

**Τ.Ε.Ι. ΠΑΤΡΩΝ**  
**ΣΧΟΛΗ: ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ**  
**ΤΜΗΜΑ: ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΠΛΗΡΟΦΟΡΙΑΚΩΝ**  
**ΣΥΣΤΗΜΑΤΩΝ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΘΕΜΑ:**  
**ΟΙ ΟΝΤΟΛΟΓΙΕΣ ΩΣ ΜΕΣΟ ΑΝΑΠΑΡΑΣΤΑΣΗΣ ΚΑΙ ΑΝΑΖΗΤΗΣΗΣ ΣΤΟ**  
**ΣΗΜΑΣΙΟΛΟΓΙΚΟ ΙΣΤΟ**

**ΣΠΟΥΔΑΣΤΕΣ:**  
**ΑΝΑΣΤΑΣΙΟΥ ΣΩΤΗΡΙΟΣ**  
**ΤΣΟΓΚΑ ΠΑΡΘΕΝΙΑ**

**ΕΠΙΒΛΕΠΩΝ: ΔΑΡΣΙΝΟΣ ΒΑΣΙΛΕΙΟΣ**

**Πάτρα, Νοέμβριος 2008**



# ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ.....	3
ΕΙΣΑΓΩΓΗ.....	6
<b>ΚΕΦΑΛΑΙΟ 1</b>	
1.1 Παγκόσμιος ιστός.....	7
1.2 Σημασιολογικός Ιστός(Semantic Web.....	8
1.3 Η Σχέση του Σημασιολογικού Ιστού με τον Παγκόσμιο Ιστό.....	10
1.4 Το Όραμα του Σημασιολογικού Διαδικτύου.....	11
1.5 Δικτυακές υπηρεσίες του σημασιολογικού ιστού.....	12
1.6 Σημασιολογικά εργαλεία.....	13
1.7 Οι έξυπνοι πράκτορες στο σημασιολογικό ιστό.....	15
1.8 Μηχανές αναζήτησης.....	16
1.9 robust hyperlinks .....	17
1.10 Στόχοι του Σημασιολογικού Ιστού.....	18
1.11 Οι τεχνολογίες που χρησιμοποιεί το Semantic Web.....	19
1.12 Η αρχιτεκτονική του σημασιολογικού ιστού σε στρώματα.....	21
1.13 Η σημασιολογία στην επιστήμη των υπολογιστών.....	23
1.14 Το ζήτημα των οντολογιών στο Σημασιολογικό Ιστό.....	32
<b>ΚΕΦΑΛΑΙΟ 2</b>	
2.1 Οντολογίες.....	33
Εισαγωγή στις οντολογίες.....	33
2.2 Βασική παρουσίαση των οντολογιών .....	33
2.3 Λόγοι ανάπτυξης οντολογιών .....	37
2.4 Τι είναι Οντολογία.....	43
2.5 Ορισμός της οντολογίας.....	44
2.6 Η ανάπτυξη οντολογιών αφορά.....	45

2.7	Σχεδιασμός Οντολογιών –Αντικειμενοστραφής σχεδιασμός.....	45
2.8	Κατηγορίες Οντολογιών.....	46
2.9	Χαρακτηριστικά της οντολογίας.....	46
2.10	περιοχές εφαρμογής οντολογιών .....	48
2.10.1.	Περιοχή εφαρμογής: Διαχείριση Γνώσης.....	49
2.10.2.	Περιοχή εφαρμογής: Ηλεκτρονικό Εμπόριο.....	50
2.11.	Το πρόβλημα – λύση.....	52
2.12.	Μεθοδολογίες ανάπτυξης οντολογιών.....	53
2.12.1	SEAL (SEmantic PortAL).....	53
2.12.2	IDEF5 (ICAM).....	55
2.12.3.	TOVE (Toronto Virtual Enterprise).....	56
2.12.4	BPD/D.....	57
2.12.5	METHONTOLOGY.....	57
2.12.6.	CEO (Core Enterprise Ontology).....	58
2.12.7.	SENSUS.....	59
2.12.8	DECOR Methodology.....	61
2.13.	Εργαλεία ανάπτυξης οντολογιών.....	62
2.13.1	Apollo.....	63
2.13.2.	LinKFactory.....	65
2.13.3	OilEd.....	67
2.13.4	OntoEdit.....	69
2.13.5	Ontolingua Server.....	70
2.13.6	OntoSaurus.....	72
2.13.7	OpenKnoME.....	73
2.13.8	Protégé.....	75
2.13.9	SymOntoX.....	77
2.13.10	WebODE.....	80
2.13.11	WebOnto.....	81
2.14	Εργαλεία αξιολόγησης οντολογιών.....	83
2.14.1	OntoAnalyser.....	83
2.14.2	OntoGenerator.....	84

2.14.3 OntoClean.....	85
2.14.4 ONE-T.....	87
2.15 Γλώσσες αναπαράστασης οντολογιών.....	90
2.15.1 OIL (Ontology Inference Layer) .....	90
2.15.2 LOOM.....	91
2.15.3 KIF (Knowledge Interchange Format).....	91
2.15.4 OWL (Ontology Web Language).....	92
2.15.5 RDF(Resource Description Framework).....	98
2.15.6 RDFS .....	104
2.15.7 RDF Data Query Language (RDQL).....	112
2.15.8 XML.....	115
2.16 Γλώσσες Περιεχομένου και Οντολογίες στη JADE.....	116
2.17 Σημασιολογικός ιστός και ηλεκτρονικές επιχειρήσεις.....	119
2.18 Συμπεράσματα.....	122
<b>BIBΛΙΟΓΡΑΦΙΑ.....</b>	<b>128</b>

## ΕΙΣΑΓΩΓΗ

Η λέξη "σημασιολογία" έχει ως ρίζα τις Ελληνικές λέξεις "σημάδι", "σημαίνω" και "σημαντικός" και σήμερα αναφέρεται στην απόδοση νοήματος στα σύμβολα μιας γλώσσας. Το *Σημασιολογικό Διαδίκτυο (Semantic Web)* αυτή τη στιγμή αποτελεί τη μεγαλύτερη προσπάθεια αυτόματης ενοποίησης συστημάτων ώστε να συνεργάζονται διαλειτουργικά, σε παγκόσμιο επίπεδο. Ο Tim Berners-Lee, που επινόησε τον Παγκόσμιο Ιστό το 1989, είχε το όραμα ενός ιστού δεδομένων αυτόματα επεξεργάσιμων από τις εφαρμογές, βάσει του νοήματος και όχι της μορφής της πληροφορίας.

Το κέντρο βάρους των περιεχομένων του διαδικτύου μετατοπίζεται συνεχώς από το ελεύθερο κείμενο που είναι πλήρως κατανοητό μόνο από τον άνθρωπο, προς την ημιδομημένη ή και πλήρως δομημένη πληροφορία η οποία μπορεί να γίνει αυτόματα κατανοητή από διαδικτυακές εφαρμογές, όπως είναι οι διαδικτυακές υπηρεσίες.

Για να φτάσει το διαδίκτυο στο μέγιστο των δυνατοτήτων του, πρέπει να εξελιχθεί σε μια τέτοια μορφή στην οποία να παρέχει μια παγκοσμίως προσβάσιμη πλατφόρμα που να επιτρέπει σε αυτοματοποιημένα εργαλεία (πράκτορες) να διαμοιράζονται και να επεξεργάζονται πληροφορίες και δεδομένα για λογαριασμό των ανθρώπων-χρηστών τους. Το Σημασιολογικό Διαδίκτυο αποτελεί πρωτοβουλία της *Κοινοπραξίας του Παγκοσμίου Ιστού (W3C)* και η σχετική δραστηριότητα (*Semantic Web Activity*) έχει δημιουργηθεί για να διαδραματίσει έναν ηγετικό ρόλο, τόσο στο σχεδιασμό προδιαγραφών, όσο και στην ανοικτή ανάπτυξη της τεχνολογίας μέσω της συνεργασίας. Εδώ θα πρέπει να διευκρινιστεί ότι η κυριολεκτική μετάφραση του αγγλικού όρου *Semantic Web* είναι *Σημασιολογικός Ιστός*. Όμως η τεχνολογία που αναπτύσσεται και το εύρος των εφαρμογών καλύπτουν και άλλους τομείς του διαδικτύου, όπως είναι για παράδειγμα το ηλεκτρονικό ταχυδρομείο. Επίσης, είναι γεγονός πώς το *Διαδίκτυο (Internet)* γίνεται πλέον αντιληπτό στο χρήστη σχεδόν αποκλειστικά μέσω του παγκόσμιου ιστού. Σαν αποτέλεσμα, οι δύο όροι *ιστός* και *διαδίκτυο* χρησιμοποιούνται σχεδόν ταυτόσημα.

# ΚΕΦΑΛΑΙΟ 1

## 1.1 Παγκόσμιος ιστός

Ο Παγκόσμιος Ιστός αποτελεί ένα μοναδικό και χωρίς προηγούμενο φαινόμενο επιτυχίας τόσο σε επίπεδο διαθέσιμης πληροφορίας όσο και σε αριθμό χρηστών. Για αυτό το λόγο έχει καταστεί η μεγαλύτερη πηγή πληροφορίας στις μέρες μας. Και η επιτυχία του αυτή συνεχίζει με αποτέλεσμα κάθε χρόνο να προστίθενται στον Ιστό 1 έως 2 exabytes καινούρια δεδομένα. Είναι όμως αυτή η επιτυχία του η οποία υποσκάπτει τα θεμέλια του; Η έλλειψη κοινής ορολογίας, οργάνωσης και σημασιολογίας στα δεδομένα που διακινούνται τον Ιστό δημιουργεί πολύ μεγάλη δυσκολία στην ανταλλαγή και στην επεξεργασία των δεδομένων. Ακόμη σημαντικές αδυναμίες του σημερινού Ιστού είναι ότι η πληροφορία που είναι διαθέσιμη δεν είναι κατανοητή από τους υπολογιστές αφού τα δικτυακά δεδομένα είναι σχεδιασμένα για κατανόηση μόνο από ανθρώπους και σχεδόν όλα τα μεταδεδομένα (HTML) που χαρακτηρίζουν σήμερα το περιεχόμενο του Παγκόσμιου Ιστού αναφέρονται σε πληροφορίες παρουσίασης (π.χ. μέγεθος και χρώμα γραμματοσειράς) και σε διασυνδέσεις σε σχετικό περιεχόμενο. Παρά τις έντονες προσπάθειες για βελτίωση και αυτοματοποίηση της συλλογής και της επεξεργασίας της πληροφορίας στο Διαδίκτυο, η πρόοδος που έχει σημειωθεί μέχρι σήμερα στον τομέα αυτό είναι πολύ μικρή. Στο σημείο αυτό έρχεται ο Σημασιολογικός Ιστός, ο οποίος σύμφωνα με τον Tim Berners Lee του W3C, εμπνευστή του διαδικτύου, θα αποτελέσει τη νέα γενιά του διαδικτύου. Ο Σημασιολογικός Ιστός δεν είναι ένας ξεχωριστός ιστός αλλά επέκταση του σημερινού Παγκόσμιου Ιστού, στον οποίο η πληροφορία είναι καλά καθορισμένη κάνοντας καλύτερη τη συνεργασία ανθρώπων και υπολογιστών. Ο Σημασιολογικός Ιστός θα είναι ένα σύνολο από πηγές και συνδέσεις και οι υπολογιστές θα μπορούν να παράγουν συμπεράσματα έχοντας πρόσβαση σε κατανοητή για αυτούς γνώση. Το μυστικό του Σημασιολογικού Ιστού δεν είναι κάποιες προχωρημένες τεχνικές Τεχνητής Νοημοσύνης εφαρμοσμένες σε δίκτυα. Το μυστικό του Σημασιολογικού

Ιστού είναι τα μεταδεδομένα τα οποία κάνουν σαφές το υπονοούμενο (“making explicit the implicit” Tim Berner Lee). Τα μεταδεδομένα θα καταστήσουν τα δεδομένα κατανοητά σε επίπεδο μηχανής ώστε να διευκολύνεται η επεξεργασία και η αναζήτηση τους, θα επιτρέψουν τη δημιουργία κοινής ορολογίας, οργάνωσης και σημασιολογίας στα δεδομένα που διακινούνται στον Ιστό, στο βαθμό που είναι εφικτό, και θα προσδώσουν ευελιξία στη σύνθεση της σημασιολογικής πληροφορίας και στη δημιουργία νέας. Ο Σημασιολογικός Ιστός βασίζεται στις Οντολογίες. Μια Οντολογία είναι ένα σύνολο από δεδομένα με σημασιολογικό περιεχόμενο. Μια Οντολογία αποτελείται από έννοιες (ή κλάσεις), που είναι οτιδήποτε μπορεί να ειπωθεί για αυτό, από σχέσεις, που αναπαριστούν τις αλληλεπιδράσεις μεταξύ των εννοιών από αξιώματα, που περιγράφουν προτάσεις οι οποίες είναι πάντα αληθείς και από στιγμιότυπα που αναπαριστούν συγκεκριμένα στοιχεία. Μια Οντολογία έχει διπλό ρόλο, από τη μια εξασφαλίζει ένα κοινό λεξιλόγιο και από την άλλη εξασφαλίζει μια κοινή κατανόηση του τομέα γνώσης τον οποίο μοντελοποιεί.[25]

## 1.2 Σημασιολογικός Ιστός(Semantic Web)

Το Semantic Web είναι ένα πρόγραμμα που σκοπεύει να δημιουργήσει ένα καθολικό μέσο για την ανταλλαγή πληροφοριών δίνοντας νόημα (σημασιολογία) στο περιεχόμενο των εγγράφων στο Web, κατά τρόπο κατανοητό από τους υπολογιστές. Υπό την καθοδήγηση του δημιουργού του Web, Tim Berners-Lee του World Wide Web Consortium (W3C), το Semantic Web επεκτείνει τις δυνατότητες του World Wide Web μέσω της χρήσης προτύπων, γλωσσών επισήμανσης (mark-up) και σχετικών εργαλείων επεξεργασίας.[3]

Ο σημασιολογικός Ιστός παρέχει ένα κοινό πλαίσιο που επιτρέπει στα δεδομένα να διαμοιράζονται και να επαναχρησιμοποιούνται πέρα από τα όρια μιας εφαρμογής, μιας επιχείρησης ή μίας κοινότητας. Είναι μια συλλογική προσπάθεια της οποίας ηγείται το W3C με τη συμμετοχή ενός μεγάλου αριθμού ερευνητών και βιομηχανικών συνεργατών. Είναι βασισμένο στο Resource Description Framework (RDF), το οποίο ενσωματώνει



ποικίλες εφαρμογές χρησιμοποιώντας XML για τη σύνταξη και URIs για την ονοματοδοσία.[3]

Το 2001 ο Tim Berners-Lee έγραψε:

"Το Semantic Web είναι μια επέκταση του τρέχοντος Web, στον οποίο δίνεται ένα καθορισμένο με σαφήνεια νόημα στις πληροφορίες, δίνοντας τη δυνατότητα σε υπολογιστές και ανθρώπους να συνεργάζονται καλύτερα."

Το μεγαλύτερο μέρος του περιεχομένου του Ιστού σήμερα σχεδιάζεται για να διαβαστεί από ανθρώπους, όχι για να το χειριστούν εννοιολογικά τα προγράμματα υπολογιστών. Βασίζεται κυρίως σε έγγραφα γλώσσας HTML η οποία χρησιμεύει στην περιγραφή ενός δομημένου κειμένου, με έμφαση στην οπτική απεικόνιση. Η HTML έχει περιορισμένες ικανότητες όσον αφορά την ταξινόμηση των τμημάτων κειμένου μιας σελίδας. Οι υπολογιστές μπορούν αποδοτικά να αναλύσουν ιστοσελίδες ως προς την εμφάνιση και τις ρουτίνες, αλλά γενικά, οι υπολογιστές δεν έχουν συνήθως κάποιον αξιόπιστο τρόπο να επεξεργαστούν τη σημασιολογία (semantics) που περιέχεται στη σελίδα αυτή.

Για παράδειγμα, οι μηχανές αναζήτησης εκτελούν απλές λεξιλογικές αναλύσεις που βασίζονται σε συμβολοσειρές αναζήτησης και παράγουν λίστες αποτελεσμάτων που απαιτούν την εξαγωγή συμπερασμάτων από τους ανθρώπινους χρήστες, όσον αφορά το ποια από τα δεδομένα είναι τα κατάλληλα για την θεματική περιοχή που ερευνάται. Χρειάζεται δηλαδή η λογική ανάλυση εκ μέρους του χρήστη για να βρεθούν τα συναφή συμπεράσματα. Αυτό συμβαίνει διότι οι ιστοσελίδες δεν περιέχουν αρκετή πληροφορία σχετικά με τη σημασιολογία των περιεχομένων τους οπότε το λογισμικό του χρήστη δεν μπορεί να βγάλει συμπεράσματα από το περιεχόμενο της σελίδας. Για να αντιμετωπιστεί αυτό το πρόβλημα, οι τεχνολογίες του Semantic Web δίνουν τη δυνατότητα σε αυτοματοποιημένα εργαλεία να εξάγουν αυτά τα συμπεράσματα.

Το Semantic Web επιχειρεί να δομήσει το σημασιολογικό περιεχόμενο των ιστοσελίδων, δημιουργώντας ένα περιβάλλον όπου οι πράκτορες λογισμικού που περιπλανώνται από σελίδα σε σελίδα μπορούν εύκολα να εκτελέσουν τους περίπλοκους στόχους των χρηστών τους.[29] Έτσι, αποζητείται ο μετασχηματισμός του Web από ένα σύνολο στατικών σελίδων σε ένα δίκτυο δυναμικών παροχών υπηρεσιών (Web services) που

ανακαλύπτουν αυτόματα τις αναζητούμενες πληροφορίες, διαπραγματεύονται για αγαθά που ο χρήστης σκοπεύει να αγοράσει ή συγκεντρώνουν πληροφορίες από διαφορετικές πηγές και τις συνενώνουν σε ομοιογενείς μορφές.

Για να λειτουργήσει το Semantic Web, πρέπει οι υπολογιστές να έχουν πρόσβαση σε δομημένες συλλογές πληροφοριών και σύνολα λογικών κανόνων που μπορούν να χρησιμοποιήσουν για να διεξάγουν την αυτοματοποιημένη εξαγωγή συμπερασμάτων. Οι ερευνητές της Τεχνητής Νοημοσύνης έχουν μελετήσει τέτοια συστήματα πολλά χρόνια πριν. Η Αναπαράσταση Γνώσης (Knowledge Representation), όπως καλείται αυτή η τεχνολογία, δεν έχει ακόμα χρησιμοποιηθεί επαρκώς ώστε να επιδείξει πλήρως τις δυνατότητές της. Για να γίνει αυτό πρέπει να υλοποιηθεί σε ένα μοναδικό παγκόσμιο σύστημα .[25]

### **1.3. Η Σχέση του Σημασιολογικού Ιστού με τον Παγκόσμιο Ιστό**

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.

Για παράδειγμα με τη χρήση της HTML και ενός προγράμματος πλοήγησης μπορούμε να δημιουργήσουμε και να παρουσιάσουμε μια ιστοσελίδα που απαριθμεί στοιχεία κάποιων προς πώληση βιβλίων. Όμως με την HTML δεν μπορεί να γίνει αντιληπτό ότι το στοιχείο “The Da Vinci Code” χαρακτηρίζει ένα βιβλίο αφού αναφέρεται στον τίτλο του ή ότι το στοιχείο “€20” αναφέρεται στην τιμή του. Δεν υπάρχει επίσης κανένας τρόπος να εκφραστεί το γεγονός ότι αυτά τα κομμάτια πληροφορίας είναι αλληλένδετα στην περιγραφή ενός συγκεκριμένου στοιχείου (δηλαδή ενός βιβλίου και μόνο), ευδιάκριτου από άλλα που ίσως απαριθμούνται στη σελίδα.

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 OWL (Web Ontology Language) οι οποίες είναι κατανοητές από τους υπολογιστές.[28]

## 1.4. Το Όραμα του Σημασιολογικού Διαδικτύου

Το ΣΔ αποτελεί μια επέκταση του υπάρχοντος διαδικτύου, στην οποία η πληροφορία που παρουσιάζεται στο χρήστη ορίζεται με τέτοιο τρόπο ώστε να είναι κατανοητή όχι μόνο από τους ανθρώπους αλλά και από τα προγράμματα-πράκτορες, ενισχύοντας έτσι τη διαλειτουργικότητα της επεξεργασίας των πληροφοριών μεταξύ των πρακτόρων αλλά και διευκολύνοντας τη λειτουργικότητα της χρήσης του διαδικτύου από τους ανθρώπους με τη βοήθεια των πρακτόρων. Το ΣΔ βασίζεται στην ιδέα της οργάνωσης και διασύνδεσης της πληροφορίας που υπάρχει στο διαδίκτυο, ώστε να μπορεί να χρησιμοποιηθεί πιο αποτελεσματικά για την ανακάλυψη, αυτοματοποίηση, ομαδοποίηση και επαναχρησιμοποίησή της από διαφορετικές μεταξύ τους διαδικτυακές εφαρμογές.

Το ΣΔ φιλοδοξεί να παρέχει στο μέλλον την υποδομή που απαιτείται για τη δημιουργία και την αξιοποίηση του συνόλου των πληροφοριών από ένα πλήθος διαδικτυακών εφαρμογών που θα περιλαμβάνουν, εκτός από τις απλές ιστοσελίδες του διαδικτύου, τις εταιρικές βάσεις δεδομένων, τις διαδικτυακές ηλεκτρονικές υπηρεσίες, τους πράκτορες, ακόμη και τις οικιακές συσκευές. Με το νέο αυτό τρόπο οργάνωσης των δεδομένων, οι ευφυείς λογισμικοί πράκτορες που μετά από αίτηση κάποιου χρήστη αναζητούν πληροφορίες ή παρεχόμενες υπηρεσίες στο διαδίκτυο, θα έχουν τη δυνατότητα να φιλτράρουν καλύτερα τα δεδομένα που συλλέγουν. Έτσι αυτά θα ανταποκρίνονται πραγματικά στις ανάγκες του χρήστη και θα παρουσιάζονται με κατανοητή μορφή. Το όραμα του ΣΔ συνεχώς ενδυναμώνεται με τη γέννηση μιας νέας γενιάς τεχνολογιών και εργαλείων. Συγκεκριμένα, δημιουργούνται νέες γλώσσες, όπως οι XML, RDF, OWL, που αναπαριστούν την πληροφορία σε μορφή εύκολα κατανοητή και επεξεργάσιμη από τους πράκτορες.

Στο σημείο αυτό θα πρέπει να διευκρινιστεί ότι το ΣΔ δεν είναι ακόμα υπαρκτό, αλλά ότι υπάρχει μια εξαιρετικά σημαντική κινητικότητα ερευνητών και εταιριών οι οποίες αναπτύσσουν τις συγκεκριμένες τεχνολογίες καθώς και κάποιες (όχι ακόμα αρκετές) εφαρμογές οι οποίες βασίζονται πάνω στις τεχνολογίες αυτές. Παρόλα αυτά το όραμα του ΣΔ ίσως να μην είναι και τόσο μακρινό αν γίνει ένας παραλληλισμός με τη διάδοση του ίδιου του διαδικτύου. Το 1989, όταν δηλαδή το διαδίκτυο έκανε τα πρώτα βήματα για

την παγκόσμια καθιέρωσή του, η ανακάλυψη και ανάκτηση πληροφοριών από απομακρυσμένα συστήματα ήταν κάτι που μόνο ένας ειδικός μπορούσε να κάνει. Το διαδίκτυο ως τεχνολογικό υπόβαθρο υπήρχε και θεωρητικά έδινε δυνατότητα πρόσβασης σε αυτά τα συστήματα. Υπήρχαν όμως σημαντικά προβλήματα ευχρηστίας εξαιτίας των διαφορετικών πρωτοκόλλων πρόσβασης, ακόμη και μέσα στο ίδιο σύστημα.

