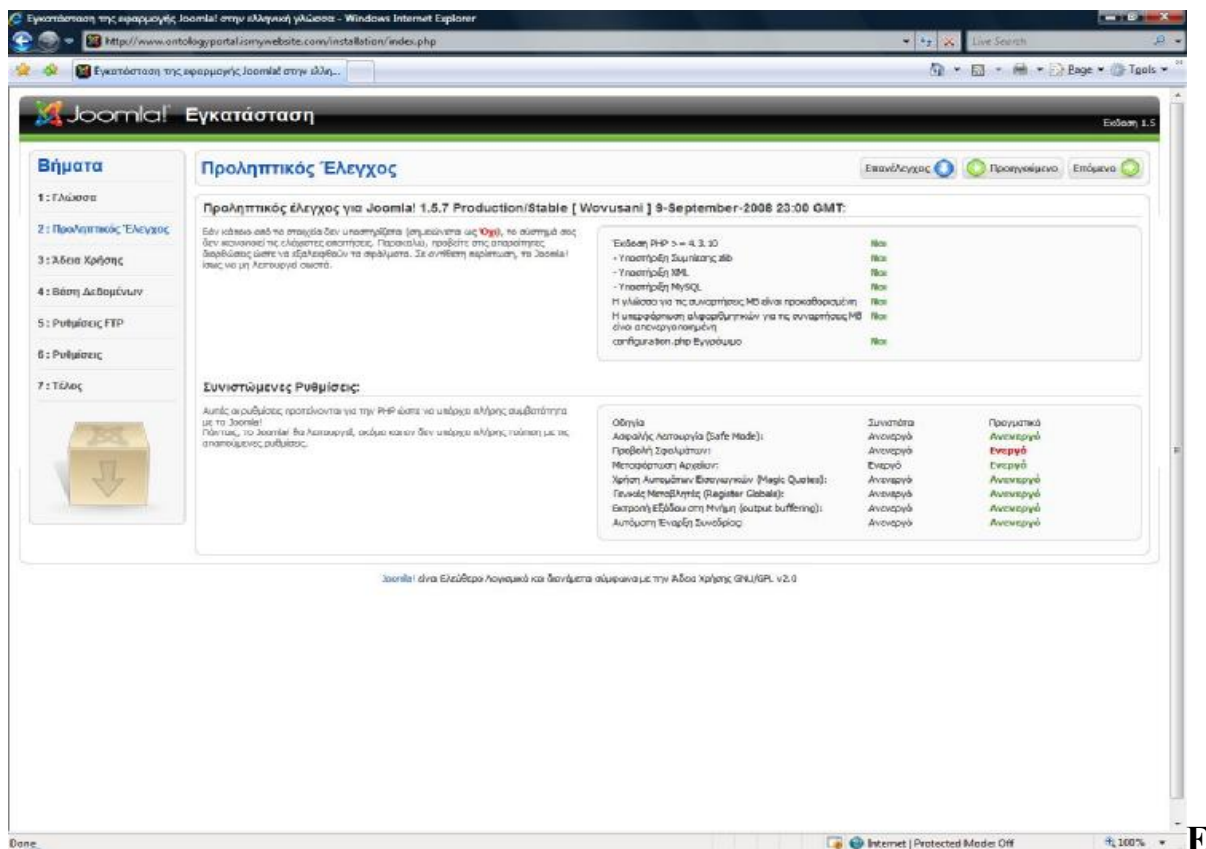
ικώνα 4.3 Σύνδεση με server μέσω FileZilla

Στη συνέχεια καλούμε το αρχείο μέσω internet explorer browser και μέσω του panel διαχείρισης του server δημιουργούμε μια βάση δεδομένων.

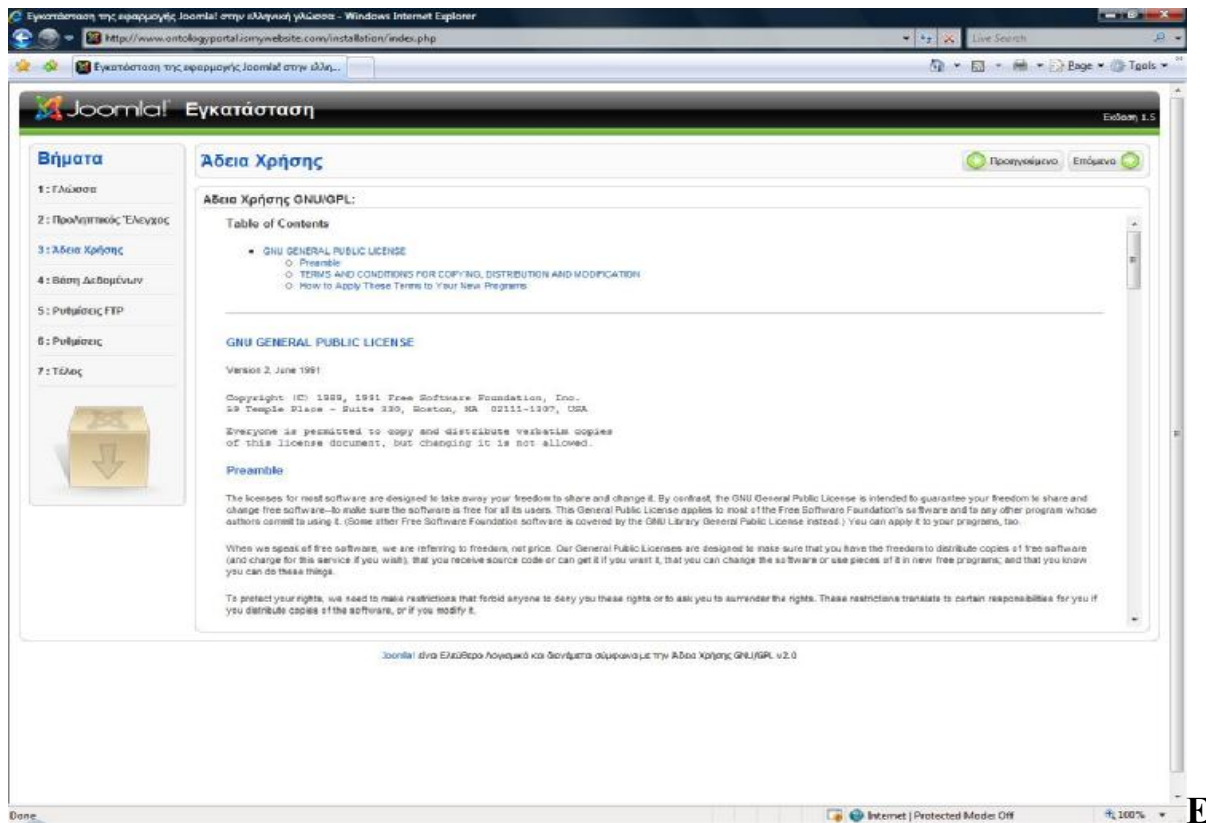
Παρακάτω παραθέτουμε στιγμιότυπα της διαδικασίας εγκατάστασης του Joomla!.