Με την ανάπτυξη του *Παγκόσμιου Ιστού* (*World Wide Web*) η διασύνδεση των διαφόρων πηγών πληροφοριών έγινε πολύ εύκολη υπόθεση, χωρίς να λαμβάνονται υπόψη ποια πρωτόκολλα χρησιμοποιούνται. Ακόμη και αν οι πληροφορίες αυτές αντλούνται από μια καλά οργανωμένη βάση δεδομένων, η δομή αυτή δεν είναι ευδιάκριτη για ένα λογισμικό πράκτορα που αναζητά στοιχεία, με αποτέλεσμα να μην είναι εύκολη η αυτοματοποίηση κάποιων χρονοβόρων και επίπονων εργασιών στο διαδίκτυο από προγράμματα-πράκτορες.[3]

## 1.5. Δικτυακές υπηρεσίες του σημασιολογικού ιστού

Εδώ θα αναφέρουμε κάποιες νέες δικτυακές υπηρεσίες που θα δημιουργηθούν στα πλαίσια του σημασιολογικού ιστού. Έτσι έχουμε:

- **Αυτόματη ανακάλυψη δικτυακής υπηρεσίας.** Αφού οι περιγραφές των δικτυακών υπηρεσιών είναι καταχωρημένες σε δημόσια μητρώα, έξυπνοι πράκτορες θα μπορούσαν να ελέγχουν αυτά τα μητρώα ψάχνοντας για την επιθυμητή από τον χρήστη υπηρεσία. Ήδη έχει γίνει έρευνα στον τομέα της υποβοηθούμενης από κάποιον πράκτορα πλοήγησης στο web.
- **Αυτόματη επίκληση διαδικτυακής υπηρεσίας.** Σήμερα οι περισσότερες υπηρεσίες, κατά την διάρκεια της εκτέλεσής τους απαιτούν ανθρώπινη παρέμβαση, όπου ο χρήστης κάνει διάφορες επιλογές για να έχει το επιθυμητό αποτέλεσμα. Στην περίπτωση της αυτόματης εκτέλεσης της δικτυακής υπηρεσίας ο χρήστης απλά θα διατυπώνει ένα ερώτημα και ένας έξυπνος πράκτορας θα ενεργοποιεί όλες τις απαραίτητες διαδικασίες αυτόματα, στη θέση του χρήστη.

- **Αυτόματη σύνθεση και λειτουργία δικτυακής υπηρεσίας.** Αν δοθεί ένα σύνολο από υπηρεσίες και ένας στόχος, είναι δυνατόν να δημιουργηθεί μια νέα υπηρεσία για να τον επιτύχει. Για να γίνει κάτι τέτοιο πρέπει να περιγραφούν επίσημα οι απαιτήσεις και τα αποτελέσματα της κάθε υπηρεσίας. Τότε η τεχνολογία για την αυτόματη δημιουργία περιβαλλόντων εργασιών ίσως να μπορεί να χρησιμοποιηθεί για να παράγει υπηρεσίες αυτόματα.
- **Αυτόματος έλεγχος εκτέλεσης δικτυακών υπηρεσιών.** Είναι χρήσιμο για υπηρεσίες που τρέχουν για μεγάλα χρονικά διαστήματα να μπορούμε να τις παρακολουθούμε και να ρωτάμε για την κατάσταση των διάφορων εργασιών τους ώστε να ξέρουμε ανά πάσα στιγμή σε πιο στάδιο βρίσκεται το έργο που επιτελούν.

Όλα αυτά έχουν ένα ευρύ πεδίο εφαρμογών στον επιχειρηματικό κόσμο. Με διάφορες σημασιολογικές τεχνικές μπορεί να αναπτυχθεί σημαντικά το ηλεκτρονικό εμπόριο, με την εμφάνιση νέων προτύπων καθώς όλες οι διεργασίες που έπρεπε να κάνει ένας χρήστης για να αγοράσει κάτι απλοποιούνται και αναβαθμίζονται ποιοτικά.[38]

## 1.6 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: Σχισιακά 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 και οι αντικειμενοστρεφείς γλώσσες διαμόρφωσης.

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 διαφόρων σελίδων 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.[39]

## **1.7 Οι έξυπνοι πράκτορες στο σημασιολογικό ιστό**

Στον σημασιολογικό ιστό οι έξυπνοι πράκτορες θα αλλάξουν τον τρόπο με τον οποίο 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 οι έξυπνοι πράκτορες 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.[36]

## 1.8 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 Boolean 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. Ακόμη θα είναι σε θέση να κατανοεί τα ερωτήματα σημασιολογικά και όχι απλά σαν λέξεις-κλειδιά. Αυτό σημαίνει ότι αν για παράδειγμα έχουμε ένα ερώτημα “*return all the reviewers for book 'The Semantic Web: an Introduction'*” η σημασιολογική μηχανή θα μας επιστρέψει ένα σύνολο από reviewers για το συγκεκριμένο βιβλίο αντί να θεωρήσει κάθε όρο του ερωτήματος ως λέξη-κλειδί και να μας φέρει ως αποτέλεσμα κάθε σελίδα που περιέχει τους όρους αυτούς και τις ρίζες τους. Επομένως μιλάμε για μεγαλύτερη ακρίβεια στην ανάκτηση, δηλαδή για πιο ποιοτικά αποτελέσματα.

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.

Στην πρώτη 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.[36]



Στην δεύτερη περίπτωση, ακριβώς επειδή έχει γίνει ανασκόπηση της κάθε σελίδας , τα αποτελέσματα φανερώνουν μεγάλη σχετικότητα, άρα υψηλή ακρίβεια. Όμως, όπως είναι λογικό, αυτή η δουλειά 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 υπάρχουν semantics 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.[11]

## **1.9 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.[42]

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.

Αυτή η τεχνική κατ'αρχάς καταπολεμά αποτελεσματικά το φαινόμενο των broken links που συχνά απαντάται στο δίκτυο. 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.[14]

## **1.10 Στόχοι του Σημασιολογικού Ιστού**

Το σημασιολογικό διαδίκτυο φιλοδοξεί να επιτρέψει τόσο στους χρήστες όσο και στα πληροφοριακά συστήματα να μπορούν να κατανοούν και να επεξεργάζονται λογικά τις ίδιες πληροφορίες και δεδομένα. Προφανώς αυτή η προοπτική θα ανοίξει νέους συναρπαστικούς ορίζοντες, αφού τα πληροφοριακά συστήματα θα μπορούν πέρα από το να αντλούν μεγάλες ποσότητες πληροφοριών όπως κάνουν σήμερα, επιπλέον να τις επεξεργάζονται λογικά και εξάγουν λογικά συμπεράσματα. Τα αποτελέσματα μιας τέτοιας διεργασίας θα μπορούν να δημοσιεύονται και αυτά στο διαδίκτυο έτσι ώστε να μπορούν να χρησιμοποιηθούν από τρίτους, αυξάνοντας έτσι τη συνολική γνώση και αξία του σημασιολογικού ιστού.

Προχωρώντας ένα βήμα πιο πέρα, θα δημιουργηθούν και θα λειτουργούν έξυπνες υπηρεσίες οι οποίες θα μπορούν να εντοπίζουν και να χρησιμοποιούν η μία την άλλη, να συνεργάζονται μεταξύ τους χωρίς να έχουν εξαρχής προετοιμαστεί για αυτό, και όλα αυτά αυτόματα χωρίς την ανθρώπινη παρέμβαση.[3]

## 1.11 Τεχνολογίες που χρησιμοποιεί το Semantic Web

Οι τεχνολογίες που χρησιμοποιεί το semantic web είναι οι εξής:

- τα **URIs** (Universal Resource Identifier): συμβολοσειρές που ταυτοποιούν μοναδικά μία οντότητα (ένα Web site, μία ιδιότητα, έναν άνθρωπο, ένα πράγμα κλπ)
- γλώσσα επισημάνσεων **XML**: επιτρέπει στους χρήστες να προσθέτουν αυθαίρετη δομή στα έγγραφά τους, χωρίς να καθορίζει την σημασιολογία αυτής της δομής
- τεχνολογία **RDF**: χρησιμοποιείται για την αναπαράσταση δεδομένων και ανταλλαγή γνώσης στο διαδίκτυο
- τεχνολογία **OWL**: χρησιμοποιείται για τη δημιουργία και διανομή οντολογιών, υποστηρίζοντας προχωρημένη αναζήτηση στο διαδίκτυο, πράκτορες λογισμικού και διαχείριση γνώσης.

Κάθε μία από τις παραπάνω τεχνολογίες βασίζεται σε εκείνες που αναφέρονται πριν από αυτή. Έτσι π.χ. η RDF βασίζεται στην XML και χρησιμοποιεί URIs. Οι τεχνολογίες αυτές συνδυάζονται ώστε να παρέχουν περιγραφές που συμπληρώνουν ή υποκαθιστούν το περιεχόμενο εγγράφων στο Web, όπως περιγράφεται παρακάτω. Αυτές οι machine-readable περιγραφές επιτρέπουν την προσθήκη νοήματος στο περιεχόμενο, διευκολύνοντας έτσι την αυτόματη αναζήτηση πληροφοριών από υπολογιστές.

Το νόημα του περιεχομένου εκφράζεται μέσω της RDF, η οποία κωδικοποιεί σύνολα από triples (τριάδες) που αντιπροσωπεύουν περίπου το υποκείμενο, το ρήμα και το αντικείμενο μιας πρότασης. Αυτές οι τριάδες μπορούν να γραφούν με XML. Ένα έγγραφο της RDF δηλώνει ότι κάποιες οντότητες (ιστοσελίδες, άνθρωποι ή πράγματα) έχουν κάποιες ιδιότητες με συγκεκριμένες τιμές (π.χ. ο X είναι συγγραφέας του βιβλίου Y). Αυτός ο τρόπος αναπαράστασης μπορεί να εκφράσει τα περισσότερα δεδομένα που κατανοούν οι υπολογιστές. Το υποκείμενο, το ρήμα και το αντικείμενο προσδιορίζονται

από ένα URI (Universal Resource Identifier), όπως ακριβώς γίνεται με τις ιστοσελίδες. Έτσι μπορεί ο καθένας να προσθέσει μία νέα οντότητα ή μία νέα ιδιότητα.

Ένα βασικό συστατικό του Semantic Web είναι οι οντολογίες. Οι οντολογίες ορίζουν με τυπικό τρόπο τις έννοιες και τις συσχετίσεις των εννοιών για κάποιο πεδίο. Περιέχουν ορισμούς κλάσεων αντικειμένων και σχέσεις ανάμεσα στις κλάσεις, καθώς και κανόνες εξαγωγής λογικών συμπερασμάτων. Μέσω των οντολογιών γίνεται δυνατόν να συνδυαστούν δεδομένα από διαφορετικές πηγές, οι οποίες μοιράζονται την ίδια οντολογία. Ακόμη, λύνονται προβλήματα ορολογίας καθώς η σημασία των όρων που εμφανίζονται σε μία σελίδα μπορεί να οριστεί με δείκτες προς την οντολογία. Οι οντολογίες μπορούν να βελτιώσουν την λειτουργία του Web αυξάνοντας την ακρίβεια των αναζητήσεων πληροφοριών, αφού αναζητούνται οι πληροφορίες που αναφέρονται μόνο σε μία συγκεκριμένη έννοια και όχι σε κάποια λέξη κλειδί. Μπορούν επίσης να χρησιμοποιηθούν για την συσχέτιση των πληροφοριών ενός site με αντίστοιχες δομές γνώσης και λογικούς κανόνες. Οι κοινές οντολογίες βοηθούν στην ανταλλαγή δεδομένων και νοημάτων μεταξύ διαφορετικών web-based υπηρεσιών.

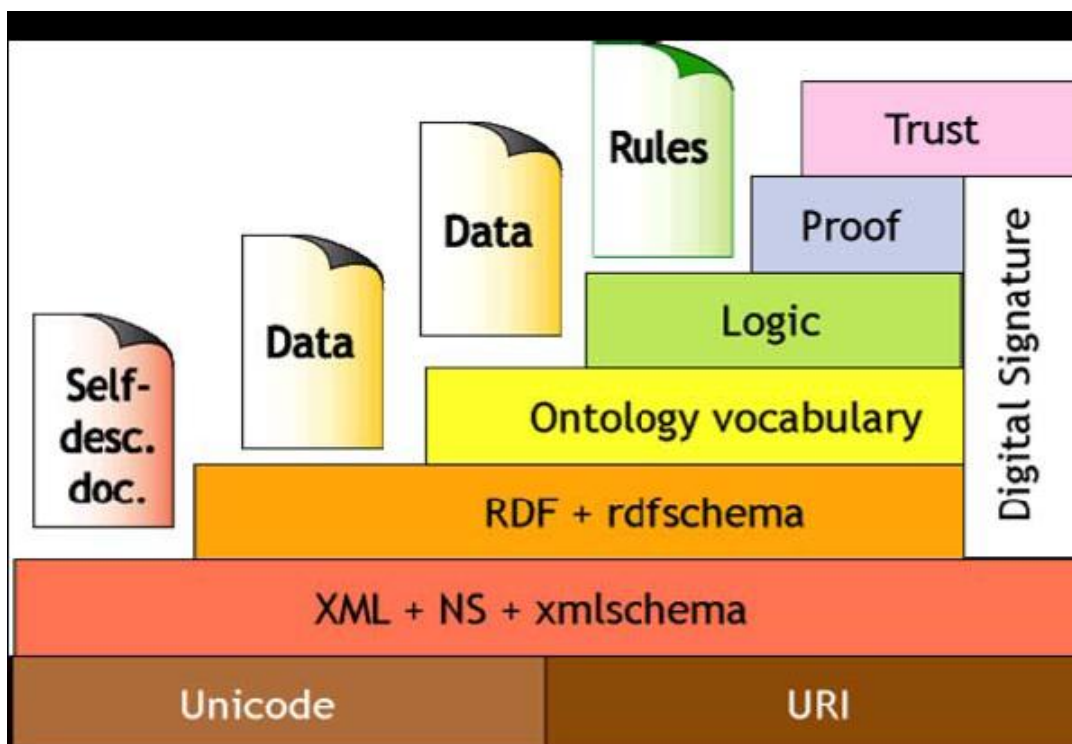
Υπάρχουν πολλές αυτοματοποιημένες υπηρεσίες βασισμένες στο διαδίκτυο (web services) που δεν χρησιμοποιούν τη σημασιολογία, όμως άλλα προγράμματα (όπως είναι οι πράκτορες λογισμικού) δεν μπορούν να εντοπίσουν μόνα τους μία τέτοια υπηρεσία για κάποια συγκεκριμένη λειτουργία. Αυτό μπορεί να συμβεί μόνο όταν υπάρχει μία κοινή γλώσσα περιγραφής παρεχόμενων υπηρεσιών σε ένα δίκτυο ούτως ώστε οι διάφοροι πράκτορες να διαφημίζουν τις υπηρεσίες τους σε μία υπηρεσία κίτρινων σελίδων.

Αφού εντοπιστεί ο επιθυμητός πράκτορας, οι δύο επικοινωνούντες πράκτορες λογισμικού μπορούν να κατανοήσουν ο ένας τον άλλο ανταλλάσσοντας οντολογίες. Το Semantic Web παρέχει αυτή την ευελιξία. Ανακαλύπτοντας νέες οντολογίες οι πράκτορες λογισμικού αποκτούν νέες δυνατότητες εξαγωγής λογικών συμπερασμάτων.

Η ενοποιητική λογική γλώσσα του Semantic Web δίνει τη δυνατότητα να ενωθούν οι έννοιες που μπορεί να ορίσει ο καθένας (μέσω ενός URI) σε ένα καθολικό Web. Έτσι οι πράκτορες λογισμικού θα μπορούν να αναλύουν σημασιολογικά τη γνώση των ανθρώπων παρέχοντας μία νέα μορφή χρήσιμων εργαλείων.[7]

## 1.12 Η αρχιτεκτονική του σημασιολογικού ιστού σε στρώματα (Ontologies in the Layered Architecture)

Ο σημαντικός ρόλος των οντολογιών για την επιτυχία του Σημασιολογικού Ιστού είναι επίσης ορατός στο διάσημο διάγραμμα της "αρχιτεκτονικής σε στρώματα" (layered architecture), όπως φαίνεται στο Σχήμα 1.12. το οποίο παρουσιάστηκε από τον Berners-Lee στο XML 2000. Αν και μόνο ένα στρώμα έχει τον τίτλο "Λεξιλόγιο οντολογίας", ("Ontology vocabulary") οι οντολογίες και τα σχήματα αντιπροσωπευσής τους είναι στην πραγματικότητα παρόντα και σε τα τρία στρώματα "RDF + rdfschema", "Λεξιλόγιο οντολογίας" ("Ontology vocabulary") και "Λογική" ("Logic"). Μέχρι σήμερα δεν υπάρχει σχεδόν καμία συγκεκριμένη λύση ή ακόμα και προτάσεις διαθέσιμες για τα δύο ανώτερα στρώματα, την "Απόδειξη" ("Proof") και την "Εμπιστοσύνη" ("Trust"), οι σχετικές με τις οντολογίες τεχνολογίες διαμορφώνουν το μεγαλύτερο μέρος της αρχιτεκτονικής σε στρώματα.[29]



Σχήμα 1.12.: Η αρχιτεκτονική σε στρώματα του Σημασιολογικού Ιστού (The Semantic Web Layered Architecture)

Προκειμένου να κατασταθεί ο Σημασιολογικός Ιστός δυνατός, πρέπει να υπάρξουν εφικτές λύσεις για όλα τα στρώματα στο διάγραμμα:

- **Unicode, URI:** Τα Unicode και URIs παρέχουν ένα τρόπο ανεξάρτητο από το σύστημα για την κωδικοποίηση και την κατεύθυνση των πληροφοριών (encoding and addressing information). Χρησιμοποιούνται ήδη από τον Παγκόσμιο Ιστό (World Wide Web) και αποτελούν επίσης τη βάση για το Σημασιολογικό Ιστό, μαζί με το δίκτυο και πρωτόκολλα μετάδοσης όπως το TCP/IP και το HTTP.[25]
- **XML:** Το XML και τα σχετικά πρότυπά του, παρέχουν έναν τρόπο να αντιπροσωπευθούν οι αυθαίρετες δομές πληροφοριών και είναι το de facto πρότυπο για την ανταλλαγή δεδομένων μεταξύ εφαρμογών. Λαμβάνοντας υπόψη την ευρεία υποστήριξή τους τόσο από εργαλεία όσο και από χρήστες, είναι μια λογική επιλογή για να χρησιμεύσει ως ένα σχήμα σειριοποίησης (serialization format) για τις οντολογίες.
- **RDF:** Τα RDF και RDFS παρέχουν έναν απλό τρόπο να αντιπροσωπευθούν ισχυρισμοί (assertions) ως τριπλέτες πληροφορίας (information triples) που μαζί δημιουργούν ένα γράφο εννοιών και των σχέσεων μεταξύ τους. Κατά συνέπεια, τα RDF και RDFS μπορούν να χρησιμοποιηθούν ως σχήμα για την αναπαράσταση πληροφοριών που περιλαμβάνονται σε μια οντολογία.
- **Οντολογία:** Τα RDF(S) δεν παρέχουν έναν τρόπο να εκφραστούν συγκεκριμένες προηγμένες ιδιότητες των εννοιών ή των σχέσεων (όπως περιορισμοί πληθικού αριθμού -cardinality-, η μεταβατικότητα -transitivity- κ.λπ....) που απαιτούνται για σημασιολογικά κατάλληλο περιεχόμενο οντολογίας και που παρέχουν τη βάση για το συμπέρασμα και τους λογικούς κανόνες. Πολλές από αυτές προηγμένες ιδιότητες παρέχονται αυτήν την περίοδο από τη (Web Ontology Language (OWL) που δομείται πάνω στο RDFS.
- **Λογική (Logic):** Πάνω από τις οντολογίες, οι μηχανές συμπεράσματος χρησιμοποιούνται για να ανακτήσουν τους ισχυρισμούς (assertions) που δεν εκφράζονται άμεσα μέσα στην οντολογία, αλλά μπορούν να εξαχθούν από τα

υπάρχοντα γεγονότα. Ενώ συνήθως δεν είναι μέρος μιας οντολογίας, οι μηχανές συμπεράσματος χρειάζονται δεδομένα από μια οντολογία για να ανακτήσουν τα εξαγόμενα γεγονότα που απαιτούνται για να μπορέσουν οι πράκτορες υπολογιστών να συμπεριφέρονται "με ευφυΐα".

Δύο ακόμα παράγοντες κλειδιά δεν εξετάζονται από το διάγραμμα της αρχιτεκτονικής σε στρώματα: πρέπει να υπάρχει μια συναίνεση στον τρόπο *καθιέρωσης ταυτότητας* για τις αυθαίρετες έννοιες (φυσικές και αφαιρέτες καθώς επίσης και για τους ηλεκτρονικούς πόρους), και πρέπει να υπάρξει κάποιο άμεσο όφελος για τους πιθανούς συντάκτες οντολογίας για να δημιουργήσουν οντολογίες και να τις μοιραστούν με άλλους μέσω του Σημασιολογικού Ιστού.[29]

### **1.13 Η σημασιολογία στην επιστήμη των υπολογιστών**

Σε ένα πληροφοριακό σύστημα τα δεδομένα είναι αποθηκευμένα στη μορφή δυαδικών ψηφίων. Ένα σύνολο ψηφίων (8, 16, 32) αποτελούν ένα χαρακτήρα. Στην περίπτωση αποθηκευμένων δεδομένων φυσικής γλώσσας ο χαρακτήρας αντιστοιχεί σε ένα αλφαριθμητικό σύμβολο. Τα αλφαριθμητικά σύμβολα με τη σειρά τους δομούν λέξεις οι οποίες με τη σειρά τους εντάσσονται σε μεγαλύτερες μορφές δεδομένων όπως κείμενο, λίστες, βάσεις δεδομένων κλπ. Όλες αυτές οι δομές συνοδεύονται από ένα σύνολο αφανών στο χρήστη δεδομένων τα οποία τους προσδίδουν κάποια λειτουργικότητα. Έτσι στην περίπτωση για παράδειγμα του κειμένου, συνοδεύονται από δεδομένα μορφοποίησης που καθορίζουν το πως θα εμφανιστεί το κείμενο σε μια οθόνη υπολογιστή ενώ σε μια βάση δεδομένων εντάσσονται σε στήλες και προστίθενται δείκτες που συνδέουν τα δεδομένα μεταξύ τους.

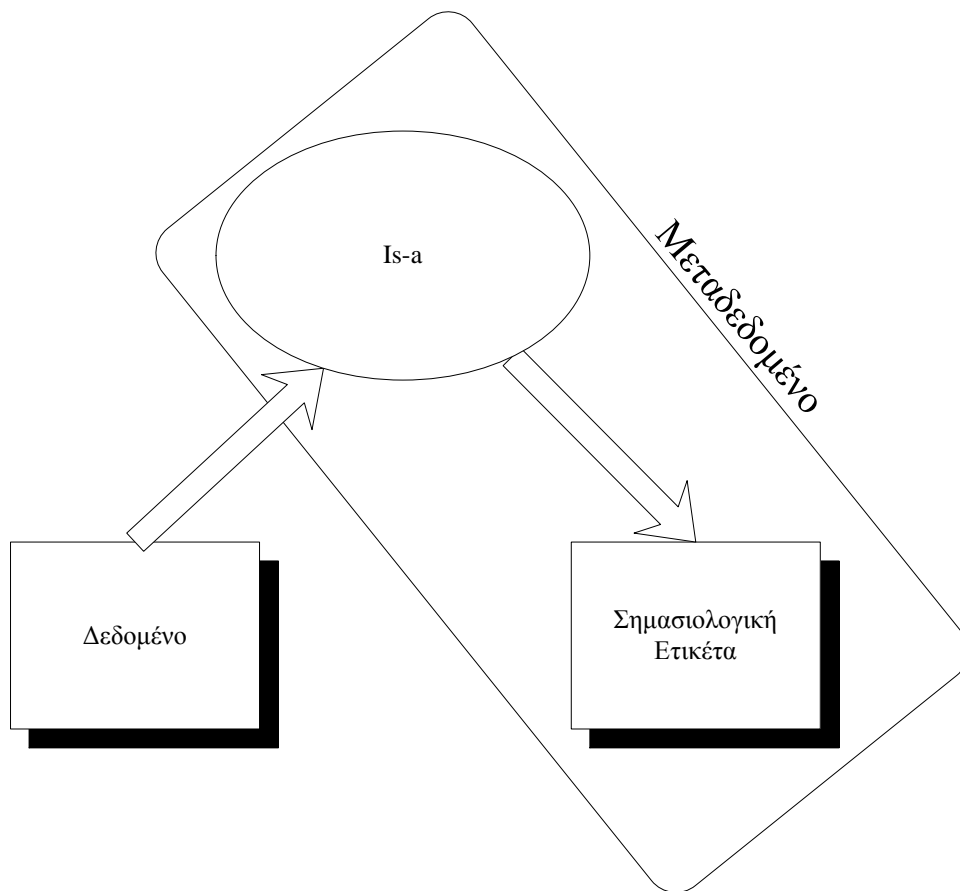
Το πρόβλημα έγκειται στο γεγονός ότι ένας υπολογιστής δεν έχει αίσθηση των δεδομένων που διαχειρίζεται. Για να είναι σε θέση να επεξεργαστεί δεδομένα θα πρέπει να υπάρχουν σαφείς οδηγίες ποια δεδομένα είναι αυτά και πώς ακριβώς

πρέπει να τα επεξεργαστεί. Με άλλα λόγια, ένας υπολογιστής δεν είναι στην παρούσα κατάσταση (και από ότι διαφαίνεται ούτε στο άμεσο μέλλον) σε θέση να εξάγει νόημα από τα συμφραζόμενα (in-context) αφού όλα τα δεδομένα για αυτόν είναι ασυνάρτητα (out-of-context). Κάτι τέτοιο είναι αποτέλεσμα του γεγονότος ότι ο υπολογιστής δεν μπορεί να αναπτύξει νοητική λειτουργία αλλά μόνο επεξεργαστική και επίσης δεν μπορεί να έχει την έμφυτη στον άνθρωπο γλωσσική διαίσθηση.