ικόνα 4.4 Επιλογή γλώσσας

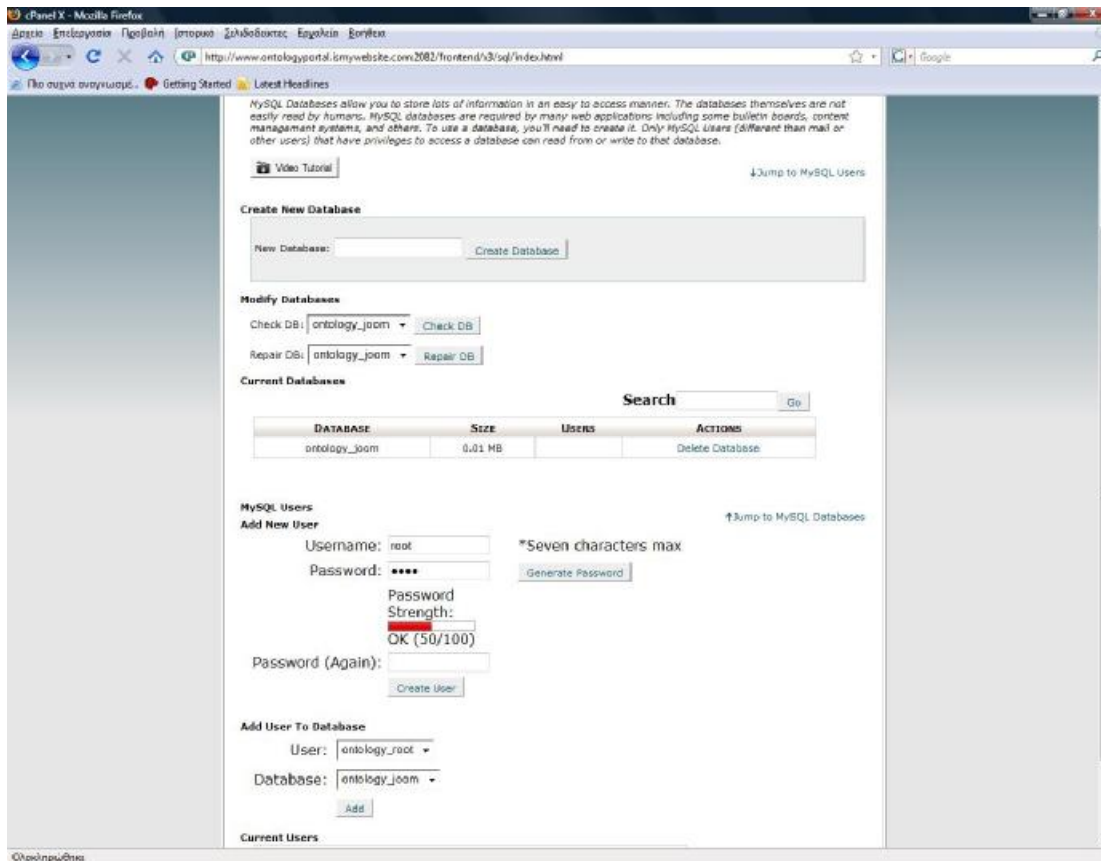


ικόνα 4.5 Προληπτικός έλεγχος

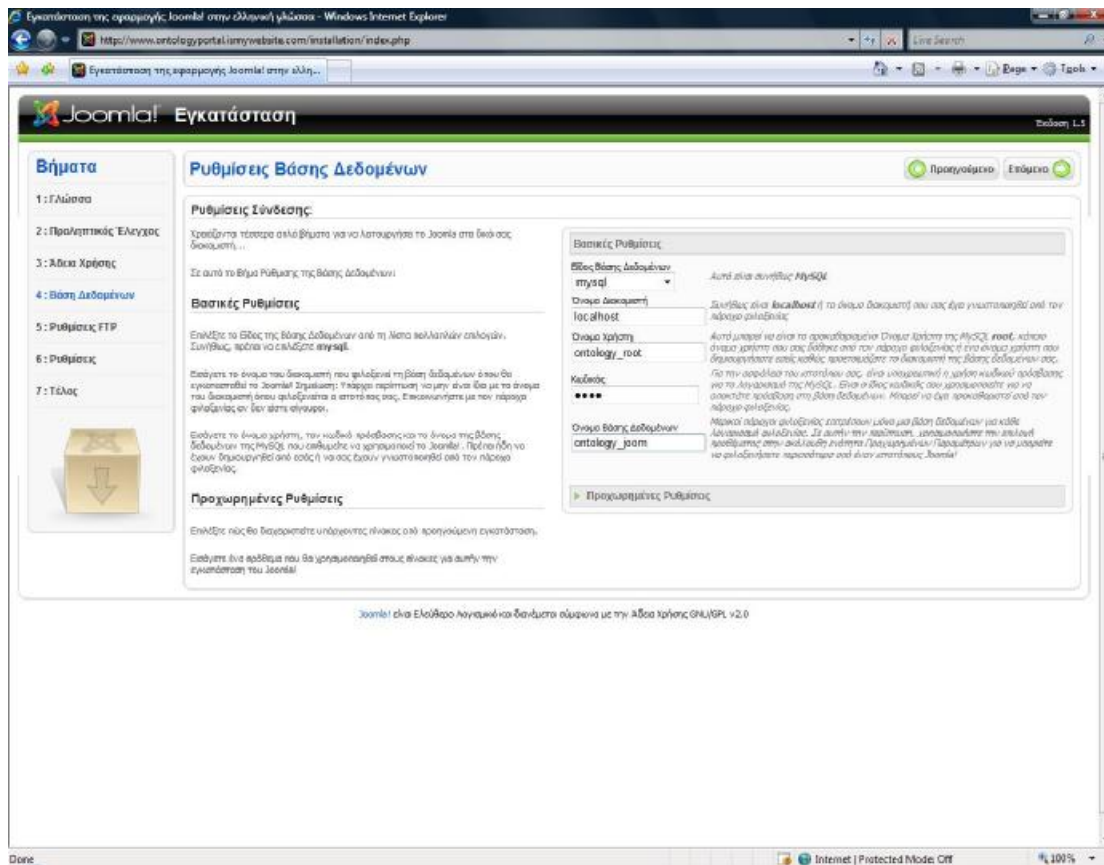


ικώνα 4.6 Άδεια χρήσης

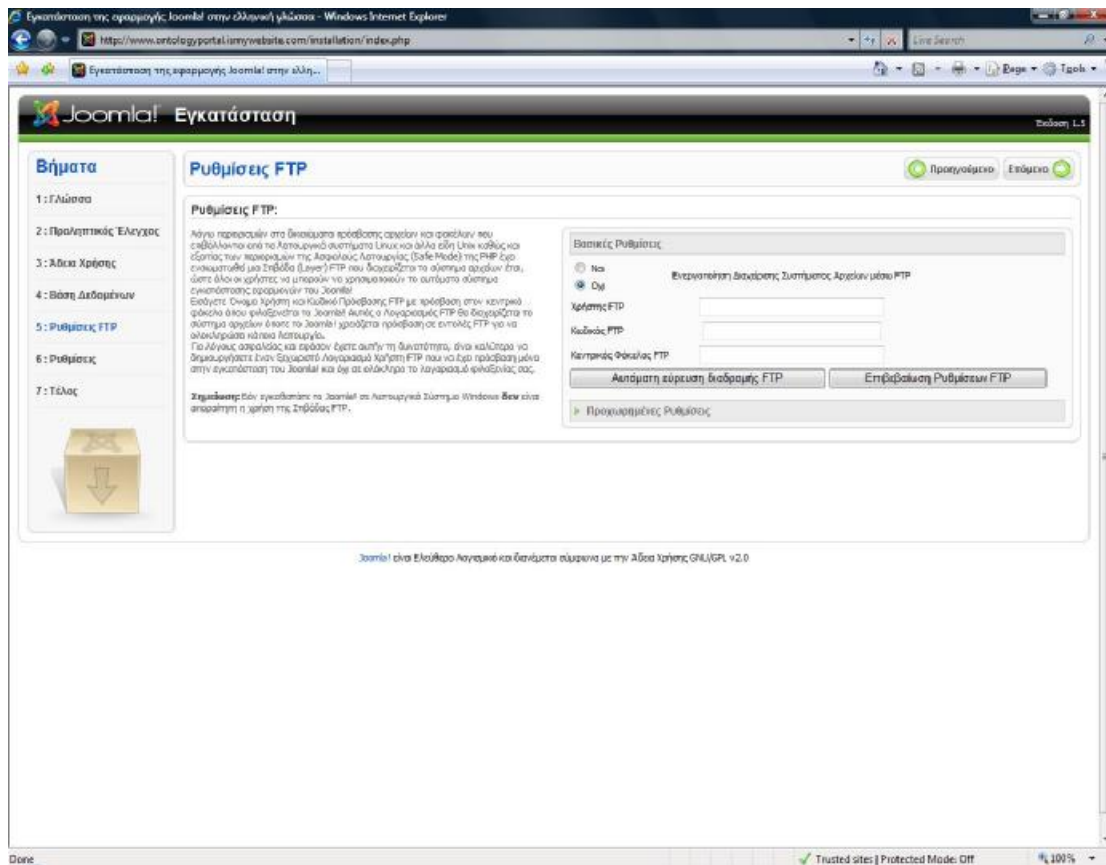
Στη συνέχεια για να ολοκληρώσουμε την εγκατάσταση βάζουμε τα στοιχεία της βάσης δεδομένων.