Στόχος όλων των σχετικών με τη σημασιολογία πρωτοβουλιών στο χώρο των πληροφοριακών συστημάτων είναι η κάμψη αυτού του εμποδίου. Η προσέγγιση είναι να προσδοθεί στα δεδομένα σημασιολογικό περιεχόμενο με άμεση παρέμβαση αφού ο άνθρωπος είναι ο μόνος που μπορεί να αντιληφθεί τη πληροφοριακή τους υπόσταση. Αυτό επιτυγχάνεται με την επισύναψη μεταδεδομένων στα δεδομένα.

Τα μεταδεδομένα είναι δεδομένα που περιγράφουν άλλα δεδομένα ή σύμφωνα με τον ορισμό που δίνει ο William R. Durrell «τα μεταδεδομένα είναι δομημένα, κωδικοποιημένα δεδομένα που περιγράφουν τα χαρακτηριστικά οντοτήτων που φέρουν πληροφορία ώστε να βοηθήσουν στην ταυτοποίηση, ανακάλυψη, εκτίμηση και διαχείριση των περιγραφόμενων οντοτήτων». Όπως προκύπτει από τον παραπάνω ορισμό, τα μεταδεδομένα είναι επίσης δεδομένα. Αυτό που τα διαφοροποιεί είναι η σύνδεσή τους με αυτά. Η σύνδεση αυτή είναι μια σχέση «κληρονομικότητας» γνωστή με τον όρο «Is-a relationship».





**Σχήμα 1.13.: Is-a Relationship και μεταδεδομένα**

Όπως φαίνεται και στο παραπάνω σχήμα, ένα μεταδεδομένο αποτελείται από τη σχέση κληρονομικότητας με το αρχικό δεδομένο και ένα επιπλέον κομμάτι πληροφορίας, τη σημασιολογική ετικέτα (semantic tag).

Για να γίνει μία αποτύπωση της χρήσης των μεταδεδομένων θα χρησιμοποιηθεί το παρακάτω παράδειγμα:

**Παράδειγμα:**

«Αλβέρτος Αϊνστάιν (14 Μαρτίου 1879, Ουλμ, Γερμανία – 18 Απριλίου 1955, Πρίνστον, Νιου Τζέρσεϊ). Γερμανός φυσικός και θεμελιωτής της θεωρίας της σχετικότητας».

Πρώτα πρέπει να γίνει μια αναγνώριση των τύπων των δεδομένων.

- Αλβέρτος: Όνομα
- Αϊνστάιν: Επώνυμο
- 14: Ημερομηνία
- Μαρτίου: Μήνας
- 1879: Χρονιά
- Ουλμ: Πόλη
- Γερμανία: Χώρα
- 18: Ημερομηνία
- Απριλίου: Μήνας
- 1955: Χρονιά
- Πρίνστον: Πόλη
- Νιου Τζέρσεϊ: Πολιτεία
- Γερμανός: Εθνικότητα
- Φυσικός: Ιδιότητα (ή επάγγελμα)
- Θεωρία της σχετικότητας: Επιστημονική θεωρία

Στη λέξη θεμελιωτής δεν προσδόθηκε κάποιος τύπος γιατί αυτή φέρει ρηματικό περιεχόμενο (θα μπορούσε να αντικατασταθεί με το ρήμα «θεμελίωσε»). Τέτοιες λέξεις (κατά κύριο λόγο τα ρήματα αλλά συχνά επίθετα, μετοχές ή και ολόκληρες προτάσεις) δεν φέρουν αυτόνομο περιεχόμενο αφού ένα ρήμα εκφράζει ενέργεια και κατά συνέπεια τη σχέση ανάμεσα σε ουσιαστικά. Το νοηματικό περιεχόμενο των ρηματικών λέξεων είναι ιδιαίτερα σημαντικό κατά τη δόμηση οντολογιών και θα μελετηθεί αργότερα. Σε μια πρώτη απόπειρα συνοδείας του άνωθι κειμένου με μεταδεδομένα το αποτέλεσμα θα ήταν ως εξής:

[Αλβέρτος]:όνομα [Αϊνστάιν]:επώνυμο ([14]:ημερομηνία [Μαρτίου]:μήνας [1879]:χρονιά, [Ουλμ]:πόλη, [Γερμανία]:χώρα – [18]:ημερομηνία [Απριλίου]:μήνας [1955]:χρονιά, [Πρίνστον]:πόλη, [Νιού Τζέρσεϊ]:πολιτεία). [Γερμανός]:εθνικότητα [φυσικός]:επάγγελμα και θεμελιωτής της [θεωρίας της σχετικότητας]:επιστημονική θεωρία.

Όλα τα μεταδεδομένα παρουσιάζονται με έντονους χαρακτήρες και φυσικά δεν θα απεικονίζονται στο χρήστη της υποθετικής ηλεκτρονικής εγκυκλοπαίδειας. Το σύμβολο «[:» εκφράζει την «Is-a relationship» ενώ η λέξη (λέξεις) που ακολουθεί τη σημασιολογική ετικέτα.

Σε ένα επόμενο στάδιο μπορούν να δημιουργηθούν μεγαλύτερες δομές με τα μεταδεδομένα μας. Για παράδειγμα, θα ήταν χρήσιμο να επισημανθεί ότι το «14 Μαρτίου 1879» είναι μια ημερομηνία γέννησης ενώ το «Ουλμ, Γερμανία» είναι ο τόπος γέννησης. Σε αυτή την περίπτωση το μοντέλο επεκτείνεται ως εξής:

[[Αλβέρτος]:όνομα [Αϊνστάιν]:επώνυμο]:πρόσωπο ([[14]:ημερομηνία [Μαρτίου]:μήνας [1879]:χρονιά]:ημερομηνία γέννησης, [[Ουλμ]:πόλη, [Γερμανία]:χώρα]τόπος γέννησης – [[18]:ημερομηνία [Απριλίου]:μήνας [1955]:χρονιά]:ημερομηνία θανάτου, [[Πρίνστον]:πόλη, [Νιού Τζέρσεϊ]:πολιτεία]:τόπος θανάτου). [Γερμανός]:εθνικότητα [φυσικός]:επάγγελμα και θεμελιωτής της [θεωρίας της σχετικότητας]:επιστημονική θεωρία.

Σε αυτή του τη μορφή το κείμενο είναι σαφώς πιο επεξεργάσιμο. Αντί να γίνονται εικασίες για τη σημασία των διαφόρων δεδομένων, αυτά συνοδεύονται από ένα σαφή ορισμό. Έτσι ένα πρόγραμμα θα ήταν σε θέση επεξεργαζόμενο τα άρθρα της εγκυκλοπαίδειας να παράγει για παράδειγμα μια λίστα όλων των ανθρώπων που γεννήθηκαν στο Ουλμ την ίδια χρονιά με τον Αϊνστάιν ή μια λίστα όλων των γερμανών φυσικών που πέθαναν τη δεκαετία του '50 και σε ηλικία μεγαλύτερη των 70 χρόνων.

Σε ένα επόμενο στάδιο, μια μηχανή αναζήτησης θα είναι σε θέση να αναζητήσει τέτοιου είδους πληροφορίες από πολλαπλές πηγές και να τις συνδυάσει κατάλληλα. Σε συνάρτηση με τις τεχνικές ανάλυσης κειμένου τελικά ο χρήστης θα είναι σε θέση να υποβάλει ερώτημα σε φυσική γλώσσα και να πάρει άμεσα την απάντηση στο ερώτημά του.

Αυτό που πρακτικά έγινε στο παραπάνω παράδειγμα είναι σε πρώτο στάδιο να ανιχνευθούν τα είδη των δεδομένων. Αφού αυτά αναγνωρίστηκαν, στο επόμενο στάδιο ορίστηκαν κλάσεις γι' αυτά. Ο όρος κλάση είναι ο ίδιος με αυτόν που χρησιμοποιείται και στις αντικειμενοστραφείς γλώσσες προγραμματισμού και πράγματι έχει και εδώ την ίδια έννοια δηλαδή της κατηγορίας ομοειδών αντικειμένων. Αντικείμενα του ίδιου τύπου (π.χ. ημερομηνίες) απαρτίζουν μια κλάση. Επίσης οι κλάσεις έχουν constructors της μορφής:

**[data]:semantic\_tag\_of\_class**

Επιπλέον, μπορούν να οριστούν υποκλάσεις των κλάσεων. Έτσι υποκλάση της κλάσης «ημερομηνία» είναι η «ημερομηνία γέννησης». Επίσης, οι κλάσεις μπορούν να έχουν ιδιότητες με τις οποίες δηλώνεται η σχέση τους με άλλες κλάσεις ή με πεδία δεδομένων. Κατά την ίδια λογική με τις αντικειμενοστραφείς γλώσσες, από τις κλάσεις προκύπτουν στιγμιότυπα (instances) με τη χρήση του constructor. Έτσι το «**[Αλβέρτος]:όνομα**» είναι στιγμιότυπο της κλάσης «όνομα». Όλη αυτή η λογική είναι σχετικά απλή και εύκολα κατανοητή.

Επέκταση αυτού του μοντέλου μπορεί να καλύψει οποιουσδήποτε τύπους δεδομένων και διάφορες ανάγκες. Σύμφωνα με τον Dave McComb ένα τέτοιο σημασιολογικό μοντέλο εξελίσσεται ως εξής:

- Γλωσσάρι (Glossary): Πρόκειται για μια απλή λίστα όρων με τους ορισμούς τους. Ένα γλωσσάρι επιδιώκει τη συγκέντρωση όρων που

χρησιμοποιούνται από έναν τομέα (domain) και χρησιμεύει στο να δημιουργηθούν απλοί και ξεκάθαροι ορισμοί. Η απλουστευμένη μορφή του δεν επιτρέπει πράξεις επεξεργασίας και συλλογισμούς.

- Ελεγχόμενο λεξιλόγιο: Σε αυτό το στάδιο υπάρχουν πλέον συγκεκριμένες διαδικασίες για την εισαγωγή και εξαγωγή όρων ώστε το τελικό αποτέλεσμα να είναι συνεπές. Οι όροι πλέον αρχίζουν να ορίζονται με βάση άλλους όρους που περιέχονται στο λεξιλόγιο και αρχίζουν να δημιουργούνται σχέσεις μεταξύ των στοιχείων.
- Λεξικό δεδομένων: Αρχίζουν πλέον να ορίζονται τύποι δεδομένων και πλέον γίνεται λόγος για το σύνολο των δεδομένων του συστήματος
- Μοντέλο δεδομένων: Σε αυτό το στάδιο αρχίζει ο ορισμός των σχέσεων μεταξύ των δεδομένων.
- Ταξινόμια: Είναι η επέκταση του μοντέλου δεδομένων σε μια ιεραρχημένη δομή όπου ισχύουν οι αρχές της κληρονομικότητας (αντίστοιχες με αυτές στις αντικειμενοστραφείς γλώσσες προγραμματισμού).
- Οντολογία: Πρόκειται για μια πλήρη ταξινόμια σε αναγνώσιμη από μηχανή μορφή που χρησιμοποιεί URI (Uniform Resource Identifier) για όλες τις μορφές δεδομένων (κλάσεις), τις ιδιότητές τους και τις σχέσεις τους.

Το πως από τον απλό ορισμό κλάσεων προκύπτει μια οντολογία εξετάζεται στη συνέχεια της παρούσας εργασίας.

Σε αυτό το σημείο πρέπει να τονιστεί το εξής: κατά τη διάρκεια ορισμού των κλάσεων και επειδή αυτό είναι το πρώτο βήμα στη δημιουργία ενός σημασιολογικού μοντέλου είναι απαραίτητο οι ορισμοί να γίνουν με τέτοιο τρόπο ώστε τελικά να ικανοποιούνται τα ακόλουθα χαρακτηριστικά:

- Ορθότητα: Κατά πόσο οι όροι και ορισμοί βρίσκονται σε αντιστοιχία μεταξύ τους. Επίσης πόσο ξεκάθαροι είναι οι ορισμοί και οι κανόνες σύνταξης του μοντέλου.
- Επάρκεια Εκφραστικότητα: Κατά πόσο το μοντέλο μπορεί να απεικονίσει όλα τα χρήσιμα στοιχεία.
- Αποτελεσματικότητα: Στο προκύπτον μοντέλο πρέπει να υπάρχει η δυνατότητα αναζήτησης, κατάθεσης ερωτημάτων αλλά και συλλογισμών.
- Πολυπλοκότητα: Το μοντέλο δεν θα πρέπει να είναι αχανές και η πολυπλοκότητά του να περιορίζεται στο βαθμό που οι άλλες αρχές τηρούνται.
- Μεταφρασιμότητα/Μετατρεψιμότητα: Κατά πόσο είναι δυνατό το μοντέλο να μεταφραστεί ώστε να είναι δυνατή η συνεργασία παρόμοιων μοντέλων μεταξύ τους ή η με σαφήνεια ένωσή τους.

Η τήρηση αυτών των αρχών είναι πολύ σημαντική. Ζητούμενο είναι να προκύψει ένα μοντέλο που θα απλοποιεί τη διαχείριση των δεδομένων. Η προφανής απαίτηση για ορθότητα καλύπτει την ανάγκη για ένα συνεκτικό τρόπο έκφρασης των δεδομένων και των σχέσεών τους.

Η επάρκεια είναι απαραίτητη ώστε να μην προκύπτει συνεχώς η ανάγκη για αναθεωρήσεις που κοστίζουν σε χρήμα και χρόνο.

Η αποτελεσματικότητα διασφαλίζει την εύρυθμη διαχείριση του μοντέλου καθώς και τη χρηστικότητά του αφού η διενέργεια συλλογισμών με βάση αυτό είναι από τα βασικότερα πλεονεκτήματα που προσφέρει.

Το ζήτημα της πολυπλοκότητας είναι μάλλον το πιο σύνθετο. Το μοντέλο που θα προκύψει θα πρέπει να είναι όσο το δυνατό πιο απλό γιατί διαφορετικά θα δημιουργήσει τα ίδια προβλήματα με αυτά τα οποία ήρθε να λύσει. Εάν για παράδειγμα ο αριθμός των ορισμένων κλάσεων είναι υπερβολικά μεγάλος τότε

από τη μία πλευρά η ένταξη των δεδομένων σε αυτές γίνεται δύσκολη και η συσχέτισή τους κοπιώδης ενώ από την άλλη η εύρεσή τους θα έγκειται στον προσδιορισμό πολύ ειδικών κριτηρίων.

Τέλος, η μετατρεψιμότητα είναι το στοιχείο που δίνει δυνατότητες για διαλειτουργικότητα. Όπως τονίστηκε και πιο πριν, η εφαρμογή σημασιολογικών προτύπων απαιτεί τη μεγαλύτερη δυνατή συναίνεση των εμπλεκόμενων φορέων. Εάν κάθε οργανισμός ακολουθεί διαφορετικά πρότυπα τότε αίρεται ένα από τα σημαντικότερα πλεονεκτήματα των σημασιολογικών μοντέλων. Σε κάποιες περιπτώσεις όμως δεν είναι δυνατό να υιοθετηθούν κοινά πρότυπα από οργανισμούς που έχουν μεγάλο ενδιαφέρον για συναλλαγές μεταξύ τους. Για παράδειγμα, διαφορετικές χώρες κρατούν διαφορετικά στοιχεία για τους πολίτες τους. Το στοιχείο «Αριθμός Φορολογικού Μητρώου» στην Ελλάδα δεν έχει κάποιο άμεσο αντίστοιχο στις Ηνωμένες Πολιτείες αν και οι αρχές των Η.Π.Α. κρατούν ένα στοιχείο με αντίστοιχα χαρακτηριστικά (μοναδικότητα, ισοβιότητα) το «Social Security Number». Για να είναι σε θέση οι αρχές των δύο χωρών να ανταλλάσσουν στοιχεία των πολιτών τους είτε θα πρέπει να χρησιμοποιήσουν ακριβώς το ίδιο σημασιολογικό μοντέλο, κάτι που θα οδηγήσει στο παράλογο συμπέρασμα ότι Α.Φ.Μ. και S.S.N. είναι το ίδιο πράγμα με μετέπειτα καταστροφικές συνέπειες, είτε να δομήσουν με τέτοιο τρόπο τα συστήματά τους ώστε να είναι σε θέση να αντιλαμβάνονται τουλάχιστον τη δομή άλλων σημασιολογικών μοντέλων με αποτέλεσμα τον υψηλό βαθμό διασυνδεσιμότητάς τους[24]

## 1.14 Το ζήτημα των οντολογιών στο Σημασιολογικό Ιστό

Απαραίτητο μέσο για πλήθος εφαρμογών στον Σημασιολογικό Ιστό, οι οντολογίες προορίζονται να προσφέρουν μία αναπαράσταση γνώσης και ένα λεξιλόγιο από κλάσεις και σχέσεις. Οι οντολογίες καταφέρνουν να συνενώσουν δύο ουσιώδη συστατικά, τα οποία συμβάλλουν στην ανάπτυξη του Παγκόσμιου Ιστού. Από τη μία ορίζουν την τυπική σημασιολογία της πληροφορίας διευκολύνοντας την επεξεργασία της από τον υπολογιστή, ενώ ταυτόχρονα από την άλλη ορίζουν σημασιολογία του πραγματικού κόσμου. Μ' αυτό τον τρόπο επιτυγχάνεται η σύνδεση του περιεχομένου, το οποίο τυχαίνει μηχανικής επεξεργασίας, με τη σημασία που του δίνουν οι άνθρωποι βασιζόμενοι σε κοινά αποδεκτή ορολογία.

Οι οντολογίες μπορούν να διαδραματίσουν έναν πολύ σημαντικό ρόλο στις υπηρεσίες εξατομικευμένης ενημέρωσης, καθώς επιτρέπουν την περιγραφή, επεξεργασία και κατηγοριοποίηση του περιεχομένου τους, μ' έναν τρόπο κοινά κατανοητό τόσο για τους χρήστες όσο και για τα συστήματα. Βασικό ενδιαφέρον των υπηρεσιών αυτών, άλλωστε, είναι η όσο το δυνατόν καλύτερη εξυπηρέτηση των πελατών τους. Η ανάπτυξη, λοιπόν, ενός σημασιολογικά δομημένου συστήματος διευκολύνει την αναζήτηση και τελικά την πρόσβαση που αποτελούν μία πλούσια πηγή διάθεσης πληροφοριών.[1]



## ΚΕΦΑΛΑΙΟ 2

### 2.1 Οντολογίες

#### Εισαγωγή στις οντολογίες

Στις μέρες μας παρατηρείται αυξημένη η χρήση των υπηρεσιών εξατομικευμένης ενημέρωσης από τον Ιστό, οι οποίες προσπαθούν να ικανοποιήσουν τις ολοένα και αυξανόμενες απαιτήσεις των χρηστών. Προκειμένου, λοιπόν, να προσφέρεται καλύτερη εξυπηρέτηση στους πελάτες τους οι εταιρείες κατευθύνονται προς την ανάπτυξη ενός σημασιολογικά δομημένου συστήματος, ώστε να επιτύχουν αποδοτική αναζήτηση και επεξεργασία δεδομένων. Προς αυτή την κατεύθυνση κινούνται και οι οντολογίες.

### 2.2 Βασική παρουσίαση των οντολογιών

Ο όρος «οντολογία» έχει μακρά ιστορία που προέρχεται από τη φιλοσοφία και αναφέρεται στον κλάδο αυτό της μεταφυσικής που ασχολείται με τη φύση και την οργάνωση των όντων. Η επιστήμη και η Τεχνητή Νοημοσύνη υιοθέτησαν τον όρο αυτό για να δηλώσουν «μία διαμοιρασμένη και κοινή κατανόηση κάποιου τομέα, η οποία μπορεί να ανταλλαγεί μεταξύ ανθρώπων και συστημάτων εφαρμογών» [12].

Για την επιστήμη οι οντολογίες είναι εργαλεία αναπαράστασης (representation) και συλλογιστικής (reasoning). Εκτενέστερα αλλά και πιο συγκεκριμένα, η οντολογία ορίζεται ως μία τυπική (formal), κατηγορηματική (explicit) προδιαγραφή μίας διαμοιρασμένης (shared) εννοιολογικής αναπαράστασης (conceptualization). Ο όρος *εννοιολογική αναπαράσταση* αναφέρεται σε ένα αφηρημένο μοντέλο φαινομένων του κόσμου στον οποίο έχουν προσδιοριστεί οι έννοιες που σχετίζονται με τα φαινόμενα αυτά. Ο όρος *κατηγορηματική* σημαίνει ότι το είδος των εννοιών που χρησιμοποιούνται, και οι περιορισμοί που αφορούν τη χρήση των εννοιών αυτών είναι προσδιορισμένα με σαφήνεια. Ο όρος *τυπική* αναφέρεται στο ότι η οντολογία πρέπει να είναι μηχανικά

αναγνώσιμη. Ο όρος *διαμοιρασμένη* αναφέρεται στην γνώση κοινής αποδοχής που πρέπει να αποτυπώνει η οντολογία στα πλαίσια μίας κοινότητας. [12]

Με την ανάπτυξη των οντολογιών γίνεται μία προσπάθεια να περιγράψει ένα πεδίο γνώσης (Knowledge Domain), ώστε να υπάρχει ένας κοινός τρόπος επικοινωνίας μεταξύ εκείνων που δραστηριοποιούνται στο συγκεκριμένο πεδίο. Επιπλέον, χρειάζεται να είναι αυστηρά θεμελιωμένη (formal model), ώστε να είναι μηχανικά κατανοητή και επεξεργάσιμη (machine readable).

Οι οντολογίες αποτελούν βασική πηγή για τις εφαρμογές στο Σηματολογικό Ιστό, και όχι μόνο, καθώς παρέχουν μία γλώσσα αναπαράστασης γνώσης και ένα λεξικό από *κλάσεις (classes)* και *σχέσεις (relations)*, τις οποίες οι διάφορες υπηρεσίες μπορούν να χρησιμοποιήσουν για να περιγράψουν το περιεχόμενό τους και να το επεξεργαστούν κατάλληλα.[1]

Τί όμως τελικά είναι οι οντολογίες και ποίος είναι ο ρόλος των *κλάσεων* και των *σχέσεων*; Μία οντολογία αποτελείται από την ταξινόμια (taxonomy), η οποία και την χαρακτηρίζει, και από ένα σύνολο συμπερασματικών κανόνων. Η ταξινόμια καθορίζει τις κλάσεις των αντικειμένων και τις μεταξύ τους σχέσεις, ενώ οι συμπερασματικοί κανόνες αναπαριστούν τη λογική. Για παράδειγμα, η πρόταση «η *παθολογία* αποτελεί κλάδο της *Ιατρικής*» αποτελεί έναν συμπερασματικό κανόνα, ο οποίος απορρέει ως συμπέρασμα μίας ταξινόμιας, στην οποία η παθολογία ορίζεται ως υποκατηγορία της Ιατρικής. Επιπλέον, οι όροι “παθολογία” και “Ιατρική” θεωρούνται κλάσεις του πεδίου «Ιατρική». Οι κλάσεις μαζί με τις παρακάτω κατηγορίες αποτελούν τα βασικά συστατικά μίας οντολογίας.[4]

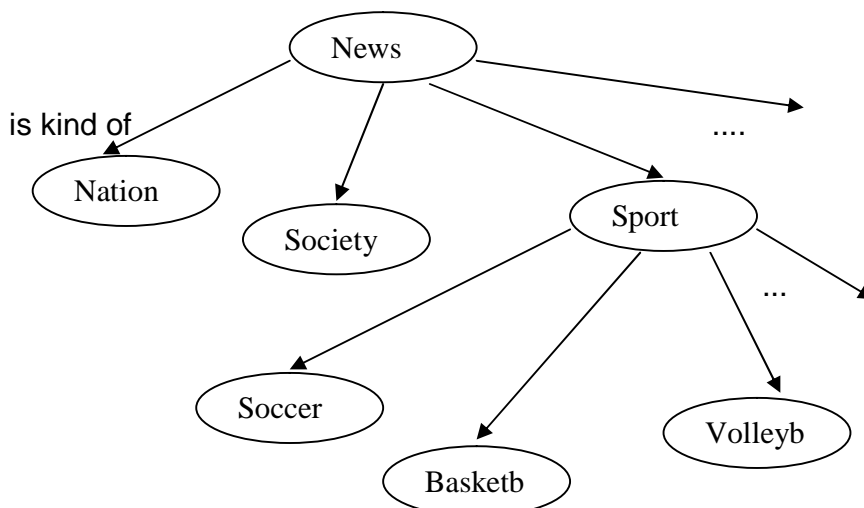
Μία οντολογία περιέχει πέντε βασικές κατηγορίες συστατικών:

- Κλάσεις (Classes): Οι κλάσεις, όπως είδαμε προηγουμένως είναι έννοιες που σχετίζονται με ένα πεδίο. Για παράδειγμα, σε μία οντολογία που

αφορά το πεδίο των ειδήσεων η «πολιτική» ή τα «αθλητικά» αποτελούν κλάσεις.

- Σχέσεις (relations): Ένας τύπος αλληλεπίδρασης μεταξύ των εννοιών ενός πεδίου. Όπως για παράδειγμα: Subclass-of, is – a και άλλα.
- Συναρτήσεις (functions): Μία περίπτωση σχέσης κατά την οποία το ν-οστό στοιχείο της σχέσης προσδιορίζεται μοναδικά από τα ν-1 προηγούμενα στοιχεία.
- Αξιώματα (axioms): αναπαριστούν προτάσεις, οι οποίες είναι αληθείς. Για παράδειγμα, αν ο φοιτητής Β. παρακολουθεί το 2<sup>ο</sup> έτος της σχολής τότε μπορεί να εγγραφεί στο μάθημα Ψ.
- Στιγμιότυπα (instances): Αναπαριστούν συγκεκριμένα στοιχεία. Για παράδειγμα, ο φοιτητής με το όνομα Νίκος είναι ένα στιγμιότυπο της κλάσης «φοιτητής».

Η παρακάτω εικόνα (εικόνα 2.2.) αποτελεί έναν απλοποιημένο σχεδιασμό μίας οντολογίας που περιέχει κλάσεις και σχέσεις.



Εικόνα 2.2: Σχεδιάγραμμα διακλάδωσης μίας οντολογίας ειδησεογραφικού περιεχομένου

Οι οντολογίες ανάλογα με το είδος της πληροφορίας που καλούνται κάθε φορά να αναπαραστήσουν χωρίζονται σε ορισμένες κατηγορίες. Ορισμένες από τις πιο βασικές περιγράφονται παρακάτω:

- *Γενικές ή κοινές οντολογίες* (generic ή common sense ontologies ή Top-level ontologies): Περιγράφουν πολύ γενικές έννοιες (όπως ο χρόνος, ο χώρος, το γεγονός), οι οποίες είναι ανεξάρτητες από ένα συγκεκριμένο πρόβλημα ή περιοχή γνώσης.
- *Οντολογίες πεδίου ορισμού* (Domain Ontologies): Είναι οι οντολογίες εκείνες που σκοπό έχουν να αναπαραστήσουν γνώση γύρω από ένα συγκεκριμένο πεδίο, όπως η ιατρική, η ειδησεογραφία κλπ.
- *Οντολογίες μεθοδολογίας ή εργασιών* (method or task ontologies): περιγράφουν λεξιλόγιο για μία γενική δραστηριότητα ή εργασία, π.χ. διάγνωση.
- *Οντολογίες μεταδεδομένων* (metadata ontologies): παρέχουν ορολογία για την περιγραφή πληροφορίας, η οποία διατίθεται ηλεκτρονικά. Η ορολογία μπορεί να περιέχει μεταδεδομένα για συγκεκριμένο πεδίο γνώσης.
- *Οντολογίες αναπαράστασης* (representational ontologies): παρέχουν οντότητες αναπαράστασης χωρίς να προσδιορίζουν τι συγκεκριμένο αναπαριστούν. Για παράδειγμα η *Frame ontology* ορίζει έννοιες όπως frames, slots κλπ.
- *Οντολογίες εφαρμογής* (application ontologies): Είναι οι πιο συγκεκριμένες οντολογίες, που σκοπό έχουν να καλύψουν τις ανάγκες μίας συγκεκριμένης εφαρμογής.

**Οι οντολογίες προσφέρουν ένα εννοιολογικό μοντέλο ενός πεδίου ενδιαφέροντος καθιστώντας τη γνώση επαναχρησιμοποιήσιμη και διαμοιραζόμενη. Λόγω, λοιπόν, της πρακτικότητας και της δυναμικότητάς τους, οι οντολογίες έχουν καταφέρει να γίνουν δημοφιλείς σε πολλές ερευνητικές ομάδες και επιχειρησιακές κοινότητες και να αποτελέσουν ακρογωνιαίο λίθο σε πολλές εφαρμογές. Ενδεικτικά αναφέρονται τα παρακάτω:**

- Συστήματα διαχείρισης γνώσης (Knowledge management), όπου χρησιμοποιούνται οντολογίες με σκοπό τη μοντελοποίηση της γνώσης και τη «μετάφραση της πληροφορίας».
- Ηλεκτρονικό εμπόριο (e-commerce), για την ανάκτηση γνώσης και πληροφορίας. Για τη διαδικασία αυτή οι οντολογίες χρησιμοποιούνται από πράκτορες λογισμικού (Software agents). Οι πράκτορες λογισμικού είναι προγράμματα που εκτελούν κάποια λειτουργία και με το πέρας της εκτέλεσης αυτής παράγουν αποτελέσματα. Συνήθως, επισκέπτονται ιστοσελίδες και επεξεργάζονται τις πληροφορίες που βρίσκουν σε αυτές που επισκέπτονται. Οι λειτουργίες για τις οποίες χρησιμοποιούνται είναι διάφορες, όπως η εύρεση, ταξινόμηση και επιλογή δεδομένων. Στο ηλεκτρονικό εμπόριο εκτελούν λειτουργίες όπως η σύγκριση τιμών του ίδιου προϊόντος σε διάφορα ηλεκτρονικά καταστήματα.
- Διαλειτουργικότητα (interoperability) μεταξύ συστημάτων: Οι οντολογίες προσφέρουν μία κοινή γλώσσα σε διάφορους χρήστες για επικοινωνητική ανταλλαγή δεδομένων ή χρησιμοποιούνται ως μεταφραστές διαφορετικών πακέτων λογισμικού. Επιπλέον, οι οντολογίες χρησιμοποιούνται για την υποστήριξη μετάφρασης μεταξύ διαφορετικών γλωσσών και αναπαραστάσεων.
- Ψηφιακές Βιβλιοθήκες (Digital libraries), για καλύτερη ταξινόμηση του υλικού τους.
- Στα πολυμέσα (Multimedia) και στις Τηλεπικοινωνίες (Telecommunications).

## 2.3 Λόγοι ανάπτυξης οντολογιών

Τα τελευταία χρόνια η ανάπτυξη των οντολογιών έχει αποτελέσει αντικείμενο ενδιαφέροντος σ' ένα ευρύ πεδίο ερευνητών. Στο Διαδίκτυο μπορεί να βρεί κανείς διάφορες οντολογίες, από μεγάλες ταξινομίες κατηγοριοποίησης Ιστότοπων, όπως το yahoo, μέχρι κατηγοριοποιήσεις προϊόντων προς πώληση (όπως στο Amazon). Γιατί κρίνεται όμως αναγκαία η ανάπτυξη και τελικά η εφαρμογή οντολογιών; Παρακάτω γίνεται μία προσπάθεια να μελετηθούν οι λόγοι που υποδεικνύουν την σκοπιμότητα των οντολογιών.

Τον πιο διαδεδομένο λόγο δημιουργίας οντολογιών αποτελεί η *διαμοιρασμένη και κοινή κατανόηση* της δόμησης της πληροφορίας τόσο ανάμεσα στους ανθρώπους όσο και ανάμεσα στους πράκτορες λογισμικού (Software agents) [Gruber] . Για παράδειγμα, διαφορετικές ιστοσελίδες που παρέχουν ιατρικές πληροφορίες ή προωθούν ιατρικές υπηρεσίες ηλεκτρονικού εμπορίου (e-commerce) μπορεί να μοιράζονται και να εκδίδουν κοινή οντολογία με όλους τους όρους που χρησιμοποιούν. Μ' αυτό τον τρόπο, οι πράκτορες λογισμικού μπορούν να εξαγάγουν και να συλλέξουν πληροφορίες από όλες αυτές τις διαφορετικές ιστοσελίδες[10].

Η δυνατότητα επαναχρησιμοποίησης της γνώσης ενός συγκεκριμένου πεδίου (domain knowledge) συντέλεσε σε μεγάλο βαθμό στην ανάπτυξη των οντολογιών. Για παράδειγμα, μοντέλα για πολλά διαφορετικά πεδία γνώσης χρησιμεύουν για την αναπαράσταση του *χρόνου*, π.χ. ώρες αφίξεων, σχετικές μετρήσεις του χρόνου κλπ. Αν, λοιπόν, αναπτυχθεί από μία ομάδα ερευνητών μία σχετική λεπτομερής οντολογία, μπορεί να επαναχρησιμοποιηθεί από κάποιους άλλους, οι οποίοι μπορεί να την χρειάζονται για διαφορετικού είδους υπηρεσίες. Επιπλέον, για την δημιουργία μία μεγάλης οντολογίας, μπορούν να χρησιμοποιηθούν ήδη υπάρχουσες οντολογίες που περιγράφουν μέρη ενός πεδίου και να ενοποιηθούν. Επίσης, γενικές οντολογίες μπορούν να επεκταθούν, προκειμένου να περιγράψει κανείς συγκεκριμένα πεδία.

Ακόμα, μπορεί κάποιος να αναπτύξει μία οντολογία προκειμένου να ορίσει ρητά (explicit) τις υποθέσεις που ισχύουν σε ένα επιστημονικό πεδίο. Με τον τρόπο αυτό, σε περίπτωση που αλλάξει η γνώση μας για ένα πεδίο μπορούν εύκολα να πραγματοποιηθούν οι απαραίτητες αλλαγές.

Μία ακόμα ευρεία χρήση των οντολογιών συνιστά ο διαχωρισμός του πεδίου γνώσης (domain knowledge) από τη λειτουργική γνώση (operational knowledge). Μ' αυτό τον τρόπο, μπορούμε να περιγράψουμε τη διαδικασία διαμόρφωσης (configuring) ενός προϊόντος από τα συστατικά του σύμφωνα με την απαιτούμενη τεχνική περιγραφή.

Επίσης, η ανάλυση γνώσης συγκεκριμένου πεδίου (Domain Knowledge) γίνεται δυνατή, καθώς διατίθεται μία προτυποποιημένη περιγραφή των όρων. Η τυπική ανάλυση των όρων είναι ιδιαίτερα χρήσιμη σε περίπτωση είτε επαναχρησιμοποίησης ήδη υπάρχοντων οντολογιών είτε επέκτασης κάποιων άλλων γενικών οντολογιών.

Η λέξη "οντολογία" είναι ένας όρος δανεισμένος από τη φιλοσοφία, η οποία αναφέρεται στην επιστήμη της περιγραφής των διαφόρων ειδών οντοτήτων στον κόσμο και του τρόπου που αυτές συσχετίζονται. Συχνά, συγχέεται με την επιστημολογία που αφορά τη γνώση, ενώ έχει δημιουργήσει πολλές διαμάχες στις συζητήσεις για την τεχνητή νοημοσύνη.

Στα πλαίσια της κατανεμημένης γνώσης, οντολογία είναι ο προσδιορισμός μιας σύλληψης, δηλαδή, μια περιγραφή των εννοιών και των σχέσεων που μπορούν να υπάρξουν για έναν πράκτορα λογισμικού ή μια κοινότητα τέτοιων πρακτόρων. Είναι μια αναγνώσιμη από τον υπολογιστή περιγραφή της γνώσης των πόρων στο δίκτυο. Αυτός ο ορισμός είναι σύμφωνος με τη χρήση της οντολογίας ως σύνολο ορισμών εννοιών, αλλά γενικότερος. Λέμε ότι ένας πράκτορας λογισμικού "δεσμεύεται" σε μια οντολογία εάν οι ενέργειές του είναι σύμφωνες με τους ορισμούς στην οντολογία. Πρακτικά, μια κοινή οντολογία καθορίζει το λεξιλόγιο με το οποίο ερωτήσεις και ισχυρισμοί ανταλλάσσονται μεταξύ των πρακτόρων λογισμικού. Οι οντολογικές δεσμεύσεις είναι συμφωνίες χρησιμοποίησης του κοινού λεξιλογίου κατά τρόπο κατανοητό και συνεπή. Οι πράκτορες λογισμικού που μοιράζονται ένα λεξιλόγιο δεν χρειάζονται να μοιράζονται μια βάση γνώσεων και καθένας από αυτούς δεν απαιτείται να απαντήσει σε όλες τις ερωτήσεις που μπορούν να διατυπωθούν στο κοινό λεξιλόγιο. Με τις οντολογίες οι πράκτορες λογισμικού γίνονται ευφείς επειδή μπορούν να χρησιμοποιήσουν τη γνώση που περιλαμβάνεται στην οντολογία στο στάδιο της διαπραγμάτευσης και της λήψης αποφάσεων .

Στην επιστήμη των πληροφοριών, μια οντολογία καθορίζει τους όρους που χρησιμοποιούνται για να περιγράψουν και να αναπαραστήσουν έναν τομέα της γνώσης, με σκοπό τη διευκόλυνση της διανομής και της επαναχρησιμοποίησης της γνώσης αυτής. Μια οντολογία είναι το προϊόν μιας προσπάθειας να διατυπωθεί ένα εξαντλητικό

και αυστηρό εννοιολογικό σχήμα για μια περιοχή. Είναι μία ιεραρχική δομή δεδομένων που περιέχει όλες τις σχετικές οντότητες, τις μεταξύ τους σχέσεις και κανόνες εξαγωγής συμπερασμάτων μέσα σε εκείνη την περιοχή. Οι οντολογίες χρησιμοποιούνται συνήθως στην τεχνητή νοημοσύνη και στην αναπαράσταση γνώσης. Τα προγράμματα υπολογιστών μπορούν να χρησιμοποιήσουν μια οντολογία για ποικίλους λόγους συμπεριλαμβανομένου του επαγωγικού συλλογισμού, της ταξινόμησης, των ποικίλων τεχνικών επίλυσης προβλημάτων, καθώς επίσης και για να διευκολύνουν την επικοινωνία και τη διανομή των πληροφοριών μεταξύ των διαφορετικών συστημάτων. Χρησιμοποιούνται από τους ανθρώπους, τις βάσεις δεδομένων, και τις εφαρμογές που πρέπει να μοιραστούν τις πληροφορίες περιοχών (μια περιοχή είναι ακριβώς μία συγκεκριμένη θεματική περιοχή ή τομέας της γνώσης, όπως η ιατρική, η κατασκευή εργαλείων, η ακίνητη περιουσία, η γεωγραφία, η οικονομική διαχείριση, κ.λπ.). Οι οντολογίες περιλαμβάνουν χρησιμοποιούμενους από υπολογιστές ορισμούς βασικών εννοιών μίας θεματικής περιοχής και τις σχέσεις μεταξύ τους. Κωδικοποιούν τη γνώση σε μια περιοχή, καθώς και τη γνώση από περισσότερες από μία περιοχές. Έτσι, καθιστούν την γνώση αυτή επαναχρησιμοποιήσιμη.[6]

Η οντολογική ανάλυση αποσαφηνίζει τη δομή της γνώσης. Δεδομένης μιας περιοχής, η οντολογία της διαμορφώνει τον πυρήνα οποιουδήποτε συστήματος αναπαράστασης γνώσης για αυτή την περιοχή. Χωρίς τις οντολογίες, ή τις συλλήψεις που κρύβονται κάτω από τη γνώση, δεν μπορεί να υπάρξει ένα λεξιλόγιο για την αναπαράσταση της γνώσης.

Η λέξη “οντολογία” έχει χρησιμοποιηθεί για να περιγράψει τεχνήματα διαφορετικών βαθμών δομής. Αυτά κυμαίνονται από τις απλές ταξονομίες, ως τα σχέδια μετα-δεδομένων και τις λογικές θεωρίες. Το Semantic Web χρειάζεται οντολογίες με έναν σημαντικό βαθμό δομής και οι οντολογίες αυτές πρέπει να διευκρινίζουν τις περιγραφές για τα ακόλουθα είδη εννοιών:

- Κλάσεις (γενικά πράγματα) στις πολλές περιοχές ενδιαφέροντος
- Οι σχέσεις που μπορούν να υπάρξουν μεταξύ των πραγμάτων
- Οι ιδιότητες (ή χαρακτηριστικά) των πραγμάτων (slots)



Οι οντολογίες εκφράζονται συνήθως σε μια βασισμένη στη λογική γλώσσα, έτσι ώστε λεπτομερείς, ακριβείς, συνεπείς και σημαντικές διακρίσεις να μπορούν να γίνουν μεταξύ των κλάσεων, των ιδιοτήτων, και των σχέσεων. Μερικά εργαλεία οντολογίας μπορούν να εκτελέσουν αυτοματοποιημένο συλλογισμό χρησιμοποιώντας τις οντολογίες, και να παρέχουν έτσι προηγμένες υπηρεσίες σε ευφυείς εφαρμογές όπως: σημασιολογική αναζήτηση και ανάκτηση, πράκτορες λογισμικού, υποστήριξη αποφάσεων, ομιλία και φυσική κατανόηση γλώσσας, διαχείριση γνώσης, ευφυείς βάσεις δεδομένων, και ηλεκτρονικό εμπόριο.

Οι οντολογίες φαίνεται ότι θα χρησιμοποιηθούν στο Semantic Web ως τρόπος αναπαράστασης της σημασιολογίας των εγγράφων και προσφοράς της ικανότητας σε web εφαρμογές και ευφυείς πράκτορες να χρησιμοποιήσουν τη σημασιολογία αυτή. Οι οντολογίες μπορούν να αποδειχθούν πολύ χρήσιμες για μια κοινότητα ως τρόπος δόμησης και ορισμού της έννοιας των μεταδεδομένων που συλλέγονται αυτήν την περίοδο και τυποποιούνται. Χρησιμοποιώντας τις οντολογίες, οι αυριανές εφαρμογές μπορούν να είναι “ευφυείς”, υπό την έννοια ότι μπορούν να λειτουργήσουν ακριβέστερα στο ανθρώπινο εννοιολογικό επίπεδο.[26]

Οι οντολογίες μπορούν να ενισχύσουν τη λειτουργία του Web από πολλές απόψεις. Μπορούν να χρησιμοποιηθούν για να βελτιώσουν την ακρίβεια των Web αναζητήσεων, το πρόγραμμα αναζήτησης μπορεί να ψάχνει μόνο εκείνες τις σελίδες που αναφέρονται σε μια ακριβή έννοια αντί όλων των άλλων εννοιών των διαφορούμενων λέξεων κλειδιών. Οι πιο προηγμένες εφαρμογές μπορούν να χρησιμοποιήσουν τις οντολογίες για να συσχετίσουν τις πληροφορίες μιας σελίδας με τις σχετικές δομές γνώσης και κανόνες εξαγωγής συμπερασμάτων. Επιπλέον, καθιστούν πολύ ευκολότερη την ανάπτυξη προγραμμάτων που μπορούν να αντιμετωπίσουν περίπλοκες ερωτήσεις, των οποίων οι απαντήσεις δεν βρίσκονται σε μία ενιαία ιστοσελίδα.

Ακόμη, οι οντολογίες μπορούν να χρησιμοποιηθούν για να υποστηρίξουν συστήματα τεχνητής νοημοσύνης (AI), με την παροχή μιας βαθύτερης και σθεναρότερης αντιπροσώπευσης της περιοχής στην οποία κάποιος επιθυμεί να εξάγει συλλογισμούς και να λύσει προβλήματα. Σε αυτές εφαρμόζεται, αυστηρά, ο παραγωγικός συλλογισμός

και όχι λογική ασαφής ή βασισμένη σε πιθανότητες, ή οποιοσδήποτε συλλογισμός που προσπαθεί να μιμηθεί τη συνείδηση.

Μια οντολογία διαφέρει από ένα σχήμα XML διότι είναι μια αναπαράσταση γνώσης, όχι ένα σχήμα μηνυμάτων. Τα περισσότερα βασισμένα στη βιομηχανία πρότυπα Ιστού αποτελούνται από έναν συνδυασμό σχημάτων μηνυμάτων και προδιαγραφών πρωτοκόλλου. Σε αυτά τα σχήματα έχει δοθεί μια λειτουργική σημασιολογία αλλά η προδιαγραφή δεν έχει ως σκοπό να υποστηρίξει το συλλογισμό έξω από τα πλαίσια αυτής της λειτουργίας.

Οι οντολογίες είναι κρίσιμες για τις εφαρμογές που θέλουν να αναζητήσουν ή να συγχωνεύσουν πληροφορίες από διαφορετικές κοινότητες. Αν και τα XML DTDs και XML σχήματα είναι επαρκή για την ανταλλαγή των στοιχείων μεταξύ των συμβαλλόμενων μερών που έχουν συμφωνήσει ορισμούς εκ των προτέρων, η έλλειψη σημασιολογίας τους αποτρέπει τις μηχανές από την αξιόπιστη εκτέλεση αυτής της εργασίας όταν τους δίνονται νέα λεξιλόγια XML. Ο ίδιος όρος μπορεί να χρησιμοποιηθεί με διαφορετική έννοια σε διαφορετικές εκφράσεις, και οι διαφορετικοί όροι μπορούν να χρησιμοποιηθούν για αντικείμενα που έχουν την ίδια έννοια. Η RDF και το RDF σχήμα αρχίζει να προσεγγίζει αυτό το πρόβλημα επιτρέποντας σε απλές σημασιολογίες να συνδεθούν με όρους. Με το σχήμα RDF, μπορούν να καθοριστούν κλάσεις, που μπορούν να έχουν πολλαπλές υποκλάσεις και υπερκλάσεις, και ιδιότητες, οι οποίες μπορούν να έχουν υποιδιότητες, πεδία ορισμού και σύνολα τιμών. Από αυτή την άποψη, το σχήμα RDF είναι μια απλή γλώσσα οντολογίας. Εντούτοις, προκειμένου να επιτευχθεί η λειτουργικότητα μεταξύ πολυάριθμων, αυτόνομα αναπτυσσόμενων και διοικούμενων σχημάτων, απαιτείται πλουσιότερη σημασιολογία. Παραδείγματος χάριν, το σχήμα RDF δεν μπορεί να διευκρινίσει ότι οι κλάσεις προσώπων και αυτοκινήτων είναι διαχωρισμένες, ή ότι ένα κουαρτέτο έχει ακριβώς τέσσερις μουσικούς ως μέλη.

Μια οντολογία που είναι συνδεδεμένη με μια ιδιαίτερη περιοχή ονομάζεται domain ontology. Η αποκαλούμενη foundation ontology ή upper ontology είναι ένα είδος οντολογίας που προσπαθεί να περιγράψει γενικές οντότητες και δεν είναι τόσο συγκεκριμένη, οπότε είναι γενικά πιο εφαρμόσιμη. Περιέχει ένα βασικό γλωσσάριο, με

τους όρους του οποίου οτιδήποτε άλλο σε μια ευρεία περιοχή μπορεί και πρέπει να περιγραφεί.

## 2.4 Τι είναι Οντολογία

Ο πιο διαδεδομένος ορισμός της οντολογίας είναι αυτός του Gruber: Οντολογία είναι ένας τυπικός, σαφής προσδιορισμός μιας κοινής αντίληψης (*An ontology is a formal, xplicit, specification of a formal conceptualization*). Ως «αντίληψη» αναφέρεται ένα αφηρημένο μοντέλο των φαινομένων του κόσμου που περιλαμβάνει τους σχετικούς όρους με αυτά τα φαινόμενα. Ο όρος «σαφής» σημαίνει ότι ο τύπος των εννοιών που χρησιμοποιούνται, και των περιορισμών στη χρήση τους είναι σαφώς ορισμένοι, και ο όρος «τυπικός» αναφέρεται στο γεγονός ότι η οντολογία πρέπει να είναι αναγνώσιμη από μηχανές. Τέλος ο όρος «κοινής» αντικατοπτρίζει το γεγονός ότι η οντολογία πρέπει να συλλαμβάνει κοινά αποδεκτή γνώση από την κοινωνία.