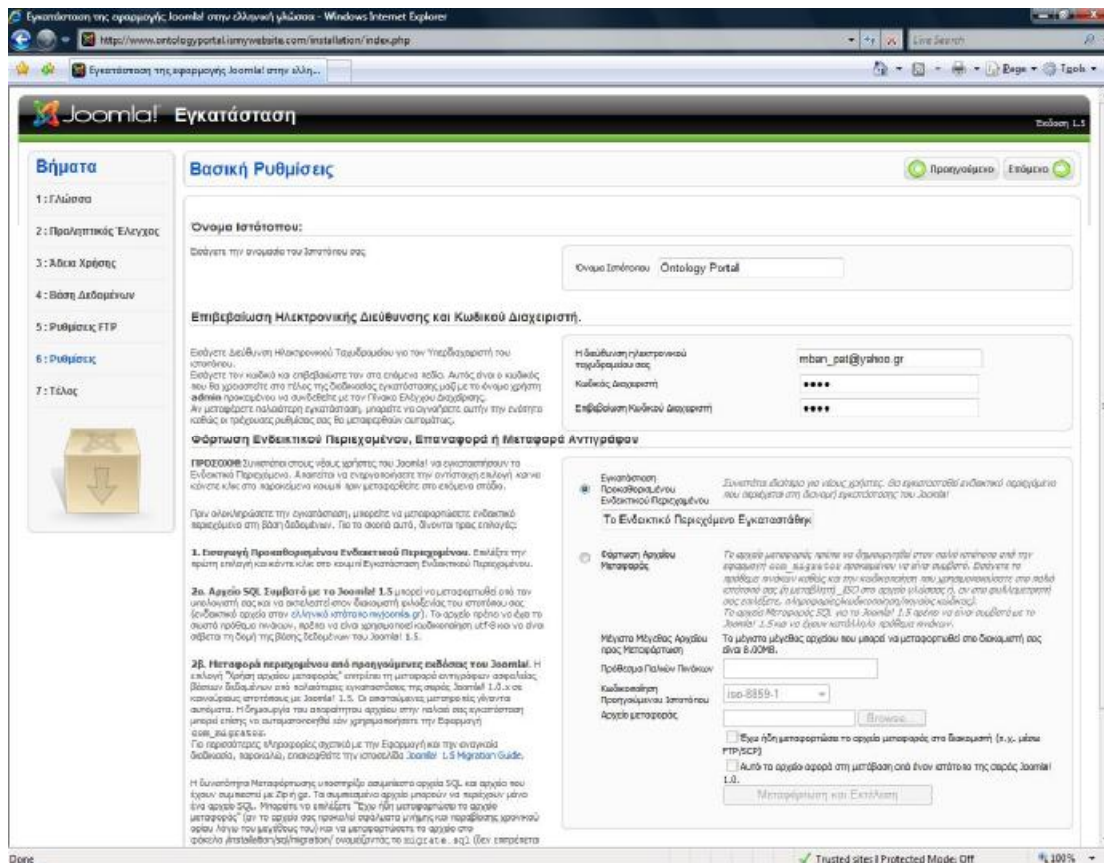
Εικόνα 4.7 Δημιουργία βάσης δεδομένων



Εικόνα 4.8 Ρυθμίσεις βάσης δεδομένων



Εικόνα 4.9 Ρυθμίσεις FTP



Εικόνα 4.10 Βασικές ρυθμίσεις

Μόλις τελειώσουμε την εγκατάσταση, διαγράφουμε τον κατάλογο /installation και το joomla site είναι online.

4.3 Σχεδιασμός του Portal – βασικά τμήματα

Κατά τον σχεδιασμό, κύριο μέλημά μας ήταν η εισαγωγή της μηχανής αναζήτησης, ενός εκ των κυρίων χαρακτηριστικών που απαρτίζει ένα portal, μια εφαρμογή που με την χρήση πρακτόρων λογισμικού επιτρέπει την αναζήτηση κειμένων και αρχείων σε αυτό.

Στη συνέχεια επιλέξαμε ποιές ενότητες και άρθρα θα παρουσιάζαμε στο κύριο μενού. Αποφασίσαμε να περιέχουν πληροφορίες για την χρήση των

οντολογιών στην πληροφορική, τις γλώσσες αναπάστασής τους, όπως επίσης και πληροφορίες για τον σημασιολογικό και παγκόσμιο ιστό. Επίσης μια τελευταία ενότητα είναι αυτή των συνδέμων, όπου δημιουργήσαμε κάποιες χρήσιμες υπερσυνδέσεις.

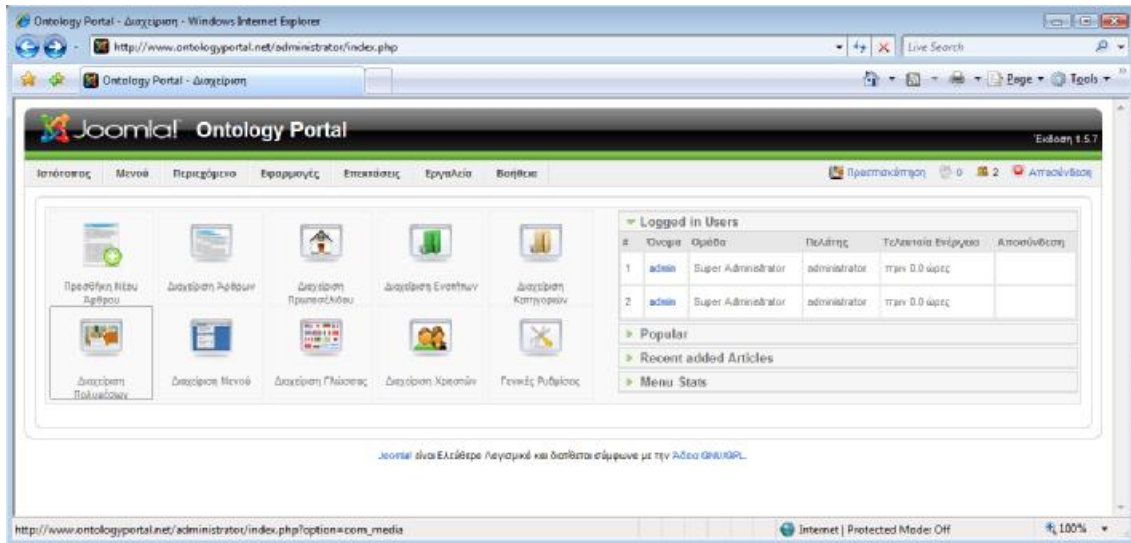
Μέσω της διαχείρισης ενθεμάτων της εφαρμογής Joomla! προσθέσαμε τον οδηγό πλοήγησης και το ένθεμα συνδεδεμένοι χρήστες, που εμφανίζει τον αριθμό των ανωνύμων και εγγεγραμμένων χρηστών που κάθε στιγμή περιηγούνται στον Ιστότοπο, όπως και μια φόρμα συνδεσης με όνομα χρήστη και κωδικό, που επίσης εμφανίζει έναν σύνδεσμο για την ανάκτηση ξεχασμένων κωδικών και έναν άλλο σύνδεσμο όπου θα μπορούν οι χρήστες να εγγράφονται μόνοι τους. Τέλος, δημιουργήσαμε δύο δημοσκοπήσεις για μια προσωπική μας έρευνα, που αφορά στο ποσοστό των επισκεπτών του portal που ήδη γνωρίζει για τις οντολογίες ή ενδεχομένως καλύφθηκε από τις πληροφορίες που δημοσιεύσαμε.

4.4Υλοποίηση Portal – Παραμετροποίηση Joomla!

Αρχικά έγινε η επιλογή του template για το portal. Το Joomla! μας δίνει περιορισμένη δυνατότητα επιλογής template και προτιμήσαμε να εγκαταστήσουμε κάποιο που ταιριάζει περισσότερο με το ύφος του portal μας, από τον ιστότοπο <http://www.siteground.com/joomla-hosting/joomla15-templates.htm?page=2>. Συνεπώς για να γίνει η εγκατάσταση, αποθηκεύουμε τοπικά το αρχείο που επιλέξαμε στο σκληρό δίσκο, το αποσυμπιέζουμε μέσω του προγράμματος Winrar και μεταφέρουμε στο server τα αρχεία με ένα πρόγραμμα FTP χρησιμοποιήθηκε το Filezilla - ftp client). Στη συνέχεια μέσω

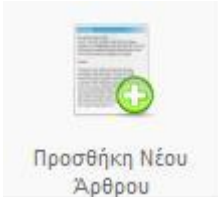
του Joomla!, από "επεκτάσεις" και "διαχείριση προτύπων", το βρίσκουμε στη λίστα και το κάνουμε προκαθορισμένο.

Στην κεντρική σελίδα δίνεται η δυνατότητα στο διαχειριστή μέσω του πίνακα ελέγχου να παραμετροποιήσει και να σχεδιάσει τη τελική μορφή του portal.



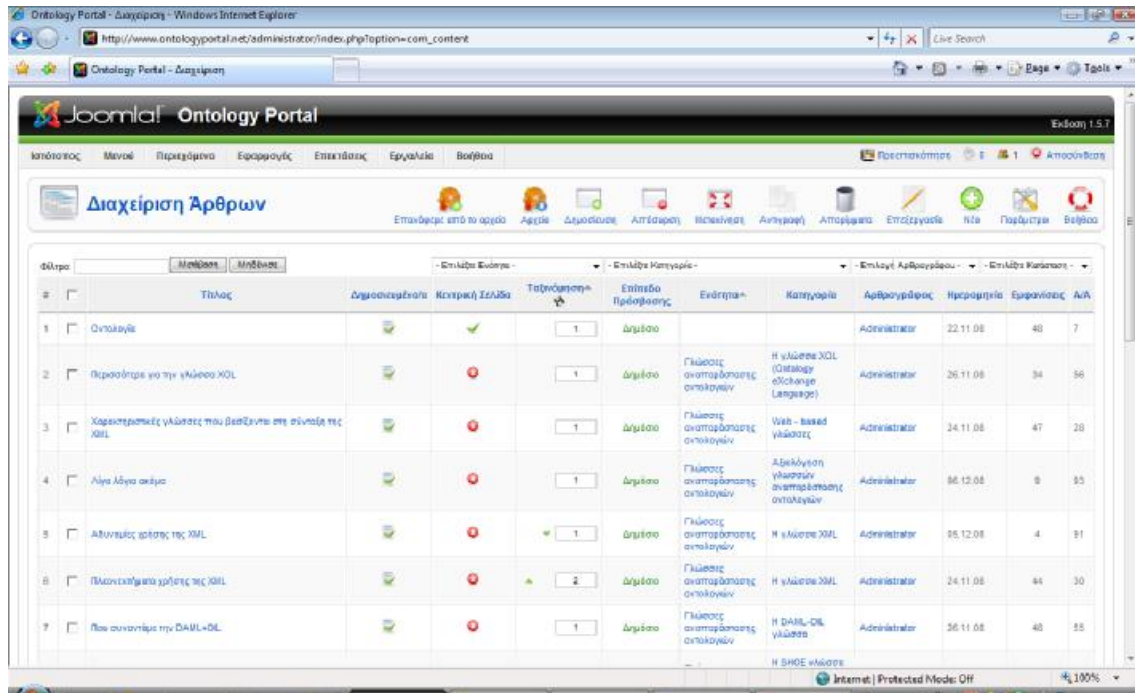
Εικόνα 4.11 Πίνακας ελέγχου

Συγκεκριμένα τα βασικά εικονίδια της διαχείρισης, αποτελούν:

-  Προσθήκη Νέου Άρθρου
Με αυτή τη διαδικασία μπορούμε να προσθέσουμε και να επεξεργαστούμε ένα νέο άρθρο. Μπορούμε επίσης να διαλέξουμε την ενότητα και την κατηγορία όπου υπάγεται το εκάστοτε άρθρο και να καθορίσουμε εάν θα δημοσιευθεί στην κεντρική σελίδα του portal.



- Με την επιλογή αυτή μπορούμε να διαχειριστούμε τη σειρά με την οποία θα εμφανίζονται τα άρθρα.



Εικόνα 4.12 Στιγμιότυπο διαχείρισης άρθρων

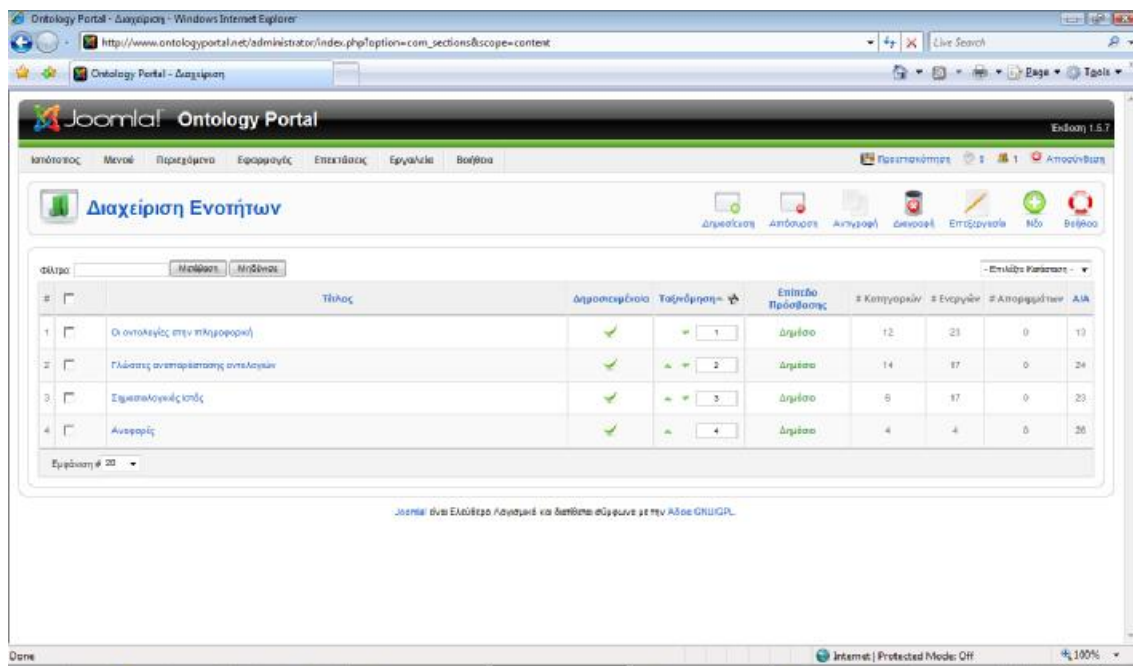


- Εδώ καθορίζουμε ποια άρθρα και με ποια σειρά θα εμφανίζονται στο πρωτοσέλιδο, που συνήθως αποτελεί την αρχική σελίδα του portal.