Η οντολογία σαν όρος είναι δανεισμένος από την φιλοσοφία, όπου οντολογία, είναι μια συστηματική εξήγηση της οντότητας (existence). Για την Τεχνητή Νοημοσύνη, ότι υπάρχει μπορεί να αναπαρασταθεί. Η γνώση για μια περιοχή μπορεί να αναπαρασταθεί από ένα σύνολο τυπικών δηλώσεων, τότε το σύνολο αντικειμένων που μπορούν να αναπαρασταθούν κατασκευάζουν το λεξικό του διαλόγου. Αυτό το σύνολο αντικειμένων και οι σχέσεις μεταξύ τους αντικατοπτρίζονται στο αντιπροσωπευτικό λεξικό το οποίο χρησιμοποιείται για να αναπαραστήσει τη γνώση. Αυτό γίνεται και στο χώρο της πληροφορικής, όπου οντολογία είναι ο κώδικας ο οποίος ορίζει σύνολα αναπαριστώμενων όρων. Τα ονόματα των αντικειμένων από το λεξικό του διαλόγου συνδέονται με λέξεις που περιγράφουν τις έννοιες που έχουν αυτά για τους ανθρώπους (καθώς οριζόμενα αξιώματα περιορίζουν την ερμηνεία τους μέσα στο συγκεκριμένο πεδίο γνώσης). Τυπικά η οντολογία είναι μία δήλωση μιας λογικής θεωρίας.[12]

## 2.5 Ορισμός της οντολογίας

**Οντολογία** είναι η περιγραφή, με τη χρησιμοποίηση ενός συγκεκριμένου λεξιλογίου, ενός συνόλου από έννοιες, αντικείμενα και σχέσεις μεταξύ τους που αφορούν μία συγκεκριμένη περιοχή γνώσης. Ουσιαστικά, μία οντολογία είναι μία ιεραρχία από κλάσεις, ιδιότητες και στιγμιότυπα των κλάσεων, που περιγράφουν ένα γνωστικό αντικείμενο.[12]

### Υπάρχουν και άλλοι ορισμοί για τον όρο οντολογία. Οντολογία είναι...

1. η τεκμηρίωση (documentation) μιας ορολογίας που χρησιμοποιείται σε κάποια περιοχή γνώσης. Περιλαμβάνει τους κανόνες που επιτρέπουν το συνδυασμό και τη χρήση της ορολογίας για την δημιουργία δηλώσεων (statements) ή την δημιουργία συμπερασμάτων που προκύπτουν από τις δηλώσεις σχετικά με την περιοχή γνώσης. Οι δηλώσεις αυτές περιορίζουν τις πιθανές ερμηνείες των ορισμένων όρων.
2. ένας κατάλογος όρων που ορίζονται σε μια μορφή που είναι «αναγνώσιμη» από ανθρώπους και μηχανές
3. ένας κατάλογος από είδη εννοιών-αντικείμενων που χαρακτηρίζουν μια περιοχή γνώσης, τις ιδιότητες τους, τις μεταξύ τους σχέσεις και τους περιορισμούς πάνω και μεταξύ των σχέσεων τους.
4. μια τυπική αναπαράσταση της γνώσης. Βασίζεται σε μια κοινή αντίληψη: αντικείμενα, έννοιες, και άλλες οντότητες που ανήκουν σε μια περιοχή γνώσης και οι σχέσεις μεταξύ τους. Η κοινή αντίληψη είναι μία αφηρημένη αλλά και απλοποιημένη άποψη του κόσμου που χρειαζόμαστε να αναπαραστήσουμε.

Η Οντολογία επιτρέπει την εξαγωγή συμπερασμάτων σχετικά με την ομοιότητα ή ανομοιότητα εννοιών και δημιουργεί αντιστοιχίες (mapping) που επιτρέπουν την εδραίωση εννοιολογικά βάσιμων καναλιών επικοινωνίας. Είναι βασικό εδώ να τονίσουμε ότι η οντολογία μαζί με το σύνολο των κλάσεων της συνιστά τη βάση γνώσης. Αυτές οι έννοιες είναι πολύ λεπτές και πρέπει να διαχωρίζονται για τις εφαρμογές.

## 2.6 Η ανάπτυξη οντολογιών αφορά:

- τον ορισμό των κλάσεων της οντολογίας
- την ιεραρχική ταξινόμηση των κλάσεων
- τον ορισμό ιδιοτήτων στις κλάσεις καθώς και τον ορισμό τιμών σε αυτές
- τη δημιουργία των instances
- τον ορισμό περιορισμών για τα instances

## 2.7 Σχεδιασμός Οντολογιών – Αντικειμενοστραφής σχεδιασμός

Παρόλο που ο σχεδιασμός των οντολογιών προήλθε βασίστηκε στην μεθοδολογία αντικειμενοστραφών συστημάτων, ωστόσο διαφέρουν, αφού στον αντικειμενοστραφή σχεδιασμό η εστίαση γίνεται στον σχεδιασμό κλάσεων και σχέσεων, δηλαδή ο προγραμματιστής παίρνει αποφάσεις σχετικά με τις *λειτουργικές* ιδιότητες της κλάσης, ενώ στην οντολογία οι αποφάσεις βασίζονται στην *δομική* ιδιότητα των κλάσεων. Δηλαδή, οι κλάσεις και οι σχέσεις μεταξύ τους στις οντολογίες διαφέρουν από τη δομή των κλάσεων σε αντικειμενοστραφή συστήματα.

## 2.8 Κατηγορίες Οντολογιών

Οι οντολογίες οργανώνονται σε τρεις κατηγορίες [Abecker ]:

- Οντολογίες πληροφοριών (Information Ontologies): περιγράφουν το πληροφοριακό μέτα-μοντέλο (meta-model) π.χ. τη δομή και μορφή των πηγών πληροφορίας. Είναι η πιο χαμηλού επιπέδου οντολογία
- Οντολογίες περιοχής γνώσης (Domain Ontologies): περιγράφουν το περιεχόμενο πηγών πληροφορίας (π.χ. ιατρική, μηχανική, κλπ)
- Οντολογίες επιχειρηματικών διαδικασιών (Enterprise Ontologies): μοντελοποιούν τις επιχειρηματικές διαδικασίες με σκοπό να μοντελοποιηθεί η ανάγκη γνώσης σε μια διαδικασία, επιτρέποντας έτσι αποτελεσματική διανομή γνώσης.[40]

## 2.9 Χαρακτηριστικά της οντολογίας

Η γνώση τυποποιείται στις οντολογίες με πέντε είδη στοιχείων:

- Κλάσεις, οργανωμένες σε ταξινομίες. Μια έννοια (κλάση) είναι οτιδήποτε, όπως μια περιγραφή μιας εργασίας, συνάρτησης, ενέργειας, στρατηγικής διαδικασίας, κλπ.
- Σχέσεις, αναπαριστούν τον τύπο αλληλεπίδρασης μεταξύ των εννοιών στο χώρο
- Συναρτήσεις, είναι ειδικού τύπου σχέσεις στις οποίες το n-οστό στοιχείο είναι μοναδικός συνδυασμός των n-1 προηγούμενων στοιχείων
- Αξιώματα, χρησιμοποιούνται για να μοντελοποιήσουν δηλώσεις που είναι πάντοτε αληθείς
- Instances, χρησιμοποιούνται για να δηλώσουν τα στοιχεία της βάσης γνώσης (γνωστικά αντικείμενα)[40]

## Τα παρακάτω συνοψίζουν χαρακτηριστικά των οντολογιών:

- **Εξαγωγή συμπερασμάτων (Inferencing):** Είναι η διαδικασία κατά την οποία χρησιμοποιούνται τα αξιώματα /κανόνες (rules/axioms) για τη δημιουργία συμπερασμάτων/κανόνων. Με άλλα λόγια είναι η δημιουργία νέας γνώσης από υπάρχουσα γνώση. Παράδειγμα τέτοιας διαδικασίας αποτελεί η μεταβατικότητα, που επιτρέπει την εξαγωγή συμπερασμάτων του τύπου:  
IF (A IS\_HIGHER\_THAN B) AND (B IS\_HIGHER\_THAN C)  
THEN (A IS\_HIGHER\_THAN C).
- **Επεκτασιμότητα (Extensibility):** Οι ακριβείς ανάγκες στη χρήση είναι δύσκολο να καθοριστούν από την αρχή, για το λόγο αυτό απαιτείται η οντολογία να είναι επεκτάσιμη τόσο ως προς instances όσο και ως προς έννοιες και σχέσεις.
- **Σαφήνεια (Clarity):** Το στοχευόμενο μήνυμα χρειάζεται να μπορεί να μεταφερθεί αποδοτικά, ελαττώνοντας την ασάφεια.
- **Συνοχή (Coherence):** Η οντολογία πρέπει να διακρίνεται από εσωτερική συνέπεια. Τα αξιώματα ορισμού και οι έννοιες πρέπει να είναι λογικά συνεπή.
- **Ελαχιστοποίηση της οντολογικής δέσμευσης (minimal ontological commitment):** Αναφέρεται στην γενίκευση των αξιωμάτων με τρόπο που να αποφεύγεται σαφής αναφορά στον κόσμο που μοντελοποιείται, έτσι ώστε να υπάρχει ελευθερία για εξειδίκευση και εφαρμογή σε συγκεκριμένες περιοχές.
- **Ελαχιστοποίηση της μεροληψίας στην κωδικοποίηση (minimal coding bias):** Η οντολογία πρέπει να στοχεύει στο γνωστικό επίπεδο και όχι στην συμβολική αναπαράσταση της. Ο στόχος είναι να μπορεί να χρησιμοποιηθεί από διαφορετικά συστήματα με διαφορετικές προσεγγίσεις αναπαράστασης.
- **Κληρονομικότητα (Inheritance):** Οι κλάσεις ταξινομούνται στην ιεραρχία με την σχέση subclass\_of κάνοντας την ερώτηση: «αν ένα αντικείμενο είναι instance μιας κλάσης τότε αυτό θα είναι απαραίτητα αντικείμενο και μιας άλλης κλάσης» ή “if a class B is subclass\_of A, then every instance of B is also an instance of A”.
- **Αντίστροφες σχέσεις (Inverse Relations):** Αντίστροφες σχέσεις είναι αυτές του τύπου: «A produces B» , «B is\_produced\_by A». Παρόλο που η αποθήκευση και των δύο σχέσεων μπορεί να φαίνεται υπερβολή, αφού είναι θέμα απλής

συνεπαγωγής, όμως στα συστήματα διαχείρισης γνώσης είναι χρήσιμο να ορίζονται οι αντίστροφες σχέσεις έτσι ώστε το σύστημα να μπορεί να κάνει από μόνο του τη συνεπαγωγή.[6]

### **2.10.1 Περιοχές εφαρμογής οντολογιών: Διαχείριση Γνώσης και Ηλεκτρονικό Εμπόριο**

Οντολογίες σχεδιάζονται για να διευκολύνουν την ανταλλαγή γνώσης και την επαναχρησιμοποίησή της. Έτσι η οντολογία είναι προδιαγραφή που ορίζει δεσμεύσεις εννοιών. Με αυτή τη λογική οι οντολογίες είναι προδιαγραφές που χρησιμοποιούνται για να γίνουν οντολογικές δεσμεύσεις. Ο τελευταίος αυτός όρος αναφέρεται σε μία συμφωνία για τη χρήση ενός λεξικού, (π.χ. για ερωτήματα και απαντήσεις με ισχυρισμούς), με τρόπο τέτοιο ώστε να είναι συνεπής (αλλά σίγουρα όχι πλήρης) αναφορικά με την θεωρία που ορίζει η οντολογία. Συνήθως κατασκευάζονται agents που χρησιμοποιούν τις οντολογίες, και αντίστροφα οντολογίες που να μπορούν να χρησιμοποιηθούν από agents [12].

Μία χρήσιμη εναλλακτική εξήγηση του ρόλου των οντολογιών είναι ότι επιτρέπουν στους ανθρώπους να συγκρίνουν την ομοιότητα ή τη διαφορά των εννοιών και να δημιουργήσουν αντιστοιχίες που θα χρησιμοποιούνται για να εδραιωθούν εννοιολογικά σωστά κανάλια επικοινωνίας.

**Οι εφαρμογές της οντολογίας κατηγοριοποιούνται σε δύο ευρείες περιοχές:**

- Διαχείριση Γνώσης (**Knowledge Management**) η οποία περιλαμβάνει τη διαχείριση γνώσης και εταιρικών Intranets (Corporate Intranet and Knowledge Management), την ανάκτηση πληροφορίας (Information Retrieval) και τις δικτυακές πύλες και κοινότητες (Portal and web communities)
- Ηλεκτρονικό Εμπόριο (**e – Commerce**) όπως το B2B και το B2C)



### **2.10.1 Περιοχή εφαρμογής: Διαχείριση Γνώσης**

Η γνώση έχει αναγνωριστεί στις μέρες μας ως ο πιο σημαντικός παράγοντας επιτυχίας των επιχειρήσεων, και κατά συνέπεια η διαχείριση της γνώσης αναγνωρίζεται ως στρατηγικά σημαντικό μέσο για τις επιχειρήσεις. Η ανταγωνιστικότητα των σύγχρονων επιχειρήσεων εξαρτάται σε μεγάλο βαθμό από τον τρόπο που αυτές συντηρούν και ανακτούν την γνώση.

Στο παρελθόν, οι τεχνολογίες της πληροφορίας για τέτοιες εφαρμογές έχουν επικεντρωθεί σε πηγές πληροφοριών βασισμένων σε κείμενα. Στο μέλλον, οι τεχνολογίες του Σημασιολογικού Ιστού (Semantic Web) κυρίως οι οντολογίες και τα σχετικά μεταδεδομένα οδηγούν στον δρόμο νέων λύσεων που βασίζονται σε εννοιολογικά συνδεδεμένα τμήματα γνώσης.

Η ανεύρεση και συντήρηση πληροφορίας αποτελεί ένα δύσκολο πρόβλημα στην τελείως αδόμητη ή έστω ανεπαρκώς δομημένη αναπαράσταση των διαφόρων μέσων. Παρόλο που όλο και περισσότερες επιχειρήσεις αναγνωρίζουν την αξία της γνώσης, επίσης συνειδητοποιούν ότι η ωμή πληροφορία που συνήθως βρίσκεται και σε μεγάλες ποσότητες, δεν συνεισφέρει στην επίλυση κανενός είδους προβλήματος, δεν προσφέρει αξία ή ανταγωνιστικό πλεονέκτημα. Είναι πρακτικά άχρηστη αν δεν συνοδεύεται με αποδοτικούς τρόπους εφαρμογής, που να μπορούν να μετασχηματίσουν την πληροφορία σε χρήσιμη γνώση. Τα εργαλεία διαχείρισης γνώσης στοχεύουν στο να αποδώσουν την συνεχώς μεταβαλλόμενη γνώση μεταξύ ανθρώπων με διαφορετικές εμπειρίες και από διαφορετικά περιβάλλοντα και έτσι να εκμεταλλευτούν την υπάρχουσα πληροφορία στις επιχειρήσεις.

Σε σχέση με τις εφαρμογές ανάκτησης γνώσης, οι οντολογίες χρησιμοποιούνται για να οδηγήσουν την αναζήτηση ώστε τα αποτελέσματα να είναι αποδοτικά. Οι οντολογίες ενισχύουν την αναπαράσταση των εννοιών που αναζητούνται και κατά συνέπεια επιτρέπουν τη βελτίωση της απόδοσης των συστημάτων.

Σε σχέση με τις θεματικές πύλες στο Internet, η ιδέα του semantic web, αναφέρεται στην παράδοση δεδομένων πάνω από εννοιολογική βάση. Αυτή η προσέγγιση βοηθά στην επίτευξη καλύτερης επικοινωνίας μεταξύ του προμηθευτή και καταναλωτή της πληροφορίας.

Συγκεκριμένες λύσεις που βελτιώνουν τις παραπάνω υπηρεσίες, σχετίζονται ή βασίζονται σε οντολογίες και αποτελούν πεδίο έρευνας, έχουν να κάνουν με:

- **διαφανή ενσωμάτωση** μηχανισμού δημιουργίας γνώσης ή νέας γνώσης (Seamless Integration of Knowledge Creation)
- **ενσωμάτωση της πληροφορίας του περιβάλλοντος** (Context). Η γνώση είναι εξαρτώμενη από αυτό (context dependent)
- **Εξατομίκευση** (personalization) της γνώσης σε σχέση με το περιβάλλον.
- KM παντού και πάντα (anywhere, anytime)
- συνδυασμός KM και eLearning

## 2.10.2 Περιοχή εφαρμογής: Ηλεκτρονικό Εμπόριο

Το ηλεκτρονικό εμπόριο βρίσκεται σε μια πορεία επαναστατικών αλλαγών όπου δημιουργούνται ηλεκτρονικές αγορές οι οποίες συνεπάγονται μια νέα γενιά υπηρεσιών μεταξύ προμηθευτών και αγοραστών, όπως ηλεκτρονικές διαπραγματεύσεις, διαφάνεια στις αγορές, online διαμόρφωση προϊόντων κλπ. Το πρόβλημα που αντιμετωπίζεται όμως είναι η έλλειψη μιας κοινής γλώσσας επικοινωνίας σχετικά με τα προϊόντα και τα χαρακτηριστικά τους. Οι πελάτες χρειάζονται να προσδιορίσουν τις ανάγκες τους σε μια γλώσσα κατανοητή για αυτούς. Αυτές οι ανάγκες πρέπει να ταιριάζουν με τις υπηρεσίες που προσφέρονται από τους προμηθευτές (και που πιθανά εκφράζονται σε κάποια άλλη γλώσσα, φυσική ή γλώσσα αναπαράστασης). Για να γίνει αυτό πιο αποδοτικά για όλους, τα δεδομένα πρέπει να παρέχονται με έναν συνεπή τρόπο από όλους. Αυτόν τον τρόπο παρέχουν οι οντολογίες, επεκτείνοντας και τυποποιώντας τον τρόπο ανταλλαγής

δεδομένων στις εμπορικές συναλλαγές, ενισχύοντας έτσι τη διαφάνεια στις διάφορες διαστάσεις των χαρακτηριστικών των προϊόντων.

Συγκεκριμένες λύσεις που βελτιώνουν τις παραπάνω υπηρεσίες, σχετίζονται ή βασίζονται σε οντολογίες και αποτελούν πεδίο έρευνας, έχουν να κάνουν με:

- Την πρόσβαση σε πληροφορίες προϊόντων από διαφορετικούς προμηθευτές, και τη δυνατότητα επαρκούς σύγκρισης των χαρακτηριστικών
- Την εμπρόθεσμη και κατάλληλα οργανωμένη παραγγελία και παράδοση των προϊόντων
- Την τυποποίηση του τρόπου οργάνωσης και παρουσίασης της πληροφορίας μεταξύ όλων των προμηθευτών συναφών προϊόντων
- Την νομική και οικονομική προσαρμογή των συναλλαγών σε διαφορετικά κράτη και κουλτούρες.

Συνοψίζοντας, οι οντολογίες χρησιμοποιούνται για:

- **Κοινή χρήση της δομής της πληροφορίας ανάμεσα σε ανθρώπους και μηχανές**, ο οποίος είναι και ο πιο διαδεδομένος λόγος ανάπτυξης οντολογιών. Όταν με τη βοήθεια μιας οντολογίας αναπτύσσονται διαφορετικές σελίδες που βασίζονται σε αυτήν, τότε οι εφαρμογές μπορούν να συνδυάσουν πληροφορίες για να απαντήσουν στις αναζητήσεις του χρήστη.
- **Επαναχρησιμοποίηση της γνώσης της περιοχής**, που πάλι αποτελεί μια από τις βασικές δυνάμεις που οδηγούν την έρευνα. Για παράδειγμα, πολλά μοντέλα οντολογιών διαφορετικών περιοχών χρειάζονται την έννοια του χρόνου. Η συγκεκριμένη αναπαράσταση περιλαμβάνει τις έννοιες χρονικού διαστήματος, χρονικής στιγμής κλπ. Μια καλά ορισμένη οντολογία από ένα γκρουπ, μπορεί να ξαναχρησιμοποιηθεί σε άλλες περιοχές και από άλλες ομάδες.
- **Σαφή ορισμό των αξιωμάτων μιας γνωστικής περιοχής**, έτσι ώστε να είναι δυνατόν εύκολα να γίνουν αλλαγές εξαιτίας αλλαγών που προκύπτουν στο πεδίο

της γνώσης. Από την άλλη μεριά ο σαφής ορισμός των εννοιών βοηθά νέους χρήστες να κατανοήσουν την υπάρχουσα γνώση.

- **Διαχωρισμό της γνώσης μια περιοχής από τη λειτουργική γνώση.** Μια διαδικασία η οποία περιγράφεται για κάποιες συσκευές, μπορεί να χρησιμοποιηθεί αυτούσια για κάποιες άλλες, αν αυτές προστεθούν στην χρησιμοποιούμενη οντολογία, χωρίς όμως να χρειάζονται αλλαγές στον ορισμό της διαδικασίας.
- **Ανάλυση της γνώσης,** η οποία είναι δυνατή με χρήση όρων που είναι διαθέσιμοι και οι οποίοι πρέπει να έχουν παραχθεί από μια τυπική διαδικασία ανάλυσης

## 2.11 Το πρόβλημα και η λύση των οντολογιών

### Το πρόβλημα:

- Άνθρωποι, οργανισμοί και προγράμματα πρέπει να επικοινωνούν μεταξύ τους.
- Οι διαφορετικές ανάγκες και το διαφορετικό υπόβαθρο οδηγούν σε αποκλίνουσες οπτικές γωνίες και παραδοχές για πράγματα που στην ουσία είναι ίδια μεταξύ τους.
- Η έλλειψη κοινής αντίληψης οδηγεί σε:
  1. προβλήματα στην επικοινωνία μεταξύ ανθρώπων και οργανισμών.
  2. δυσκολίες στον προσδιορισμό των απαιτήσεων και κατά συνέπεια στην ανάπτυξη των προδιαγραφών των συστημάτων.
- Οι ανομοιόμορφες μέθοδοι μοντελοποίησης, οι γλώσσες και τα εργαλεία λογισμικού περιορίζουν σοβαρά:
  1. τη δια-λειτουργικότητα
  2. την επαναχρησιμοποίηση και το διαμοιρασμό εφαρμογών.
- Τελικά όλα τα παραπάνω οδηγούν στο να ξαναανακαλύπτουμε τον τροχό.

### η λύση:

- Η εξάλειψη ή η μείωση της σύγχυσης σχετικά με τις έννοιες και τους όρους και τελικά η απόκτηση κοινής αντίληψης.
- Αυτή η κοινή αντίληψη μπορεί να αποτελέσει το ενοποιητικό πλαίσιο ανάμεσα στις διαφορετικές οπτικές γωνίες και να συμβάλει στην βελτίωση της επικοινωνίας, της δια-λειτουργικότητας, και να οδηγήσει σε πλεονεκτήματα αναφορικά με την μηχανική των συστημάτων παρέχοντας δυνατότητες επαναχρησιμοποίησης, βελτιώνοντας την αξιοπιστία, και διευκολύνοντας την ανάπτυξη προδιαγραφών.[13]

#### **2.12.1 Μεθοδολογίες ανάπτυξης οντολογιών**

Μια βασική διαφορά σε μια τεχνολογία που βρίσκεται στην εμβρυακή της ηλικία και σε μια άλλη που έχει φτάσει στο στάδιο της ωριμότητας είναι ότι η τελευταία έχει σαφείς μεθοδολογίες ενώ η άλλη όχι ακόμα. Από την άλλη μεριά, η δημιουργία οντολογιών είναι μια πολύπλοκη, χρονοβόρα και ευαίσθητη σε σφάλματα εργασία. Για το λόγο αυτό διάφορες προσεγγίσεις διαδικασίας έχουν προταθεί από διάφορες ερευνητικές κυρίως ομάδες, το θέμα όμως ακόμα δεν έχει λυθεί με την τυποποίηση κάποιας διαδικασίας. Στην παράγραφο αυτό επιχειρείται με μια σύντομη επισκόπηση διαφόρων μεθοδολογιών.[8]

#### **2.12.1 SEAL (Semantic Portal)**

Το SEAL είναι ένα ευρύτερο πλαίσιο εργασίας (framework), που δημιουργήθηκε για να λύσει προβλήματα που σχετίζονται με εννοιολογικές αναζητήσεις (semantic search) γνώσης στο web και συγκεκριμένα στην ανάπτυξη εννοιολογικών web portals. Βασίζεται σε οντολογίες, και τις χρησιμοποιεί για να ορίσει και να δομήσει όρους, υποστηρίζοντας

έτσι, εννοιολογικές ερωτήσεις (semantic querying) ή δυναμική παρουσίαση υπερκειμένων (dynamic hypertext views). Η μεθοδολογία SEAL αποτελεί μια γενική μεθοδολογία για την ανάπτυξη εννοιολογικών πυλών (semantic portals). Είναι μια πρωτοβουλία υλοποίησης της βασικής ιδέας του όρου «Semantic Web», αυτής της διανομής δεδομένων/ γνώσης πάνω σε εννοιολογική βάση, αναμένοντας αποτελέσματα που θα βελτιώνουν την επικοινωνία μεταξύ του προμηθευτή και του καταναλωτή της πληροφορίας.