- Τα άρθρα στο Joomla! οργανώνονται σε ενότητες και κατηγορίες. Οι ενότητες είναι το ανώτερο επίπεδο οργάνωσης και οι

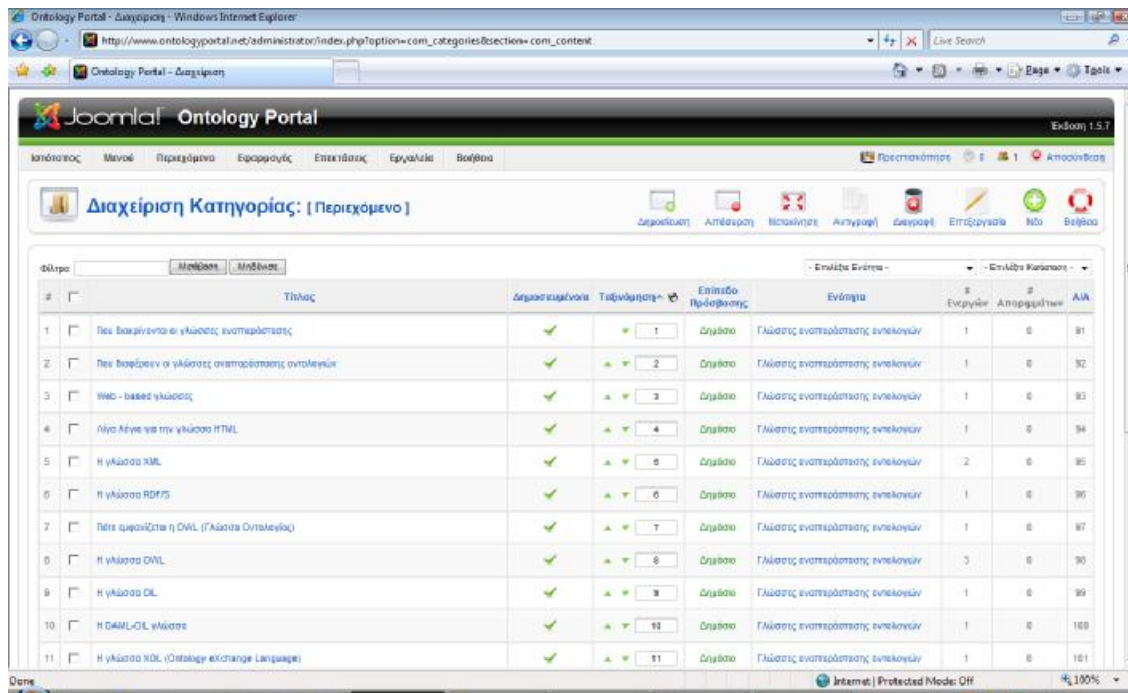
κατηγορίες πάνε κάτω από αυτές. Στην διαχείριση ενότητων επεξεργαζόμαστε ήδη υπάρχουσες ενότητες και δημιουργούμε νέες. Οι ενότητες χρησιμοποιούνται για δύο λόγους. Ο πρώτος είναι για την οργάνωση των άρθρων στο portal. Ο δεύτερος στο να βοηθάει τον διαχειριστή, όταν το πλήθος των άρθρων είναι μεγάλο, να μπορεί να τα φιλτράρει ανά ενότητα και να βλέπει μόνο όσα άρθρα εμπεριέχονται σε αυτήν.




Εικόνα 4.13 Στιγμιότυπο διαχείρισης ενότητων

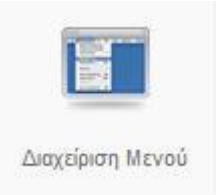



- Επεξεργαζόμαστε ήδη υπάρχουσες κατηγορίες όπως και δημιουργούμε νέες.



Εικόνα 4.14 Στιγμιότυπο διαχείρισης κατηγορίας

- 

Είναι ένα πολύ χρήσιμο εργαλείο για την μεταφόρτωση ή την διαγραφή αρχείων στην τοποθεσία των εικόνων στον server.
- 

Ρυθμίζει το πώς θα φαίνεται στην πραγματικότητα το μενού του Joomla!
- 

Όπου επιλέγουμε την προκαθορισμένη γλώσσα που θα εμφανίζεται στο portal.



Διαχείριση Χρηστών

- Εδώ μας εμφανίζεται η λίστα των εγγεγραμμένων χρηστών, όπου μπορούμε να δημιουργούμε και να επεξεργαζόμαστε τους λογαριασμούς των χρηστών.



Γενικές Ρυθμίσεις

- Επιτρέπει να διαμορφώσουμε το portal με τις προσωπικές μας ρυθμίσεις.

4.5 Παρουσίαση υλοποίησης – οθόνες από τα κύρια τμήματα του portal



Εικόνα 4.15 Πρωτοσέλιδο ιστοτόπου

Ο επισκέπτης του portal, καθοδηγείται από το κύριο μενού “Επιλέξτε κατηγορία” και μέσω των υπερσυνδέσμων καταλήγει στα άρθρα. Λόγου χάριν, στην εικόνα 4.16 παρουσιάζεται η λίστα από άρθρα που αφορούν την αρχιτεκτονική του σημασιολογικού ιστού.

The screenshot shows the 'ontology portal' website. At the top, there is a navigation bar with 'Κύριο Γε Ανάλυση' and 'Αρχήματα'. Below the header, the article title 'Η αρχιτεκτονική του σημασιολογικού ιστού' is displayed. The main content area contains a diagram illustrating the architecture of the Semantic Web, showing layers from 'Διευθυνσιοδότηση URI' and 'Κωδικοποίηση Unicode' at the bottom to 'Διατεθέντα χρήστη & εφαρμογές' at the top. The diagram includes components like 'Αξιωματικά', 'Αποδείξεις', 'Ενσωματωμένη Λογική', 'Αιτήματα SPARQL', 'Οντολογίες OWL', 'Κανόνες RIF', 'RDF Schema (RDFS)', and 'Ανταλλαγή δεδομένων: Resource Description Framework'. A vertical bar on the right is labeled 'εμπειρογένητα'. Below the diagram is a table listing 11 articles with their titles, authors, and view counts.

#	Τίτλος άρθρου	Αρθρογράφος	Προβολές
1	Αξιώματα	Administrator	52
2	Αποδείξεις	Administrator	41
2	Ενσωματωμένη Λογική	Administrator	49
4	Αιτήματα	Administrator	27
5	Κανόνες	Administrator	45
6	Οντολογίες - OWL	Administrator	43
7	Resource Description Framework Schema (RDFS)	Administrator	41
8	Ανταλλαγή Δεδομένων (RDF - Resource Description Framework)	Administrator	46
9	Extensible Markup Language (XML)	Administrator	45
10	Διευθυνσιοδότηση συστάσεων	Administrator	48
11	Κωδικοποίηση χαρακτήρων	Administrator	45

Εικόνα 4.16 Στιγμιότυπο άρθρων “Η αρχιτεκτονική του σημασιολογικού ιστού.”

Στην επόμενη οθόνη παρουσιάζεται η δημοσκόπηση “Γνωρίζετε για την χρήση οντολογιών στην πληροφορική;”. Εμφανίζεται στον χρήστη όταν θα

συμμετάσχει στην ψηφοφορία, ή όταν θα πατήσει στον υπερσύνδεσμο “Αποτελέσματα”.



Εικόνα 4.17 Στατιστικά αποτελέσματα δημοσκόπησης την 10^η Ιανουαρίου του 2009

Παρακάτω έχουμε την φόρμα που καλείται να συμπληρώσει ο χρήστης που επιθυμεί να εγγραφεί στο portal. Εμφανίζεται με την επιλογή “Δημιουργία λογαριασμού”, που βρίσκεται με την μορφή υπορσυνδέσμου στην φόρμα σύνδεσης, επάνω - δεξιά.