Η επικοινωνία μεταξύ του προμηθευτή και του καταναλωτή της πληροφορίας περιγράφεται από την τριάδα των όρων: «σύμβολο», «έννοια», «αντικείμενο» του φυσικού κόσμου. Η σχέση μεταξύ μιας λέξης και ενός αντικειμένου είναι έμμεση, και μπορεί μόνο να επιτευχθεί μέσω της «έννοιας».

Η οντολογία αποτελεί μια γενική λογική θεωρία που αποτελείται από όρους (vocabulary) και ένα σύνολο δηλώσεων για μια περιοχή της γνώσης. Η θεωρία αυτή ορίζει σχέσεις και έτσι περιορίζει το σύνολο των πιθανών ερμηνειών των συμβόλων σε αντικείμενα.

Η μεθοδολογία αυτή έχει μελετηθεί ως προς πιθανές επεκτάσεις σε διάφορες κατευθύνσεις:

- **Εννοιολογική Ταξινόμηση-Ομοιότητα-Συνωνυμία (Semantic Ranking-similarity):** Η εφαρμογή εξετάζει τρόπους εύρεσης της προτιμότερης λύσης από ένα σετ λύσεων που ικανοποιούν μια συγκεκριμένη αναζήτηση.
- **Εννοιολογική εξατομίκευση (Semantic personalization):** Ανάλυση και παρακολούθηση της συμπεριφοράς του χρήστη σε μια αναζήτηση, σε εννοιολογικό επίπεδο (semantic bookmarks, semantic logfiles).

## **2.12.2. IDEF5 (ICAM Integrated Computer-Aided Manufacturing] DEFinition)**

Η IDEF5 αποτελεί μια μεθοδολογία για την δημιουργία οντολογιών και παρέχει μία γλώσσα για την αναπαράσταση των εννοιών. Έχει σχεδιαστεί ειδικά για να εξυπηρετήσει τη διαδικασία συλλογής γνώσης για φυσικά ή εννοιολογικά αντικείμενα μιας περιοχής γνώσης και να καταγράψει τις σχέσεις τους. Η πρωτοβουλία ανάπτυξης της μεθόδου, προέκυψε από την ανάγκη δημιουργίας μιας πρακτικής μεθόδου για την δημιουργία οντολογιών. Το θέμα έχει μελετηθεί σε ακαδημαϊκό επίπεδο αλλά δεν υπήρξε μια πρακτική μέθοδος που να υποδεικνύει τον μηχανισμό απόκτησης, αποθήκευσης, και συντήρησης οντολογιών με επαναχρησιμοποιούμενη και επεκτάσιμη προοπτική.

Η μεθοδολογία IDEF5 χρησιμοποιείται κυρίως για την αναπαράσταση Επιχειρησιακής Γνώσης και πιο συγκεκριμένα για την αναπαράσταση επιχειρησιακών διαδικασιών (Business Processes) και θεωρεί ότι:

- Οντολογία είναι (α) ένας κατάλογος από όρους οι οποίοι χρησιμοποιούνται σε μια γνωστική περιοχή, (β) κανόνες/περιορισμοί που προσδιορίζουν τους τρόπους που αυτοί οι όροι συνδυάζονται για να περιγράψουν καταστάσεις του χώρου, (γ) το μοντέλο- κανόνων που μπορεί να συνεπάγεται επιπλέον κανόνες (inference).
- Σχέση (relation) περιγράφει σχέση (association) στον πραγματικό κόσμο
- Όρος χρησιμοποιείται για να περιγράψει ένα αντικείμενο ή μια κατάσταση στον πραγματικό κόσμο
- Το μοντέλο κανόνων περιλαμβάνει εκείνους τους κανόνες που είναι χρήσιμοι, στο συγκεκριμένο περιβάλλον (context)
- Είδος (Kind) είναι μια κατηγορία αντικειμένων που συνδέονται από μια κοινή φύση, ένα σύνολο κοινών ιδιοτήτων (properties) που χαρακτηρίζουν όλα τα αντικείμενα.

### 2.12.3 TOVE (Toronto Virtual Enterprise)

Η μεθοδολογία αναπτύχθηκε για την μοντελοποίηση δραστηριοτήτων σε επιχειρήσεις. Η μεθοδολογία περιγράφει τη δόμηση ενός μοντέλου γνώσης που πρέπει να προσδιοριστεί από την οντολογία. Το μοντέλο χτίζεται μέσα από μια διαδικασία που αρχικά αποσκοπεί σε μια άτυπη περιγραφή των προδιαγραφών που πρέπει να εξυπηρετηθούν και στη συνέχεια οδηγείται στην τυπική περιγραφή. Χαρακτηριστικό της οντολογίας είναι η υποστήριξη της διαδικασίας αξιολόγησης, που έχει τη φιλοσοφία ορισμού ερωτημάτων τα οποία πρέπει να απαντηθούν με την οντολογία. Έτσι αυτά τα ερωτήματα βοηθούν στην αξιολόγηση της αποδοτικότητας της οντολογίας.

Το έργο TOVE έχει δημιουργήσει μια επαναχρησιμοποιήσιμη οντολογία των βιομηχανικών εννοιών. Οι οντολογίες του TOVE δημιουργήθηκαν για να περιγράψουν οντολογίες σε επιχειρησιακές διαδικασίες και μοντελοποίηση δραστηριοτήτων. Στο έργο TOVE αναπτύχθηκαν μεθοδολογίες για δραστηριότητες, καταστάσεις, οργάνωση, πόρους, προϊόντα, υπηρεσίες, παραγωγή, κόστος, και ποιότητα. Οι οντολογίες αυτές συνιστούν ένα ολοκληρωμένο μοντέλο επιχείρησης και υποστηρίζουν πολύπλοκους συλλογισμούς σε προβλήματα που απαιτούν αλληλεπίδραση μεταξύ των διαφορετικών οντολογιών. το έργο TOVE θεωρεί ότι:

- Οι έννοιες ομαδοποιούνται σε θεματικές κατηγορίες.
- Η οντολογία περιλαμβάνει γενικές έννοιες όπως «χρόνος» (time), «σχέση αιτίου-αιτιατού» (causality), «ενέργεια» (activity) και «περιορισμός» (constraint).
- Οι έννοιες δομούνται σε ταξονομίες, και αναπαρίστανται από σταθερές και μεταβλητές.



## **2.12.4 BPD/D (Business Process and Data Driven) Ontological Engineering**

Η μεθοδολογία BPD/D βασίζεται στη μεθοδολογία TOVE και προσαρμόζει τα βήματα της για να δημιουργήσει κανόνες για την κατασκευή όχι μόνο του περιεχομένου της οντολογίας αλλά και του πλαισίου λειτουργίας της οντολογίας στην εφαρμογή. Επίσης, υποστηρίζει τη συλλογή στοιχείων που βοηθούν στην λήψη αποφάσεων σχετικά με το σχεδιασμό και την αρχιτεκτονική της εφαρμογής. Τέλος, υποστηρίζει τη δημιουργία ομάδων (clusters) οντολογιών συντελώντας στην επαναχρησιμοποίηση τους. Η ομάδα ανάπτυξης της μεθοδολογίας αυτής αναγνωρίζει δύο τύπους προσεγγίσεων για την ανάπτυξη οντολογιών που υποστηρίζουν τη λήψη αποφάσεων: στη μία, οι οντολογίες δημιουργούνται για να υποστηρίξουν καινούργιες διαδικασίες στην επιχείρηση, αλλά βασίζονται στα υπάρχοντα δεδομένα, και στην άλλη, οι οντολογίες δημιουργούνται με βάση τα υπάρχοντα δεδομένα αλλά συχνά δεν υποστηρίζουν νέες διαδικασίες και αποφάσεις. Αυτή η μεθοδολογία που προτείνεται συνδυάζει και τις δύο προσεγγίσεις.

Οι οντολογίες δημιουργούνται με προσεγγίσεις ανάλυσης επιχειρησιακών διαδικασιών αλλά και με προσεγγίσεις ανάλυσης δεδομένων. Χρησιμοποιούνται για εφαρμογές που σχετίζονται με τις επιχειρησιακές διαδικασίες ή για εφαρμογές των οποίων τα δεδομένα και πληροφοριακά συστήματα συνιστούν ανάπτυξη οντολογιών ενώ χρειάζονται να ανταλλάσσουν ή να μοιράζονται οντολογία για την καλύτερη λειτουργία τους.

## **2.12.5 METHONTOLOGY**

Η μεθοδολογία METHONTOLOGY αυτή παρουσιάζει τη διαδικασία δημιουργίας οντολογιών, είτε από το μηδέν, είτε επαναχρησιμοποιώντας υπάρχουσες οντολογίες όπως είναι, ή μετασχηματίζοντας κάποιες υπάρχουσες. Η διαδικασία που προτείνει, εξετάζει και υποστηρίζει την ανάπτυξη οντολογίας με όλες τις σχετιζόμενες δραστηριότητες, όπως Διαχείριση έργου, Ανάπτυξη, Υποστήριξη, αλλά και δραστηριότητες συντήρησης στη διάρκεια ζωής της οντολογίας. Η ομάδα ανάπτυξης, έχοντας μελετήσει τις διάφορες προσπάθειες που έχουν γίνει για τη δημιουργία

οντολογιών, αναγνωρίζει την έλλειψη μιας τυπικής μεθόδου που θα μπορούσε να μετατρέψει τη διαδικασία ανάπτυξης μεθοδολογίας από τέχνη σε επιστήμη, που να περιλαμβάνει τυπικές δραστηριότητες, τυπική πορεία ζωής μεθοδικές μεθοδολογίες, αλλά επίσης και ένα σύνολο εργαλείων και τεχνικών. Η προσπάθεια της ομάδας ανάπτυξης της METHODOLOGY, στοχεύει ακριβώς εδώ: στο να μετατρέψει την διαδικασία από τέχνη σε επιστήμη μηχανικής, δίνοντας έναν ορισμό και μια τυπική διαδικασία για όλο τον κύκλο ζωής της οντολογίας, από την συλλογή των απαιτήσεων ως τη συντήρηση του ολοκληρωμένου προϊόντος, καθώς και μεθοδολογιών και εργαλείων που οδηγούν στην ανάπτυξή τους.

Οι οντολογίες, σε αυτήν την μεθοδολογία, κατασκευάζονται είτε για να μπορούν να επαναχρησιμοποιηθούν, είτε ως κάποιο κοινό εργαλείο ανεξάρτητα από χρόνο, χώρο και τακτική ή περιοχή εφαρμογής. Η μεθοδολογία βασίζεται στη διαχείριση του κύκλου ζωής των εξελισσόμενων πρωτότυπων.

### **2.12.6 CEO (Core Enterprise Ontology)**

Η μεθοδολογία CEO αποτελεί μια αρχική προσπάθεια μοντελοποίησης των εννοιών που σχετίζονται με τις επιχειρησιακές πρακτικές. Η προσέγγιση που υιοθετείται, στοχεύει στο να δημιουργήσει μια οντολογία η οποία γενικεύει τις βασικές έννοιες που αντιπροσωπεύουν τις επιχειρηματικές πρακτικές στην μεγάλη τους πλειοψηφία, και ανεξάρτητα από βιομηχανικό τομέα. Για παράδειγμα υιοθετεί την έννοια του «πελάτη», ως κάποιος που λαμβάνει ένα αγαθό ή υπηρεσία αντί κάποιου είδους πληρωμής. Αυτή η έννοια στη συνέχεια προσαρμόζεται σε κάτι διαφορετικό, ανάλογα με τον βιομηχανικό τομέα εφαρμογής, και έτσι μπορεί να είναι ένας τουρίστας ή επιβάτης ή οδηγός. Η μεθοδολογία αυτή αναφέρεται στο πρόβλημα που αντιμετωπίζεται στον γρήγορα εξελισσόμενο σύγχρονο κόσμο, όπου νέες επιχειρήσεις εμφανίζονται και νέου τύπου συνεργασίες δημιουργούνται. Σ' αυτό το πλαίσιο, ένα από τα βασικά προβλήματα προκύπτει από τη δυσκολία δημιουργίας συνεργασίας μεταξύ διαφορετικών

επιχειρήσεων που χρειάζονται να μοιραστούν μια κοινή γνώση μιας περιοχής ενώ η δημιουργία κοινής οντολογίας θεωρείται ως μια υποσχόμενη λύση.

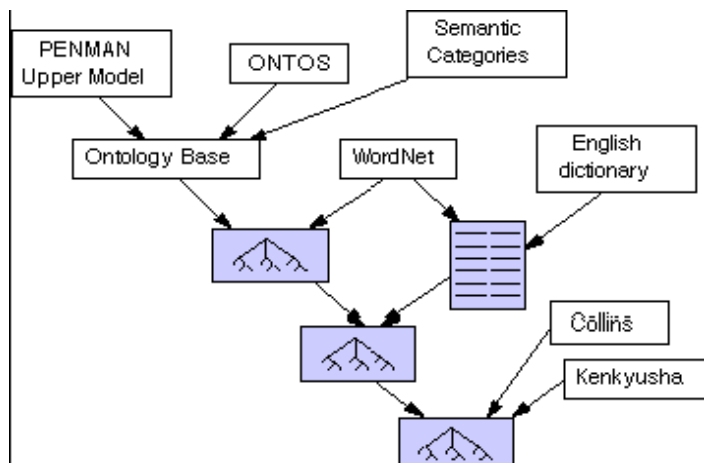
Η μεθοδολογία υποστηρίζει την κατασκευή επιχειρησιακών οντολογιών. Τέτοιες οντολογίες μπορεί να διαφέρουν από βιομηχανία σε βιομηχανία, αλλά είναι πολύ χρήσιμες για εφαρμογές Management Information Systems (MIS), Business Process Reengineering (BPR), Virtual Enterprise (VE), όπου η μοντελοποίηση και κατανόηση μιας επιχείρησης και του τομέα δραστηριότητάς της είναι απαραίτητα.

Η μεθοδολογία αποδέχεται ότι από πλευράς τεχνολογίας η οντολογία μπορεί να θεωρηθεί σαν μια βάση δεδομένων εννοιών. Οι έννοιες είναι πολύπλοκες δομές, πολύ στενά αλληλοσυνδεδεμένες. Για τη μοντελοποίηση της γνώσης συνηθίζεται τριεπίπεδη οργανωτική δομή. Στο κατώτερο επίπεδο, το ονομαζόμενο extensional που υλοποιεί ένα μοντέλο, αποθηκεύονται πληροφορίες για τα αντικείμενα. Η αντιστοίχιση στην τεχνολογία βάσεων δεδομένων είναι η ίδια η βάση δεδομένων. Το μεσαίο επίπεδο (intentional) αφορά το εννοιολογικό μοντέλο που επιτρέπει το χτίσιμο οντολογίας. Η αντιστοιχία με τις βάσεις δεδομένων είναι το σχήμα της βάσης. Το ανώτερο επίπεδο (metalevel) περιλαμβάνει τις ιδέες μοντελοποίησης ή δομές που υλοποιούνται στην μοντελοποίηση εννοιών του μεσαίου επιπέδου. Πρακτικά, το περιεχόμενο του πάνω επιπέδου αντιστοιχεί στον ορισμό του του μεσαίου, το οποίο θα αναπαραστήσει το εννοιολογικό μοντέλο μιας περιοχής γνώσης.

### **2.12.7 SENSUS**

Η SENSUS είναι μια οντολογία που περιλαμβάνει περίπου 50.000 όρους και κατασκευάστηκε εξάγοντας και συνθέτοντας πληροφορίες από υπάρχουσες πηγές, όπως το WORDNET, λεξικά, και οντολογίες. Η ομάδα ανάπτυξής της χρησιμοποιεί αυτή την γενική οντολογία και με ένα ημιαυτόματο τρόπο που επιτρέπει να αναγνωρίσει κανείς όρους στη γενική οντολογία, οι οποίοι σχετίζονται με κάποια περιοχή γνώσης. Έτσι μπορούν περικόπτοντας τους περιττούς όρους, έτσι ώστε να εξυπηρετεί τις ανάγκες της περιοχής γνώσης. Επιπρόσθετα, είναι δυνατόν να εμπλουτίσουν την

παραγόμενη οντολογία με νέους όρους που απαιτούνται στη εφαρμογή. Η προσέγγιση αυτή προέκυψε από την παρατήρηση ότι πολλές υπάρχουσες οντολογίες είχαν τα απαραίτητα χαρακτηριστικά για να γενικευτούν, ωστόσο καμιά δεν τα είχε όλα. Οι PENMAN και ONTOS (υπάρχουσες οντολογικές βάσεις) παράγουν υψηλού επιπέδου οργάνωση, αλλά προσφέρουν περιορισμένο αριθμό εννοιών και έτσι δεν μπορούν να δώσουν την απαραίτητη ευρεία κάλυψη. Από την άλλη μεριά το WORDNET, αποτελεί ένα εννοιολογικό δίκτυο σε μορφή θησαυρού, έχει γενική κάλυψη αλλά δεν έχει την απαραίτητη υψηλού επιπέδου οργάνωση. Διάφορα λεξικά επίσης έχουν πολλούς όρους, αρκετά ευρύ πεδίο κάλυψης, αλλά δεν αποτελούν εννοιολογικά δίκτυα.



στρατηγική συγχώνευσης *sensus*

Με βάση την περίληψη η μεθοδολογία παραδέχεται πως μια γενική οντολογία πρέπει:

- να καλύπτει ένα πολύ ευρύ τομέα γνώσης
- να περιέχει μεγάλο αριθμό όρων
- να περιέχει κάποια μορφή οργάνωση υψηλού επιπέδου, που να αποτελεί τον οδηγό για την ελάττωση παρερμηνειών

### **2.12.8 DECOR Methodology**

Η μεθοδολογία DECOR αυτή δεν αποτελεί αποκλειστικά μια μεθοδολογία ανάπτυξης οντολογίας. Είναι μια πλήρης προσέγγιση για την διαχείριση γνώσης που βασίζεται σε επιχειρηματικές διαδικασίες (Business Process Oriented Knowledge Management, BPOKM). Η περιοχή αυτή έρευνας βασίζεται στο γεγονός ότι η επιχειρηματική διαδικασία αποτελεί για ένα οργανισμό σημαντική γνώση. Η μεθοδολογία συνδυάζει μεθόδους από την περιοχή της μοντελοποίησης επιχειρηματικών διαδικασιών και διαχείρισης γνώσης για να δημιουργήσει έναν οδηγό εκπόνησης έργων BPOKM.

Η μεθοδολογία προφανώς σχεδιάστηκε για εφαρμογές διαχείρισης της γνώσης μέσα στις επιχειρηματικές διαδικασίες οργανισμών. Πιλοτικά εφαρμόστηκε για τη μοντελοποίηση της διαδικασίας απονομής πλήρους σύνταξης γήρατος, για το ΙΚΑ. Η διαδικασία που επιλέχθηκε περιλαμβάνει κάποια στάδια λήψης απόφασης που στηρίζονται σε έγγραφα και γνώση, ενώ παράλληλα κάθε στάδιο πρέπει να είναι συμβατό από την υπάρχουσα νομοθεσία.

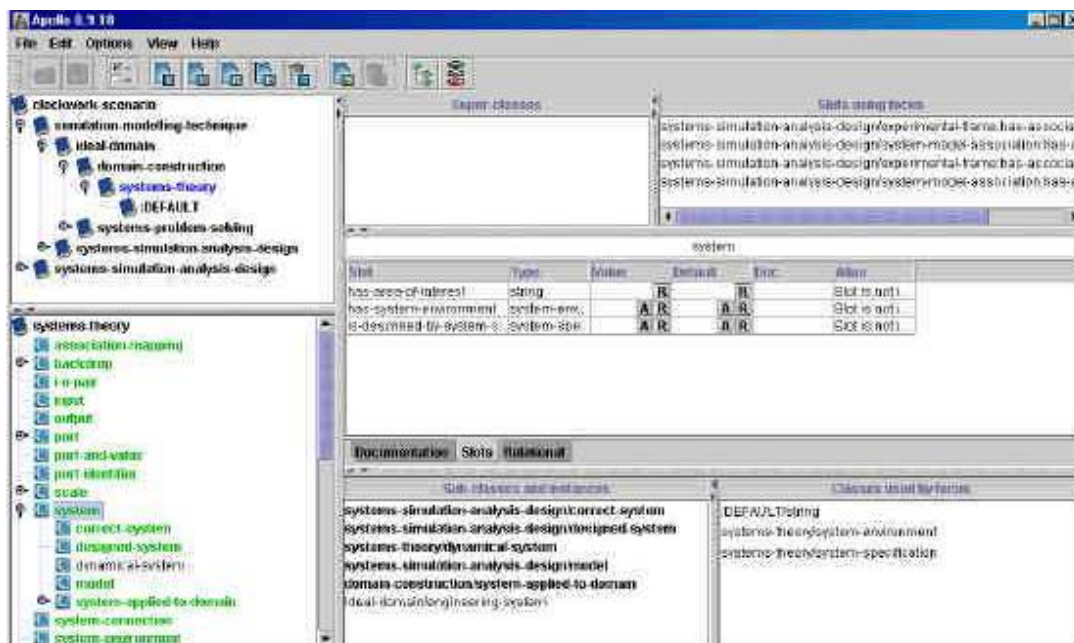
## 2.13 Εργαλεία ανάπτυξης οντολογιών`

Τα τελευταία χρόνια έχουν αναπτυχθεί και αναπτύσσονται αρκετά εργαλεία που παρέχουν ένα πλαίσιο εργασίας για την κατασκευή οντολογιών. Τα εργαλεία αυτά έχουν ως σκοπό να κάνουν τη διαδικασία ανάπτυξης μιας οντολογίας ευκολότερη και είναι γνωστά ως *συστήματα διαχείρισης οντολογιών* (ΣΔΟ)[[39]. Τα ΣΔΟ έχουν αντίστοιχη λειτουργικότητα με τα ολοκληρωμένα περιβάλλοντα ανάπτυξης (integrated development environments - IDE) που χρησιμοποιούνται συνήθως για την ανάπτυξη λογισμικού. Τα ΣΔΟ επιτρέπουν στους σχεδιαστές οντολογιών να κατασκευάζουν οντολογίες μέσω ενός γραφικού περιβάλλοντος (graphical user interface - GUI). Επιπλέον κάποια από αυτά τα εργαλεία παρέχουν τη δυνατότητα διασύνδεσης εννοιών της υπό κατασκευή οντολογίας με εξωτερικές πηγές δεδομένων, όπως βάσεις δεδομένων και υπηρεσίες web. Αφού η κατασκευή της οντολογίας ολοκληρωθεί, αυτή μπορεί να πλαισιωθεί από μια μηχανή εξαγωγής συμπερασμάτων (inference engine). Η μηχανή εξαγωγής συμπερασμάτων παρέχει την νοημοσύνη που απαιτείται για τη λογική ανάλυση της οντολογίας και την παραγωγή νέων δεδομένων και σχέσεων. Οι διαθέσιμες μηχανές εξαγωγής συμπερασμάτων ποικίλουν σε βαθμό πολυπλοκότητας και λειτουργικότητας. Τέλος, μετά το πέρας της ανάπτυξης τη οντολογίας, μια μηχανή υποβολής ερωτημάτων που μπορεί είτε να περιέχεται στην οντολογία είτε να είναι εξωτερική και να διασυνδέεται με την οντολογία, χρησιμεύει για την υποβολή σημασιολογικά δομημένων ερωτήσεων στη μηχανή εξαγωγής συμπερασμάτων και την εξαγωγή των αντίστοιχων δεδομένων. Συνήθως τα ΣΔΟ παρέχουν τη δυνατότητα υποβολής ερωτημάτων στη μηχανή εξαγωγής συμπερασμάτων μέσα από το περιβάλλον τους.