The image shows a web registration form for the 'ontology portal'. The header features the 'ontology portal' logo on the left and navigation links for 'Κύριο', 'Το Αιτίωμα', and 'Αφήσεις' on the right. Below the header, there are three main sections:

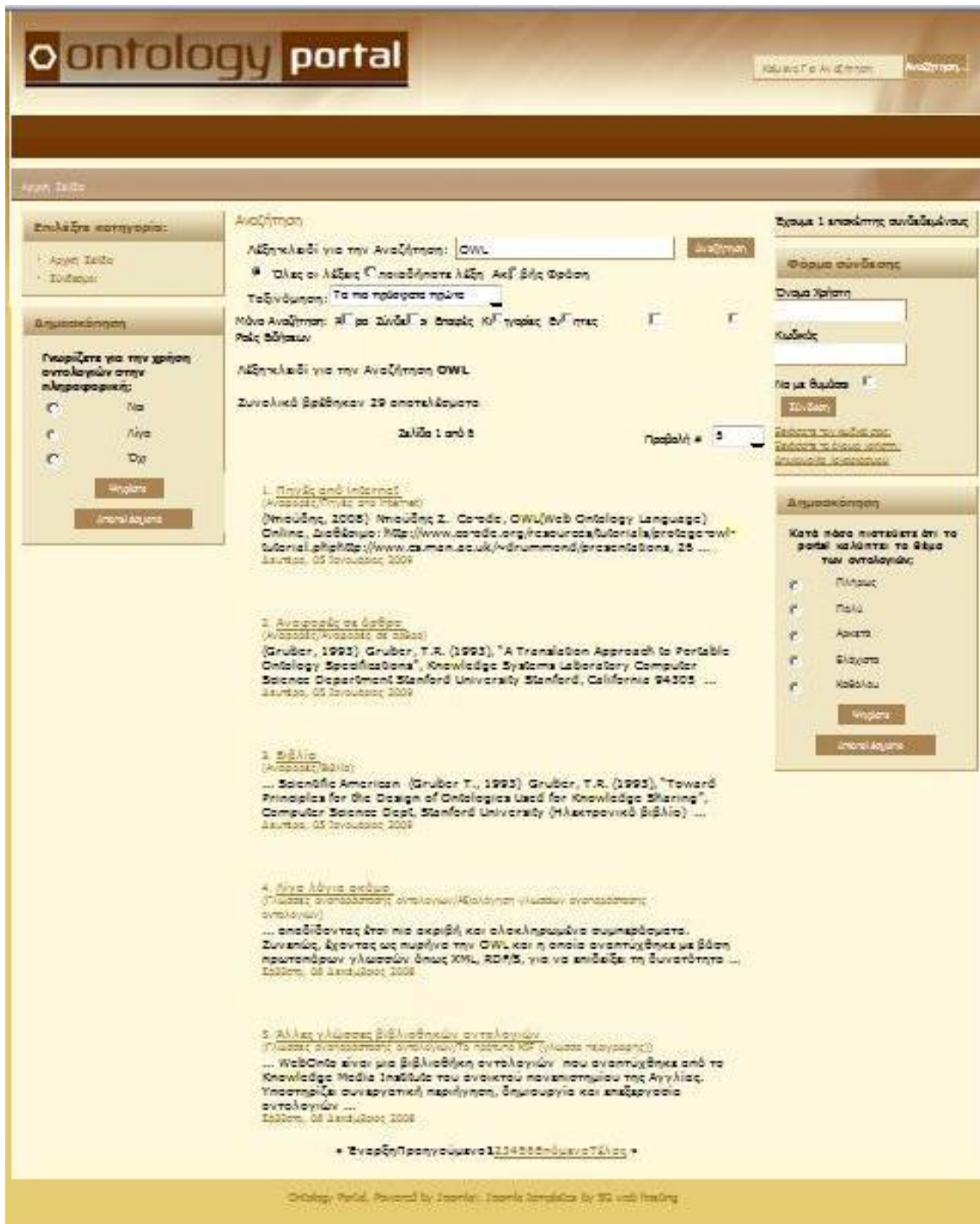
- Επιλέξτε κατηγορία:** A dropdown menu with options 'Ασκήτ. Ξαβό' and 'Σύνταγμα'.
- Δημοσκόπηση:** A section asking for consent to use the portal for informational purposes, with radio buttons for 'Ναι', 'Λίγο', and 'Όχι', and buttons for 'Υποβολή' and 'Αποχή Δέχεται'.
- Εγγραφή:** A registration form with fields for 'Όνομα:', 'Όνομα Χρήστη:', 'Ηλεκτρονικό Ταχυδρομείο:', 'Κωδικός πρόσβασης:', and 'Επιβεβαίωση Κωδικού Πρόσβασης:'. A note states: 'Τα πεδία που σημειώνονται με αστέρικο (*) είναι υποχρεωτικά.' Below these fields is an 'Εγγραφή' button.

On the right side of the form, there is a section for 'Έχετε 1 αποσπαστή συνδεδεμένους' and a 'Φόρμα σύνδεσης' section with fields for 'Όνομα Χρήστη:' and 'Κωδικός:', a 'Να με θυμάσαι' checkbox, and a 'Σύνδεση' button. Below this is another 'Δημοσκόπηση' section asking for trust in the portal's quality, with radio buttons for 'Πολύς', 'Πολύ', 'Αρκετά', 'Ελάχιστο', and 'Καθόλου', and buttons for 'Υποβολή' and 'Αποχή Δέχεται'.

At the bottom of the page, there is a footer: 'Ontology Portal. Powered by Joomla! Joomla! Copyright by 3C web Hosting'.

Εικόνα 4.18 Φόρμα εγγραφής του επισκέπτη

Το επόμενο στιγμιότυπο παρουσιάζει ένα παράδειγμα εφαρμογής της μηχανής αναζήτησης.



Εικόνα 4.19 Αναζήτηση της λέξης “OWL” μέσα στο portal

ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο στόχος της εργασίας αυτής ήταν αφενός μεν η μελέτη και η καταγραφή του πεδίου της έρευνας των οντολογιών και αφετέρου η συγκέντρωση και παρουσίαση της πληροφορίας αυτής σε ένα δυναμικό ιστότοπο – portal του οποίου έγινε η ανάλυση, σχεδίαση και υλοποίηση.

Ένας ηλεκτρονικός υπολογιστής ποτέ δε θα μπορεί πλήρως να κατανοήσει - με την έννοια της τεχνητής νοημοσύνης - τι σημαίνει ένα έγγραφο, είναι δεδομένο. Οι οντολογίες αποτελούν το κλειδί στην δημιουργία του σημασιολογικού ιστού. Βοηθούν δίνοντας την λύση με την οποία οι ηλεκτρονικοί υπολογιστές, συμφωνώντας στο νόημα εννοιών και όρων, μπορούν να ανταλλάσσουν μηνύματα και πληροφορίες.

Στην επιστήμη των υπολογιστών υπάρχει συμφωνία σχετικά με τον ορισμό της οντολογίας οντολογιών η οποία είναι μια αυστηρή, τυπική προδιαγραφή μιας κοινά αποδεκτής εννοιολογικής μορφοποίησης.

Το πεδίο έρευνας των οντολογιών είναι σχετικά νέο και συνεχώς αναπτυσσόμενο, ενώ παρουσιάζεται μεγάλη ποικιλομορφία στα πεδία εφαρμογής και έρευνα από τους σύγχρονους επιστήμονες. Σύμφωνα με τη μελέτη διαπιστώσαμε ότι η αναπαράσταση της πληροφορίας σε μορφή κατανοητή από τα υπολογιστικά συστήματα ανοίγει νέους ορίζοντες στην εξέλιξη εφαρμογών και υπηρεσιών. Το οικοδόμημα του Σημασιολογικού Ιστού, που τώρα θεμελιώνεται, προβάλλει νέους ρόλους στους σύγχρονους διαχειριστές της πληροφορίας - οι υπάρχουσες πληροφορίες και γνώσεις θα μεταφερθούν σε μηχαναγνώσιμη μορφή και θα αποτελέσουν τη μνήμη, τη γνώση που θα αξιοποιηθεί από πλήθος επιστημονικών, επιχειρηματικών και καθημερινών εφαρμογών του μέλλοντος.

Κατά την διαδικασία ανάπτυξης μιας οντολογίας καταρχάς πρέπει να καθοριστεί ο σκοπός ανάπτυξής της και κατόπιν να γίνει η σύλληψη αυτής, δηλαδή η εύρεση των εννοιών της οντολογίας, των σχέσεων μεταξύ των

εννοιών αυτών, ο ορισμός διάφορων αξιωμάτων και τέλος η εισαγωγή των στιγμιότυπων. Η κωδικοποίηση της οντολογίας πρέπει να γίνει σε μια συγκεκριμένη γλώσσα. Οι γλώσσες αναπαράστασης οντολογιών μπορεί να ανήκουν στις web – based γλώσσες, τις γλώσσες που αναπτύχθηκαν για να αναπαραστήσουν συγκεκριμένες οντολογίες και να χρησιμοποιηθούν σε συγκεκριμένες εφαρμογές, τις γλώσσες που αναπτύχθηκαν για να αναπαραστήσουν τη λογική των πλαισίων περιγραφικά.