Μερικά από τα πιο γνωστά εργαλεία ανάπτυξης και ΣΔΟ που μπορούν να χρησιμοποιηθούν είτε για την ανάπτυξη οντολογιών από μηδενικής βάσης, είτε για την επεξεργασία και επαναχρησιμοποίηση οντολογιών κατασκευασεμένων από τρίτους, είναι τα παρακάτω:

### 2.13.1. Apollo

Το Apollo (<http://apollo.open.ac.uk>)[45] είναι μια φιλική προς το χρήστη εφαρμογή ανάπτυξης οντολογιών. Σχεδιάστηκε με στόχο την κάλυψη των αναγκών μηχανικών που επιθυμούσαν να χρησιμοποιήσουν τεχνικές μοντελοποίησης γνώσης, αλλά παράλληλα απαιτούσαν ευκολία στη χρήση, εύκολη σύνταξη και κατανοητό περιβάλλον ανάπτυξης. Ένα στιγμιότυπο χρήσης του Apollo φαίνεται στην εικόνα 2.13.1. Μια ιεραρχική αναπαράσταση των διαθέσιμων οντολογιών φαίνεται στο αριστερό πλαίσιο. Η ιεραρχία των κλάσεων απεικονίζεται στο κάτω δεξιά πλαίσιο, ενώ όταν ο χρήστης επιλέξει κάποια



εικόνα:2.13.1

κλάση ή στιγμιότυπο αυτό αναπτύσσεται λεπτομερώς στα πλαίσια που βρίσκονται στη δεξιά περιοχή της οθόνης. Οι ιδιότητες και οι τιμές του επιλεγμένου αντικειμένου εισάγονται από τον χρήστη με τρόπο παρόμοιο με την συμπλήρωση ενός λογιστικού φύλλου (spreadsheet). Το Apollo υποστηρίζει όλες τις βασικές δομές μοντελοποίησης γνώσης : οντολογίες, κλάσεις, στιγμιότυπα, συναρτήσεις και σχέσεις. Το σύστημα πραγματοποιεί αυτόματα ελέγχους συνέπειας (όπως ο εντοπισμός σημείων χρήσης

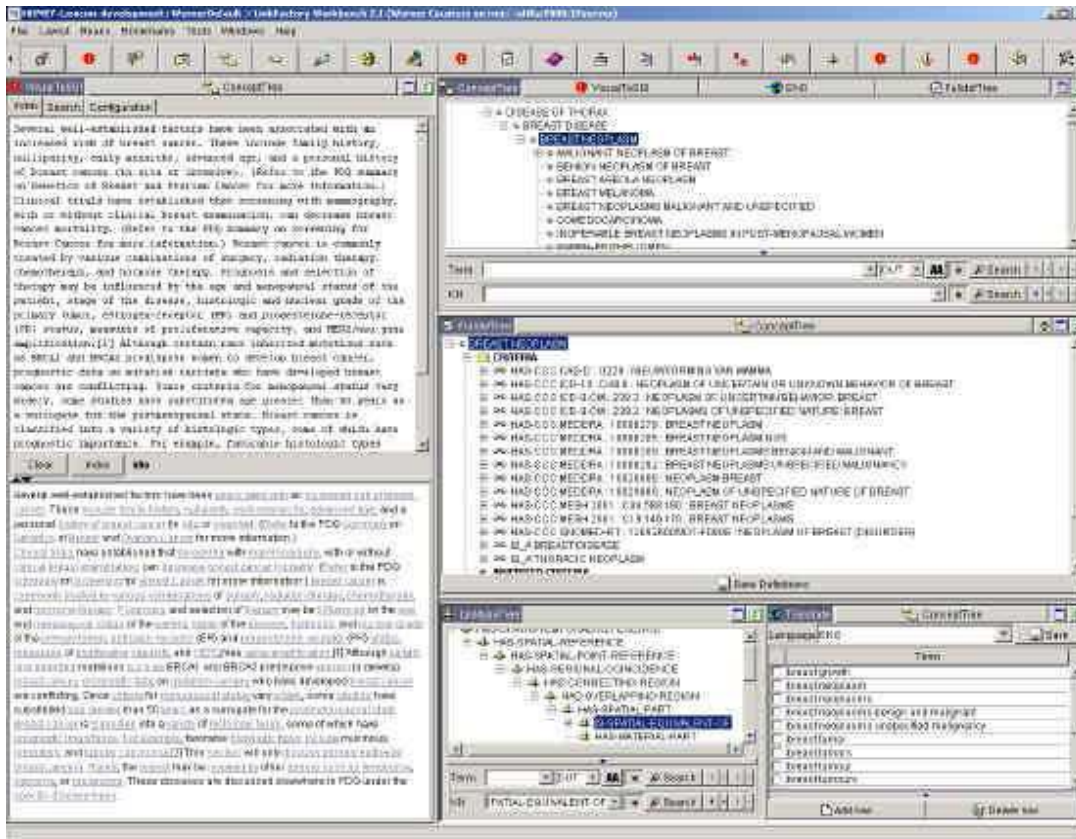
κλάσεων που δεν έχουν προηγουμένως οριστεί) καθώς ο χρήστης επεξεργάζεται τις οντολογίες.

Το Apollo είναι υλοποιημένο σε Java και έχει την δικιά του γλώσσα για την εσωτερική αναπαράσταση και □ ύλαξητων οντολογιών, ενώ μπορεί να εξάγει τις οντολογίες σε διάφορες γλώσσες αναπαράστασης.



## 2.13.2. LinKFactory


Το LinKFactory (<http://www.landcglobal.com/pages/linkfactory.php>)[46] είναι ένα ΣΔΟ που έχει αναπτύξει η εταιρία Language & Computing και είναι σχεδιασμένο για την κατασκευή και διαχείριση μεγάλων και πολύπλοκων οντολογιών. Το σύστημα αποτελείται από δύο βασικές υποεφαρμογές που είναι υλοποιημένες σε Java: το LinKFactory Server και το LinKFactory Workbench (εφαρμογή-πελάτης). Στη μεριά του server, το LinKFactory αποθηκεύει τα δεδομένα σε μια σχεσιακή βάση δεδομένων. Η πρόσβαση στη βάση δεδομένων γίνεται μέσω μιας συλλογής  $\square$  ιλικών προς το χρήστη



εικόνα 2.13.2: Στιγμιότυπο οθόνης του LinkFactory


συλλογής  $\square$  ιλικών προς το χρήστη συναρτήσεων όπως: get-children, findpath, join concepts, get terms for concept X, κλπ. Οι συναρτήσεις αυτές είναι προσβάσιμες από προγράμματα-πελάτες μέσω ενός προτυποποιημένου API που επιτρέπει την ανάπτυξη

εφαρμογών πάνω από τη σημασιολογική βάση δεδομένων χωρίς να απαιτείται γνώση της εσωτερικής δομής της βάσης. Το τμήμα αυτό έχει την ικανότητα να εξυπηρετεί πολλούς ταυτόχρονους χρήστες και είναι ανεξάρτητο υπολογιστικής πλατφόρμας (έχει δοκιμαστεί σε περιβάλλοντα MS Windows, Sun Solaris, UNIX και GNU/Linux). Η εφαρμογή απαιτεί την λειτουργία ενός ειδικού μητρώου (registry) προκειμένου να καταχωρείται σε αυτό και να μπορεί να εντοπιστεί και να προσπελαστεί από τα προγράμματα-πελάτες.

Το LinkFactory Workbench δίνει τη δυνατότητα στο χρήστη να δημιουργεί, να αναζητά και να διασυνδέει οντολογίες όπως  αίνεταστην εικόνα.

Το Workbench είναι ένα δυναμικό πλαίσιο εργασίας που αποτελείται από υπομονάδες υλοποιημένες σε Java. Κάθε υπομονάδα έχει ξεχωριστή και συγκεκριμένη λειτουργικότητα και περιορισμένη πρόσβαση στα μέρη της οντολογίας, αλλά ο συνδυασμός τους μπορεί να προσφέρει στο χρήστη ένα πολυδύναμο εργαλείο για την επισκόπηση και τον χειρισμό των δεδομένων.

Παραδείγματα τέτοιων υπομονάδων είναι τα εξής : Concept tree, Concept criteria and full definitions, Linktype tree, Criteria list, Term list, Search pane, Properties panel, Reverse relations, και πολλά άλλα. Ο χρήστης μπορεί να δημιουργήσει πολλαπλές απεικονήσεις της οντολογίας χρησιμοποιώντας τις διαθέσιμες υπομονάδες. Οι απεικονίσεις αυτές ονομάζονται "Layouts". Κάθε layout μπορεί να αποτελείται από ένα ή περισσότερα πλαίσια στα οποία τοποθετούνται οι υπομονάδες. Η δημιουργία ενός νέου layout ή η προσθήκη περισσότερων πλαισίων σε ένα υπάρχον γίνονται απλά με την επιλογή των αντίστοιχων επιλογών από ένα μενού. Κάθε πλαίσιο μπορεί να χωριστεί σε υποδοχείς, όπου τοποθετούνται οι υπομονάδες. Ο χρήστης μπορεί να επιλέξει την υπομονάδα που επιθυμεί να χρησιμοποιήσει και να την τοποθετήσει με τη χρήση του ποντικιού στη θέση που θέλει. Στη συνέχεια μπορεί να διασυνδέσει τις υπομονάδες μεταξύ τους πάλι με απλούς χειρισμούς του ποντικιού.

Κάθε υπομονάδα έχει τις δικές της ιδιότητες και ο χρήστης μπορεί να αλλάξει τις τιμές τους κατά την εκτέλεση του προγράμματος. Με αυτό τον τρόπο ο χρήστης μπορεί να δημιουργεί κάθε  οράο layout που είναι πιο κατάλληλο για τον τύπο της εργασίας που επιτελεί.

Πέραν από τη χρήση και τον συνδυασμό τους μέσα στο περιβάλλον του LinKFactory Workbench, οι υπομονάδες αυτές μπορούν να χρησιμοποιηθούν ανεξάρτητα και έξω από αυτό, έτσι ώστε να είναι δυνατή η χρήση και ενσωμάτωσή τους σε προγράμματα τρίτων.

Για τον σχεδιασμό των οντολογιών στο σύστημα είναι ενσωματωμένοι διάφοροι μηχανισμοί διασφάλισης ποιότητας, όπως: η δυνατότητα διαχείρισης πολλαπλών εκδόσεων, η καταγραφή των ενεργειών του χρήστη, ιεραρχίες χρηστών, πολιτικές δικαιωμάτων, ιεραρχία συνδέσμων κλπ.

Όσον αφορά την αναπαράσταση της γνώσης και την δυνατότητα εξαγωγής συμπερασμάτων, το LinKFactory έχει τα παρακάτω χαρακτηριστικά και δυνατότητες :

- ενσωματωμένες σχέσεις IS-A, DISJOINT και SAME-AS,
- δυνατότητα ορισμού της ιεραρχίας των σχέσεων από το χρήστη,
- μεθόδους ελέγχου περιορισμών,
- αυτόματη κατηγοριοποίηση νέων εννοιών βασιζόμενη τόσο σε  $\square$  υσικήγλώσσα όσο και σε τυπικούς ορισμούς,
- μηχανισμούς για σύνδεση και συγχώνευση οντολογιών,
- δυνατότητα αυτόματης ανάλυσης κειμένου για την δημιουργία της οντολογίας.

### 2.13.3 OilEd

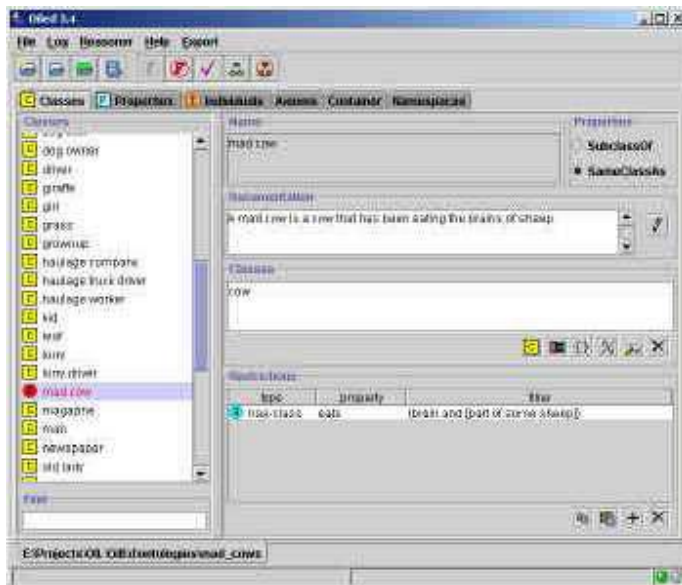
Ο OilEd (<http://oiled.man.ac.uk/>)[47] είναι ένας επεξεργαστής οντολογιών με γραφικό περιβάλλον. Αναπτύσσεται από το πανεπιστήμιο του Manchester και παρέχει στο χρήστη τη δυνατότητα δημιουργίας οντολογιών χρησιμοποιώντας τη γλώσσα DAML+OIL.

Το γνωστικό μοντέλο του OilEd βασίζεται σε αυτό της DAML+OIL και επεκτείνεται με τη χρήση παρουσίασης όμοιας με αυτή των πλαισίων. Έτσι το OilEd προσφέρει ένα οικείο περιβάλλον με πλαίσια, ενώ υποστηρίζει και την πλούσια εκφραστικότητα της DAML+OIL. Οι κλάσεις ορίζονται με βάση τις υπερκλάσεις τους και τις ιδιότητές τους, ενώ άλλες σχέσεις (όπως η disjointness) εκφράζονται μέσω συναρτήσεων. Το εκφραστικό γνωστικό μοντέλο του OilEd επιτρέπει τη χρήση πολύπλοκων και σύνθετων

παρουσιάσεων ως μοντέλα, σε αντίθεση με πολλούς επεξεργαστές που βασίζονται σε πλαίσια, όπου τα ανώνυμα πλαίσια πρέπει να ονοματίζονται για να μπορούν να χρησιμοποιηθούν ως μοντέλα.

Η βασική λειτουργία του OilEd είναι η επεξεργασία οντολογιών ή schemas, και όχι η εξόρυξη γνώσης ή η κατασκευή μεγάλων βάσεων γνώσης που περιέχουν στιγμιότυπα. Αν και το OilEd παρέχει τη δυνατότητα ορισμού ατόμων, η λειτουργία αυτή προορίζεται μόνο για τον ορισμό των ουσιαστικών που μετέχουν στη σχέση one-of της DAML+OIL.

Ένα σημαντικό χαρακτηριστικό του OilEd είναι η χρήση της μηχανής λογισμού FaCT για την κατηγοριοποίηση των οντολογιών και τον έλεγχο συνέπειας.



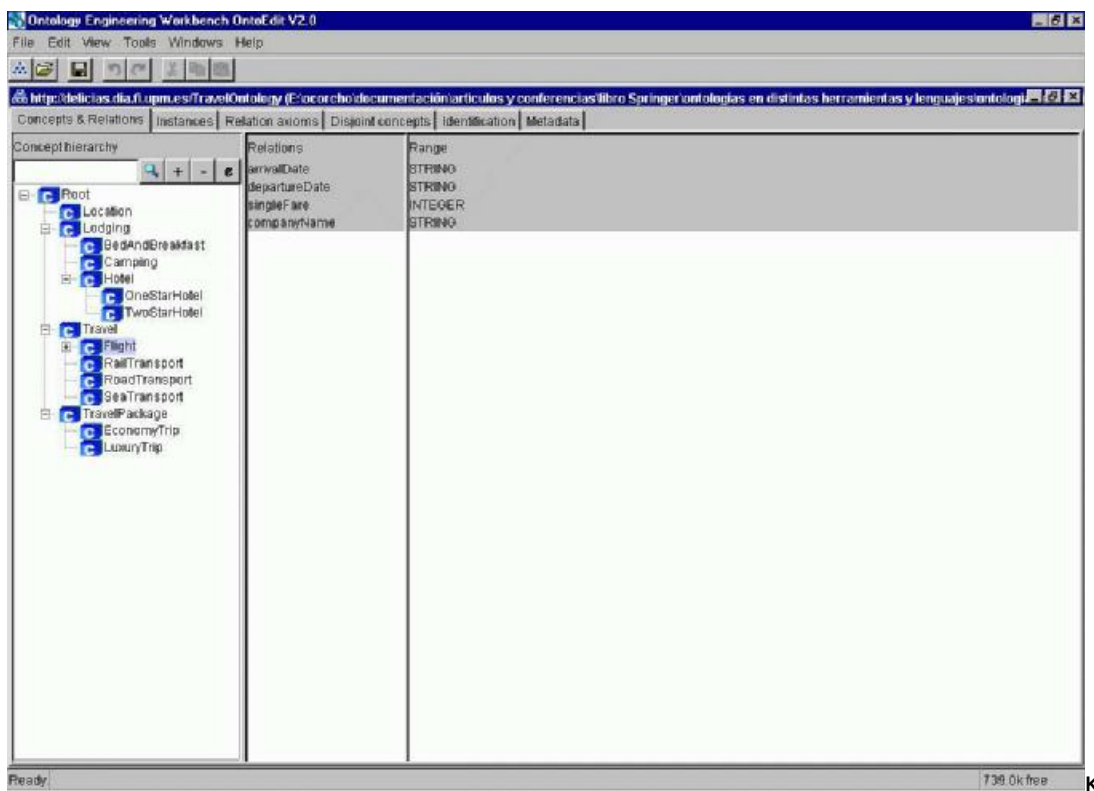
εικόνα 2.13.3: Στιγμιότυπο οθόνης του OilEd

Για το σκοπό αυτό η οντολογία μεταφράζεται εσωτερικά από την DAML+OIL στην γλώσσα αναπαράστασης λογικής SHIQ. Αυτό επιτρέπει το χρήστη να ορίσει μια κλάση και να αφήσει την μηχανή λογισμού να του υποδείξει την κατάλληλη θέση στην ιεραρχία για την τοποθέτηση του ορισμού.

Η εικόνα 2.13.3 δείχνει ένα στιγμιότυπο χρήσης του OilEd. Το DAML+OIL Schema (<http://www.daml.org/2001/03/daml+oil.daml>)[48] χρησιμοποιείται για την  όρτωση και αποθήκευση των οντολογιών. Το εργαλείο έχει τη δυνατότητα να διαβάσει και να εξάγει

ιεραρχίες εννοιών σε RDF και να παρουσιάζει οντολογίες σε μορφή HTML και SHIQ. Η ιεραρχική δομή των εννοιών μπορεί να εξαχθεί και σε μορφή κατάλληλη για επεξεργασία με το πρόγραμμα οπτικοποίησης graphviz (<http://www.research.att.com/areas/visualization/projects/software/graphviz.php>).[49] Η έκδοση 3.4 του Oiled είναι γραμμένη σε Java και διατίθεται ελεύθερα.

## 2.13.4 OntoEdit



εικόνα 2.13.4: Στιγμιότυπο οθόνης του OntoEdit

Το OntoEdit (<http://www.ontoknowledge.org/tools/ontoedit.shtml>)[50] είναι ένα ΣΔΟ κατάλληλο για την ανάπτυξη και διαχείριση οντολογιών μέσω γραφικού περιβάλλοντος. Το OntoEdit βασίζεται σε ένα πολυδύναμο εσωτερικό μοντέλο οντολογιών. Έτσι δίνει τη δυνατότητα μοντελοποίησης και ανάπτυξης οντολογιών ανεξάρτητα από τη γλώσσα αναπαράστασης.

Ο χρήστης μπορεί να οπτικοποιεί την οντολογία καθώς την αναπτύσσει μέσω πολλών ειδών γραφικών απεικονίσεων.

Το εργαλείο επιτρέπει στο χρήστη να επεξεργάζεται ιεραρχικά δομημένες έννοιες ή κλάσεις (εικόνα 2.13.4). Οι έννοιες μπορούν να οριστούν ως αφηρημένες ή συγκεκριμένες και κατά συνέπεια να επιτρέπεται ή να μην επιτρέπεται η δημιουργία άμεσων στιγμιοτύπων τους. Μια έννοια μπορεί να έχει πολλαπλά ονόματα. Με αυτό τον τρόπο μπορούν να οριστούν συνώνυμα μιας έννοιας.

Το εργαλείο παρέχει λειτουργικότητα παρόμοια με την 'αντιγραφή και επικόλληση' για την αναδιοργάνωση της ιεραρχίας των εννοιών.

Το OntoEdit βασίζεται σε ένα ευέλικτο πλαίσιο εργασίας πρόσθετων λειτουργικών μονάδων (plugins). Αυτό επιτρέπει στον χρήστη την επέκταση του εργαλείου με την απλή προσθήκη των απαιτούμενων λειτουργικών μονάδων.

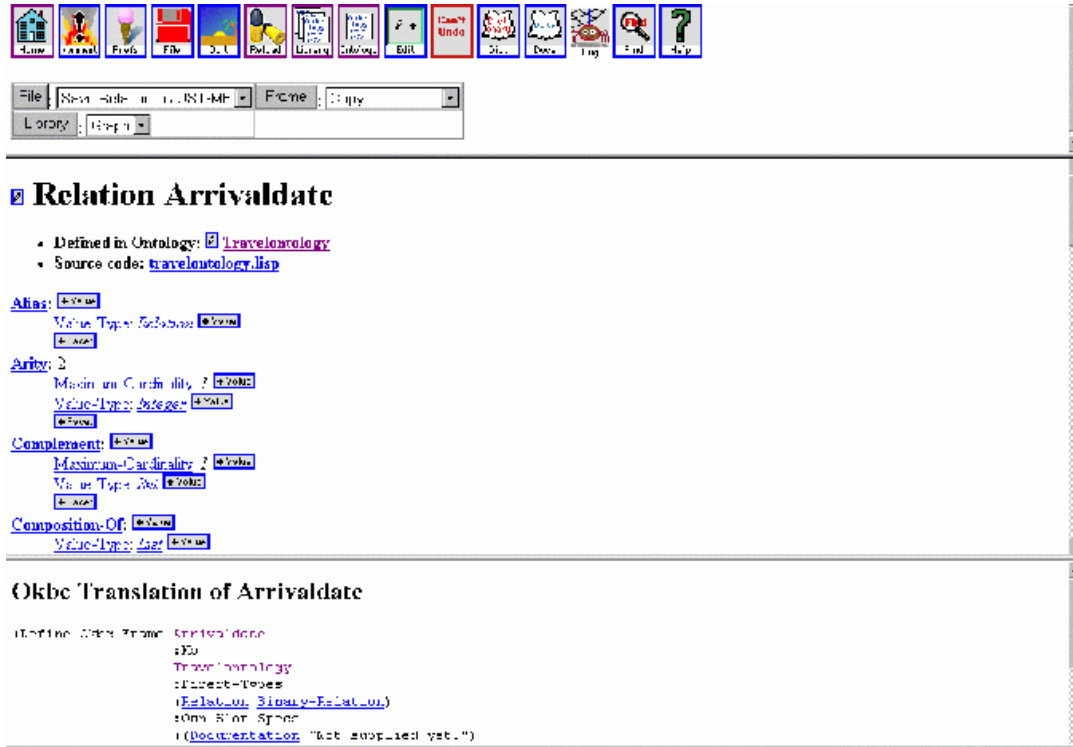
Έχοντας ο χρήστης στη διάθεσή του μια συλλογή από λειτουργικές μονάδες, όπως ένα θεματικό λεξικό, μια μηχανή εξαγωγής συμπερασμάτων, και μονάδες εισαγωγής και εξαγωγής από και προς διάφορες μορφές και γλώσσες αναπαράστασης οντολογιών, μπορεί να προσαρμόζει το εργαλείο σε διάφορα σενάρια χρήσης.

Όλες οι εκδόσεις του OntoEdit είναι διαθέσιμες σε δύο σειρές: ελεύθερη και επαγγελματική. Οι επαγγελματικές εκδόσεις συνήθως περιλαμβάνουν κάποιες επιπλέον πρόσθετες λειτουργικές μονάδες, όπως περιβάλλοντα συνεργασίας και δυνατότητες εξαγωγής συμπερασμάτων, που δεν περιέχονται στις ελεύθερες εκδόσεις.