Οι διαφορές των παραπάνω γλωσσών είναι σχετικά με τη σύνταξη, την ορολογία, την εκφραστικότητα και τη σημασιολογία. Κάτι που μπορούμε να εκφράσουμε σε μία γλώσσα δεν μπορούμε σε μία άλλη και η ίδια δήλωση μπορεί να σημαίνει διαφορετικά πράγματα σε διαφορετικές γλώσσες. Τελικές φάσεις αποτελούν η αξιολόγηση της οντολογίας και η τεκμηρίωσή της.

Η ανάγκη για αποτελεσματικότερη ανάπτυξη και διαχείριση των οντολογιών σε συνεχώς μεταβαλλόμενα περιβάλλοντα, όπως είναι ο σημασιολογικός ιστός, έχει σα σκοπό την ακριβέστερη αναζήτηση της γνώσης που περιέχεται σε διάφορες πηγές πληροφορίας. Η ανάγκη αυτή είναι που ώθησε τη ερευνητική δραστηριότητα προς τον σχεδιασμό και ανάπτυξη μεθοδολογιών μηχανικής οντολογιών και εργαλείων που τις υποστηρίζουν.

Όσον αφορά το τεχνικό μέρος αυτής της πτυχιακής εργασίας, η υλοποίηση του οντολογικού portal έγινε μέσω του συστήματος διαχείρισης περιεχομένου Joomla! Το περιβάλλον αυτού του συστήματος ήταν οικείο και προσιτό και έτσι δεν υπήρξαν ιδιαίτερες δυσκολίες στην παραμετροποίηση του Joomla! και την διαδικασία μεταφοράς των αρχείων του συστήματος σε server. Αρχικά επιλέχθηκε η φιλοξενία του portal να γίνει από το server του ιστοτόπου <http://www.ismywebsite.com>, όπου δίνεται η δυνατότητα να δημιουργηθεί λογαριασμός δωρεάν. Το σημαντικότερο πρόβλημα που αντιμετωπίστηκε και είχε ως αποτέλεσμα την καθυστέρηση ολοκλήρωσης της εργασίας, δημιουργήθηκε μετά την εγκατάσταση του Joomla! στην διεύθυνση <http://www.ontologyportal.ismywebsite.com>, όπου υπήρξαν σοβαρά τεχνικά

προβλήματα στον εξηρηρητητή, με αποτελέσμα να πρέπει να ξεκινήσει από την αρχή η διαδικασία υλοποίησης του portal. Έτσι αποφασίστηκε η ενοκίαση χώρου να γίνει από τον ιστότοπο <http://www.pair.com> και κατοχυρώθηκε URL το <http://www.ontologyportal.net>. Το Joomla!, ένα πλήρες σύστημα διαχείρισης περιεχομένου, έδωσε την δυνατότητα να ολοκληρωθεί το portal φτάνοντας στην επιθυμητή μορφή.

Μέσα από το πόρταλ που αναπτύξαμε επιτυγχάνουμε με δυναμικό τρόπο την παροχή πληροφοριών και αναφορών σχετικά με τις οντολογίες σε κάθε διαδικτυακό επισκέπτη αποτελώντας ένα βασικό σημείο αναφοράς σχετικά με την μελέτη μας επάνω στις οντολογίες.

Έχοντας ως βάση τον Παγκόσμιο Ιστό (WWW) και τη δημιουργία του Σημασιολογικού που ολοένα και περισσότερο εξελίσσονται, το διαδίκτυο σήμερα κινείται σε εφύεστερους και αυτοματοποιημένους τρόπους αναζήτησης και εξαγωγής της πληροφορίας. Οι οντολογίες παρέχουν μια μέθοδο αναπαράστασης της γνώσης και μπορούν να χρησιμοποιηθούν πλέον από τα υπολογιστικά συστήματα για τη διεξαγωγή ακριβή συμπερασμάτων. Ωστόσο όμως η αυξημένη πληροφορία και σε συνδιασμό με το ότι οι μηχανές αναζήτησης είναι ακόμη σε αναπτυσσόμενο και μη ολοκληρωμένο στάδιο, θα πρέπει σε σταθερή βάση να πραγματοποιείται έρευνα πάνω στην ανάπτυξη ενός portal, το οποίο θα είναι πιο ευέλικτο, ακριβές και ομοιογενές, καταστάσεις οι οποίες ακόμη εμφανίζονται ως μειονεκτήματα στην ανάπτυξη μιας οντολογίας.

Μελλοντικές εργασίες μπορούν να επικεντρωθούν είτε στην ύπαρξη του σημασιολογικού ιστού που βρίσκεται ακόμη σε πρώιμο στάδιο και που ολοένα και περισσότερο εξελίσσεται, είτε στην ακριβή ανάπτυξη μιας μεμονωμένης μηχανής αναζήτησης η οποία θα στηρίζεται στις γλώσσες οντολογιών, είτε ακόμη και στις εφαρμογές των οντολογιών στους διάφορους κλάδους των επιστημών γεγονός που τοποθετεί την εφαρμογή τους στη κορυφή. Οι οντολογίες μετά την κοινοπραξία του 2003 μεταξύ σημασιολογικού και

παγκόσμιου ιστού δίνουν νέα τροπή στα μέχρι τώρα δεδομένα και αποτελούν πλέον αναπόσπαστο κομμάτι αυτών.

ΑΝΑΦΟΡΕΣ – REFERENCES

BIBΛΙΑ

(Βλαχάβας et al., 2005) Βλαχάβας Ι., Κεφαλάς Π., Βασιλειάδης Ν., Κόκκορας Φ., Σακελλάριου Η., (2005), “Τεχνητή Νοημοσύνη” Β' Έκδοση, ISBN: 960-7013-35-2, Εκδόσεις Γαρταγάνη

(Αυτομ, 2006) “Επιχειρηματικά Συστήματα Αυτοματισμού Γραφείου” – Σημειώσεις μαθήματος Αυτοματισμού - 7^ο εξάμηνο – ΑΤΕΙ Πατρών (2006)

(Berners-Lee et al., 2001) Berners-Lee T., Hendler J., Lassila O. (May 2001), “The Semantic Web”, Scientific American

(Gruber T., 1993) Gruber, T.R. (1993), “Toward Principles for the Design of Ontologies Used for Knowledge Sharing”, Computer Science Dept, Stanford University

(Ηλεκτρονικό βιβλίο)

(Uschold et al., 1995) Uschold M, King M (1995), “Towards a Methodology for Building Ontologies.” Workshop on Basic Ontological Issues in Knowledge Sharing.

(Gruninger et al., 1995) Gruninger M, Fox M.S. (1995) “Methodology for the Design and Evaluation of Ontologies”, Proceedings of IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal Canada.

(Pinto et al., 2004) Pinto S., Staab S., Tempich C., Diligent (2004), “Towards a fine – grained methodology for Distributed, Loosely – controlled and evolving Engineering ontologies.”, Spain.

(Pinto and Martins, 2004) Pinto S., and Martins P., (2004), “Ontologies: How can they be built? ”, Knowledge and Information Systems, London, σελ. 441-464.

(Noy et al., 2000) Noy N. F., Fergerson, R. W., Musen. M. A. (2000), “The knowledge model of protege-2000: Combining interoperability and flexibility.”
In 12th International Conference on Knowledge Engineering and Knowledge Management (EKAW'2000), Juan-les-Pins, France

(Noy et al., 2001) Noy N. F., Sintek M., Decker S., Crubezy M , Fergerson R. W., Musen M. A. (2001), “Creating semanticweb contents with protege-2000.”
IEEE Intelligent Systems, 16(2):60–71

(Berners-Lee et al., 2000) Berners-Lee T., Cailliau R., Groff J-F. , Pollermann B., (2000) “World-Wide Web: The Information Universe”, CERN.

ΑΝΑΦΟΡΕΣ ΣΕ ΑΡΘΡΑ

(Gruber, 1993) Gruber, T.R. (1993), “A Translation Approach to Portable Ontology Specifications”, Knowledge Systems Laboratory Computer Science Department Stanford University Stanford, California 94305

(Guarino, 1998) Guarino, N. (1998), “Formal Ontology and Information Systems”, Proceedings of FOIS’98, Trento, Italy

(Fernandez-Lopez et al., 1999) Fernandez-Lopez. M., Gornez-Perez, Rojas A., (1999), “Ontologies crossed life cycles.” *Workshop on Knowledge Acquisition, Modelling and Management (EKAW)*., Editor Springer Verlag. Jean Les Pins (Francia). σελ. 65-79.

(Sure et al., 2002) Sure Y., Erdmann M., Angele J., Staab S., Studer R. and Wenke D. (2002), “OntoEdit: Collaborative Ontology Engineering for the Semantic Web.”, *In Proceedings of the International Semantic Web Conference*, June 2002, Sardinia, Italia.

(Giunchiglia et al., 2004) Giunchiglia F., Shvaiko P., Yatskevich M. (2004), “S-Match: An Algorithm and Implementation of Semantic Matching.” *The Semantic Web: Research and Applications*, Vol. 3053, σελ. 61-75 (Lecture Notes in Computer Science)

(Kalfoglou et al., 2003) Kalfoglou Y., Schorlemmer M. (2003), “Ontology mapping: the state of the art.”, *The Knowledge Engineering Review* 18(1): 1—31

(Serafini et al., 2003) Serafini L., Bouquet P., Magnini B., Zanobini S. (2003), “An Algorithm for Matching Contextualized Schemas via SAT.” *In Proceeding of CONTEX 03*

(Noy, 2003) N. Noy. (2003), “What do we need for ontology integration on the Semantic Web.” *Proceedings of the Semantic Integration Workshop Collocated*

with the Second International Semantic Web Conference (ISWC -03) Sanibel Island, Florida

(Antoniou et al., 2004) Antoniou G., Van Harmelen F., (2004), “A Semantic Web Primer”, Massachusetts , MIT Press.

(Berners-Lee, 2006) Berners-Lee T. (2006), “Artificial Intelligence and the Semantic Web”. *AAAI*, 18 July 2006.

(Carroll et al., 2004) Carroll J.J., Kline G., (10 February 2004) “Resource Description Framework (RDF):Concepts and Abstract Syntax”, *W3C Recommendation*,

(Horridge et al., 2004) Horridge M., Knublauch H., Rector A., Stevens R., Wroe C., (27 Αυγούστου 2004) “A Practical Guide to Building OWL Ontologies Using the Protégé-OWL Plugin and CO-ODE Tools”, The University of Manchester, Edition 1.0

ΠΗΓΕΣ ΑΠΟ INTERNET

(Ντιούδης, 2008) Ντιούδης Σ. Co-ode, *OWL(Web Ontology Language)* Online, Διαθέσιμο: <http://www.co-ode.org/resources/tutorials/protege-owl-tutorial.php>

<http://www.cs.man.ac.uk/~drummond/presentations>, 26 Αυγούστου 2008

(Wiki, 2008) Wikipedia, *Ontology (computer science)*. Online. Διαθέσιμο:[http://en.wikipedia.org/wiki/Ontology %28computer science%29](http://en.wikipedia.org/wiki/Ontology_%28computer_science%29) , 5 Μαΐου 2008

(WikiMer, 2008) Wikipedia, *Meronymy* Online. Διαθέσιμο: <http://en.wikipedia.org/wiki/Meronymy>, 3 Μαΐου 2008

(Berkeley, 2005) UC Berkeley Digital Library Project. Online. Διαθέσιμο: <http://elib.cs.berkeley.edu/>, 13 Σεπτεμβρίου 2008

(ISSEL, 2006) The Intelligent Systems and Software Engineering Labgroup (ISSEL) belongs to the Information Processing Laboratory (IPL) of the Department of Electrical and Computer Engineering at the Aristotle University of Thessaloniki. Online. Διαθέσιμο: <http://issel.ee.auth.gr/el/semanticweb>, 12 Σεπτεμβρίου 2008

(WikiSoft, 2008) Wikipedia, *Software agent*, Online. Διαθέσιμο: [http://en.wikipedia.org/wiki/Software agents](http://en.wikipedia.org/wiki/Software_agents), 23 Μαΐου 2008

(WikiMob, 2008) Wikipedia, *Mobile agent*, Online. Διαθέσιμο: [http://en.wikipedia.org/wiki/Mobile agent](http://en.wikipedia.org/wiki/Mobile_agent), 23 Μαΐου 2008

(WikiChar, 2008) Wikipedia, *Character encoding*, Online. Διαθέσιμο: [http://en.wikipedia.org/wiki/Character encoding](http://en.wikipedia.org/wiki/Character_encoding), 28 Μαΐου 2008

(WikiRule, 2008) Wikipedia, *Rule of inference*. Online. Διαθέσιμο: [http://en.wikipedia.org/wiki/Rule of inference](http://en.wikipedia.org/wiki/Rule_of_inference), 27 Μαΐου 2008

(Cover, 2005) Robin Cover, *Business Rules and Web Architecture: W3C Creates Rule Interchange Format WG*, Online. Διαθέσιμο: <http://xml.coverpages.org/ni2005-11-09-a.html>, 7 Οκτωβρίου 2008

(WikiXml, 2008) Wikipedia, *XML*, Online. Διαθέσιμο: <http://en.wikipedia.org/wiki/XML>, 17 Ιουλίου 2008

(W3C, 2004) *Η κοινοπραξία του Παγκόσμιου Ιστού εκδίδει τι; Συστάσεις RDF και OWL*, Online. Διαθέσιμο: <http://www.w3c.gr/office/pressreleases/2004/02/sws-pressrelease.el.html>, 2004

(Mc Guinness et al., 2004) Mc Guinness D. L., Harmelen F. *OWL Web Ontology Language Overview*. W3C Recommendation, Online. Διαθέσιμο: <http://www.w3.org/TR/owl-features/>, 10 Φεβρουαρίου 2004

(Smith et al., 2004) Smith M.K. , Welty C. , Mc Guinness D.L., *OWL Web Ontology Language Guide*, Online. Διαθέσιμο: <http://www.w3.org/TR/owlguide/>, Φεβρουάριος 2004

(XOL, 2008) *XOL Ontology Exchange Language*, Online. Διαθέσιμο: <http://www.ai.sri.com/pkarp/xol/>, Οκτώβρης 2008

(SHOE, 2008) SHOE, Online. Διαθέσιμο: <http://www.cs.umd.edu/projects/plus/SHOE/>, Οκτώβρης 2008

(WikiServer, 2008) Web server, Online. Διαθέσιμο: http://en.wikipedia.org/wiki/Web_server, Δεκέμβρης 2008

(WikiApache, 2008) Apache HTTP, Online. Διαθέσιμο:
[http://en.wikipedia.org/wiki/Apache HTTP](http://en.wikipedia.org/wiki/Apache_HTTP), Δεκέμβρης 2008

ΕΡΓΑΣΙΕΣ – ΔΗΜΟΣΙΕΥΣΕΙΣ

(Γιουβανάκης Α., 2006) Γιουβανάκης Α. (Νοέμβριος 2006), “Μεθοδολογία και Πλαίσιο Ανάπτυξης Προσαρμοσμένου Συστήματος Ηλεκτρονικής Μάθησης σε Περιβάλλον Σημασιολογικού Ιστού.”, *Διδακτορική Διατριβή*, Πανεπιστήμιο Μακεδονίας, Θεσσαλονίκη.

(Τόγιας Κ.Ι., 2007) Τόγιας Κ.Ι. (2007), “Αξιολόγηση Οντολογιών Μαθησιακού Υλικού”, *Διατριβή Μεταπτυχιακού Διπλώματος Ειδίκευσης*, Πανεπιστήμιο Πατρών, Πάτρα.