### **2.13.5 Ontolingua Server**

Ο Ontolingua Server (<http://www.ksl.stanford.edu/software/ontolingua/>)[51] είναι μια συλλογή εργαλείων και υπηρεσιών για την κατασκευή οντολογιών διαμοιραζόμενων μεταξύ διασκορπισμένων ομάδων. Αναπτύσσεται από το Knowledge System Laboratory (KSL) στο πανεπιστήμιο του Stanford. Ο Ontolingua Server παρέχει πρόσβαση σε μια βιβλιοθήκη οντολογιών, υπηρεσίες μετάφρασης μεταξύ γλωσσών όπως Prolog, CORBA IDL, CLIPS, Loom και άλλες, και έναν επεξεργαστή για την κατασκευή και παρουσίαση οντολογιών (εικόνα 2.13.5). Απομακρυσμένοι επεξεργαστές μπορούν να αναζητήσουν και να επεξεργαστούν οντολογίες και απομακρυσμένες ή τοπικές εφαρμογές μπορούν

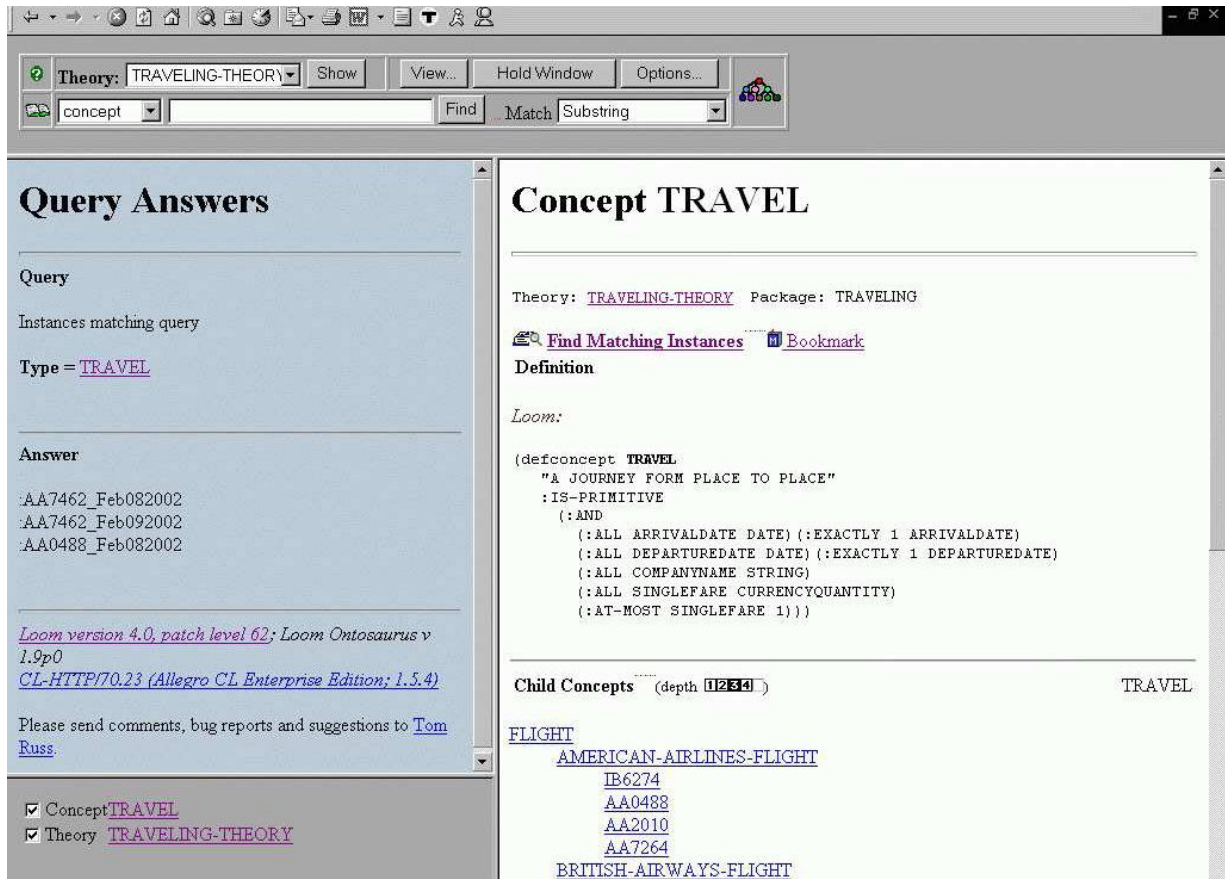
να προσπελάσουν τις οντολογίες της βιβλιοθήκης με τη χρήση του πρωτοκόλλου OKBC (Open Knowledge Based Connectivity).



εικόνα 5: Στιγμιότυπο οθόνης του Ontolingua Server



## 2.13.6 OntoSaurus



The screenshot displays the OntoSaurus web interface. The top navigation bar includes a 'Theory' dropdown set to 'TRAVELING-THEORY', and buttons for 'Show', 'View...', 'Hold Window', and 'Options...'. Below this is a search bar with 'concept' selected and a 'Find' button. The main content area is split into two panels. The left panel, titled 'Query Answers', shows the query 'Instances matching query' with 'Type = TRAVEL' and a list of three identifiers: AA7462\_Feb082002, AA7462\_Feb092002, and AA0488\_Feb082002. It also includes version information for Loom and a link to send comments. The right panel, titled 'Concept TRAVEL', shows the theory 'TRAVELING-THEORY' and package 'TRAVELING'. It features a 'Definition' section with a Loom code block: 

```
(defconcept TRAVEL
  "A JOURNEY FORM PLACE TO PLACE"
  :IS-PRIMITIVE
  (:AND
    (:ALL ARRIVALDATE DATE) (:EXACTLY 1 ARRIVALDATE)
    (:ALL DEPARTUREDATE DATE) (:EXACTLY 1 DEPARTUREDATE)
    (:ALL COMPANYNAME STRING)
    (:ALL SINGLEFARE CURRENCYQUANTITY)
    (:AT-MOST SINGLEFARE 1)))
```

 Below the definition is a 'Child Concepts' section showing a tree structure: TRAVEL (depth 1234) containing FLIGHT, which further contains AMERICAN-AIRLINES-FLIGHT (with instances IB6274, AA0488, AA2010, AA7264) and BRITISH-AIRWAYS-FLIGHT.

εικόνα 2.13.6: Στιγμιότυπο οθόνης του OntoSaurus

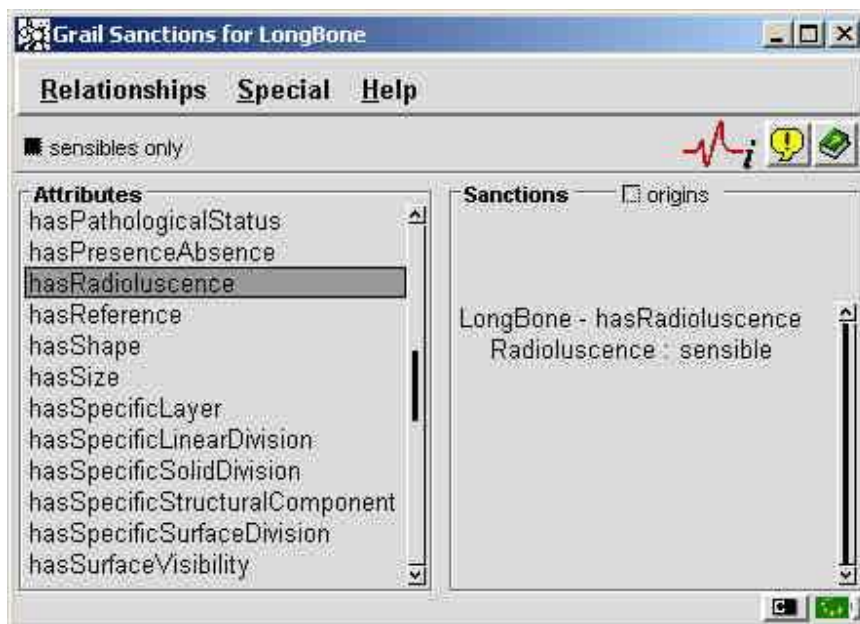
Το OntoSaurus (<http://www.isi.edu/isd/ontosaurus.html>)[52] αναπτύσσεται από το Information Sciences Institute (ISI) στο πανεπιστήμιο της Νότιας Καλιφόρνιας. Αποτελείται από δύο υποσυστήματα: Έναν server προσπέλασης οντολογιών που χρησιμοποιεί τη γλώσσα LOOM για την αναπαράσταση γνώσης, και ένα server παρουσίασης οντολογιών που δημιουργεί δυναμικά σελίδες html στις οποίες απεικονίζεται η ιεραρχική δομή των οντολογιών (εικόνα 2.13.6). Ο χρήστης μπορεί να



επεξεργαστεί οντολογίες μέσω  ορμώνhtml, ενώ από το σύστημα παρέχεται η δυνατότητα μετάφρασης απόLOOM σε Ontolingua, KIF, KRSS και C++

### 2.13.7 OpenKnoME

Το OpenKnoME (<http://topthing.com>)[53] αποτελεί μια μεγάλη σουίτα ε<sup>κ</sup> αρμογώνγια συνεργατική ανάπτυξη οντολογιών. Το OpenKnoME χρησιμοποιεί τη γλώσσα μοντελοποίησης εννοιών GRAIL. Μια σημαντική εφαρμογή αυτής της σουίτας είναι το Tigger που δημιουργήθηκε με σκοπό τη γρήγορη και εύκολη εξαγωγή γνώσης από οντολογίες για εμπειρογνώμονες και ειδικούς που δεν γνωρίζουν μηχανική οντολογιών. Τα εργαλεία που απαρτίζουν τη σουίτα είναι ελεύθερα διαθέσιμα και έχουν αναπτυχθεί



Σχήμα 2.13.7: Στιγμιότυπο οθόνης του OpenKnoME

απότο πανεπιστήμιο του Manchester. Η ανάπτυξή τους βασίστηκε σε διάφορα μεγάλα ιατρικά και  αρμακευτικάπρογράμματα που κάνουν χρήση οντολογιών, όπως τα GALEN (1992 - 1995), GALEN-IN-USE (1996 - 1999) και PRODIGY (1999 - 2001).Το

OpenKnoME χρησιμοποιείται για ανάπτυξη σε GRAIL, μια γλώσσα που αναπτύχθηκε στο πανεπιστήμιο του Manchester για να χρησιμοποιηθεί στα προγράμματα GALEN. Η GRAIL χρησιμοποιείται για την περιγραφή λογικής και λογικών γράφων. Το γνωστικό μοντέλο του OpenKnoME είναι επηρεασμένο σε μεγάλο βαθμό από την GRAIL. Έτσι εφόσον η GRAIL είναι σχεδιασμένη για την περιγραφή εννοιών, το OpenKnoME δεν μπορεί να χρησιμοποιηθεί για τη συλλογή στιγμιότυπων δεδομένων.

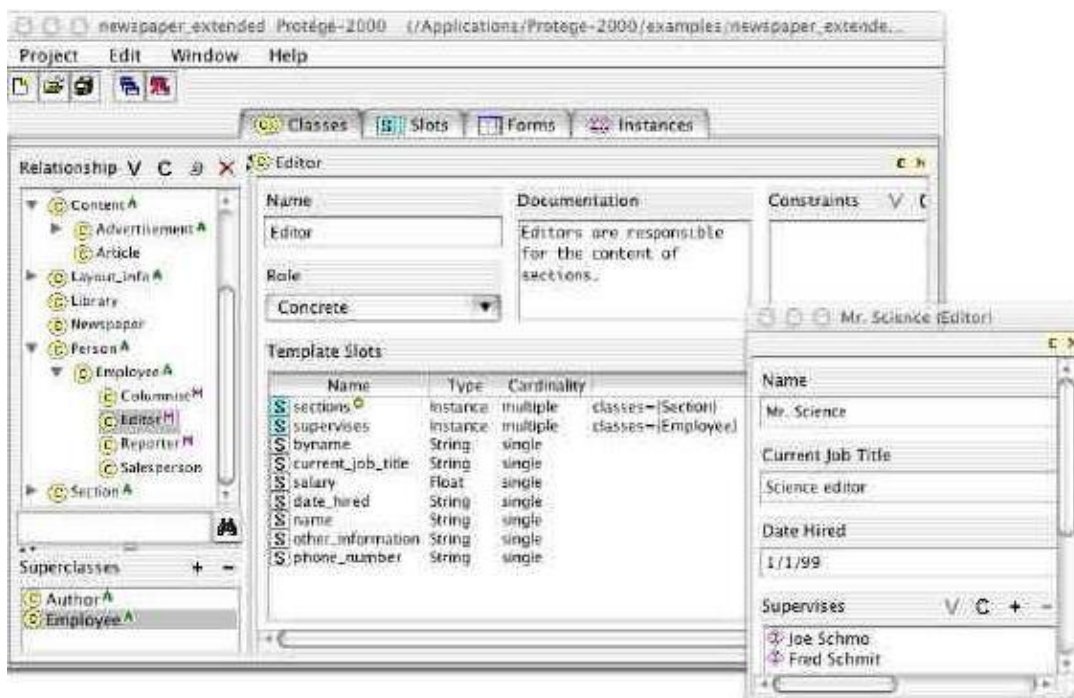
Το OpenKnoME δεν είναι μια 'stand-alone' εφαρμογή: επικοινωνεί με έναν GALEN Terminology Server (TeS) μέσω ενός καλώς ορισμένου API. Ο γραμμένος σε GRAIL πηγαίος κώδικας μετατρέπεται σε ένα μεταγλωττισμένο εννοιολογικό μοντέλο. Ο TeS αποθηκεύει και διαχειρίζεται το μοντέλο με τη βοήθεια ξεχωριστών υπομημάτων που παρέχουν διάφορα είδη υπηρεσιών: εννοιολογικές, γλωσσικές, συγγραφή κώδικα και άλλες υπηρεσίες. Το API παρέχει ένα σαφή διαχωρισμό της οντολογίας από τους πελάτες που την χρησιμοποιούν. Έτσι η οντολογία παρουσιάζεται ως υπηρεσία και όχι σαν μια δομή δεδομένων. Μέσω αυτής της υπηρεσίας το OpenKnoME μπορεί να εξετάζει, να αναζητά, να παρουσιάζει την οντολογία και να ελέγχει την ποιότητά της. Εφόσον η οντολογία πρέχεται με τη μορφή υπηρεσίας, η χρήση της δεν απαιτεί την εξαγωγή της σε κάποια στατική μορφή, αλλά γίνεται μέσω ενός TeS που δέχεται αιτήσεις από πελάτες. Ωστόσο το OpenKnoME παρέχει ένα παραμετροποιήσιμο εργαλείο εξαγωγής που υποστηρίζει εξαγωγή της οντολογίας σε HTML και CLIPS.

Το OpenKnoME έχει χρησιμοποιηθεί για την ανάπτυξη οντολογιών για διάφορα ευρωπαϊκά έργα συστηματοποίησης της ιατρικής ορολογίας, στο έργο NHS για τη δημιουργία μιας οντολογίας □ αρμακευτικών σκευασμάτων, στην ανάπτυξη του προτύπου HL7 για εφαρμογές υγείας και σε άλλα έργα. Περιέχει μια σουίτα εργαλείων για τη διαχείριση κώδικα οντολογιών από πολλούς παράλληλους συγγραφείς και τον έλεγχο του κώδικα. Η εφαρμογή Tigger διευκολύνει την εισαγωγή θεματικής γνώσης στο σύστημα. Οι ειδικοί εκπαιδεύονται στη χρήση ενός □ ιλικού προς τον χρήστη διαμεσολαβητικού περιβάλλοντος (Intermediate Representation - IR). Οι ειδικοί συντάσσουν μέσα στο περιβάλλον αυτό έννοιες με τη χρήση ενός γραφικού περιβάλλοντος ή ενός απλού επεξεργαστή κειμένου. Στη συνέχεια οι έννοιες που συνέταξαν διάφοροι ειδικοί συνδυάζονται και μεταφράζονται αυτόματα μέσω του Tiger σε κώδικα GRAIL. Έτσι αντιμετωπίζεται το πρόβλημα της εισαγωγής γνώσης στο


σύστημα. Το OpenKnoME είναι υλοποιημένο με τη σουίτα ανάπτυξης εφαρμογών Smalltalk και διατίθεται ελεύθερα.

### 2.13.8 Protege

Το Protege (<http://protege.stanford.edu>)[54] είναι ένα εργαλείο για διαχείριση γνώσης που αναπτύσσεται από το πανεπιστήμιο του Stanford. Χιλιάδες χρήστες σε όλο τον κόσμο χρησιμοποιούν το Protege σε έργα που ποικίλουν από τη μοντελοποίηση της νόσου του καρκίνου, έως την μοντελοποίηση πυρηνικών σταθμών παραγωγής ενέργειας. Το Protege είναι ελεύθερο λογισμικό και διατίθεται κάτω από την άδεια χρήσης ελεύθερου κώδικα “Mozilla open-source license”. Το Protege 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 προσπελάσουν τις σχετικές πληροφορίες γρήγορα όταν τις χρειάζονται και να διαμορφώνουν και να χειρίζονται οντολογίες. Η ιεραρχική δομή της οντολογίας



εικόνα 2.13.8: Στιγμιότυπο οθόνης του Protege

αναπαριστάται με τη μορφή δένδρου δίνοντας στο χρήστη γρήγορη και εύκολη πρόσβαση στις κλάσεις και υπο-κλάσεις. Για την εισαγωγή τιμών στις ιδιότητες το Protege χρησιμοποιεί ειδικές  όρμες (εικόνα 2.13.8). Το γνωστικό μοντέλο του Protege 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 Protege διακρίνεται από τα υπόλοιπα περιβάλλοντα διαχείρισης οντολογιών για τις δυνατότητες εξέλιξης και επέκτασης που προσφέρει.


Το Protege 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 Protege περιέχει συνδρομές από χιλιάδες προγραμματιστές από όλο τον κόσμο. Κάποιες από τις διαθέσιμες προσθήκες παρέχουν λειτουργίες όπως :

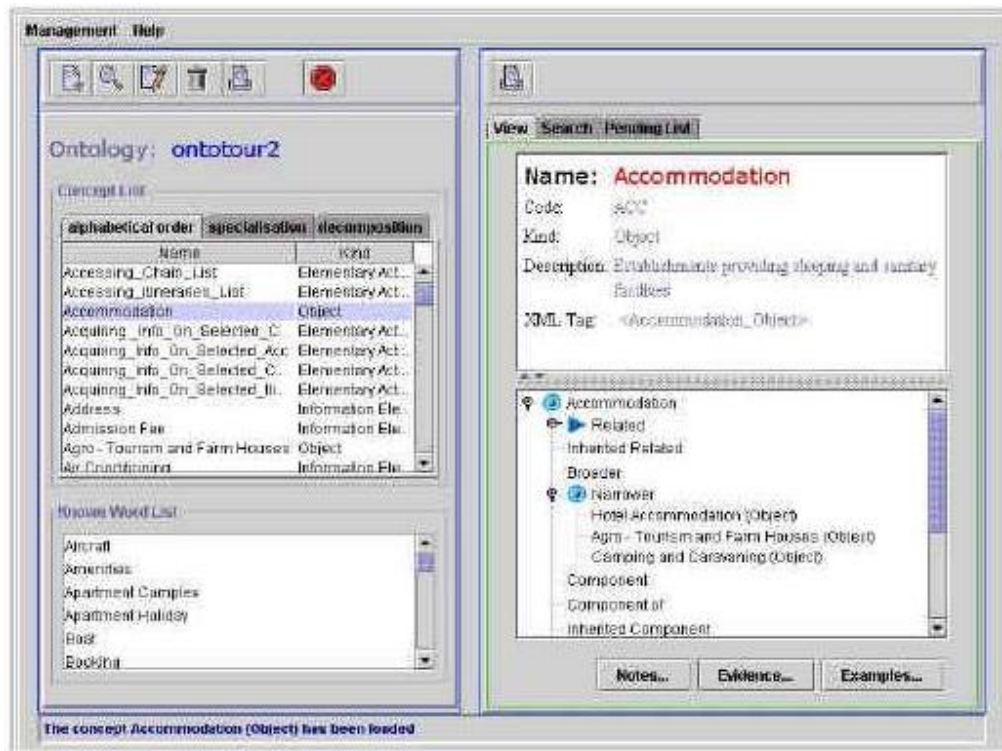
- υποστήριξη για αποθήκευση και εισαγωγή οντολογιών γραμμένων σε RDF Schema, αρχεία XML με DTD, αρχεία XML Schema και DAML+OIL,
- εμφάνιση εικόνων μορφής GIF, βίντεο και ήχου, καθώς και έναν επεξεργαστή διαγραμμάτων που επιτρέπει στον χρήστη να κατασκευάσει μια βάση γνώσης σχεδιάζοντας ένα διάγραμμα 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.

### 2.13.9 SymOntoX

Το SymOntoX (Symbolic OntologyManager XML savvy) είναι ένα πρότυπο εφαρμογής για τη διαχείριση θεματικών οντολογιών (domain ontologies). Έχει αναπτυχθεί από το LEKS (Laboratory for Enterprise Knowledge and Systems) στο IASI-CNR. Ένα στιγμιότυπο της εφαρμογής  αίνετα στην εικόνα 2.13.9. Το SymOntoX μοντελοποιεί τις θεματικές έννοιες και τις μεταξύ τους σχέσεις σύμφωνα με το OPAL (Object, Process and Actor modelling Language), μια μεθοδολογία για αναπαράσταση οντολογιών που αναπτύχθηκε στο IASI-CNR. Το SymOntoX χρησιμοποιείται στο ευρωπαϊκό έργο Harmonise που δραστηριοποιείται στο θεματικό πεδίο του τουρισμού. Σύμφωνα με το OPAL, οι έννοιες οργανώνονται ως εξής :

- *Στοιχείο δράσης* (actor) είναι κάθε οντότητα του  εματικού πεδίου που είναι ικανή να ενεργοποιήσει ή να εκτελέσει κάποια διαδικασία (πχ. ένας τουρίστας, ένα γραφείο ταξιδίων).
- *Αντικείμενο* (object) είναι κάθε παθητική οντότητα πάνω στην οποία πραγματοποιείται μια εργασία (πχ. Ξενοδοχείο, Πτήση).
- *Διαδικασία* (process) είναι κάθε ενέργεια που αποσκοπεί στην επίτευξη κάποιου στόχου (πχ. κάνω\_κράτηση).



εικόνα 2.13.9: Στιγμιότυπο οθόνης του SymOntoX

Πέρα από αυτά τα βασικά ήδη εννοιών, το OPAL προτείνει τα παρακάτω συμπληρωματικά είδη:

- *Τμήμα πληροφορίας* (Information Component) είναι κάθε δέσμη πληροφορίας που σχετίζεται με κάποιο στοιχείο δράσης ή κάποιο αντικείμενο (πχ. πληροφορίες πτήσης, διεύθυνση ξενοδοχείου).
- *Στοιχειώδης πληροφορία* (Information Element) είναι κάθε ατομικό στοιχείο πληροφορίας που είναι μέρος ενός τμήματος πληροφορίας (πχ.τιμή πτήσης, αριθμός δωματίων).
- *Πράξη* (Action) είναι μια ενέργεια που αναπαριστά ένα τμήμα διαδικασίας που μπορεί να αναλυθεί περαιτέρω (πχ. αίτηση για δωμάτιο).
- *Στοιχειώδης πράξη* (Elementary Action) είναι μια ενέργεια που αναπαριστά ένα τμήμα διαδικασίας που δεν μπορεί να αναλυθεί περαιτέρω (πχ. ακύρωση κράτησης).

- *Στόχος (Goal)* είναι μια επιθυμητή κατάσταση στην οποία ένα στοιχείο δράσης προσπαθεί να □ τάσει(πχ. πάω διακοπές).
- *Κατάσταση (State)* είναι κάθε χαρακτηριστικό στιγμιότυπο που μπορεί να αναγνωριστεί (πχ. η πτήση είναι γεμάτη).
- *Κανόνας (Rule)* είναι μια έκφραση που έχει ως στόχο να περιορίσει τις πιθανές τιμές ενός στιγμιότυπου ή να εξορύξει νέα πληροφορία (πχ. αγορά εισιτηρίου 30 μέρες πριν την αναχώρηση).

Οι παραπάνω χαρακτηρισμοί είναι απαραίτητοι για τον ορισμό εννοιών. Σύμφωνα με το OPAL οι έννοιες διασυνδέονται μεταξύ τους μέσω των οντολογικών σχέσεων Specialisation, Decomposition, Predication, Similarity και Relatedness. Το SymOntoX προορίζεται να λειτουργήσει ως διαδικτυακή υπηρεσία προσπελάσιμη μέσω ενός □ υλλομετρητήΒασίζεται κυρίως σε τεχνολογίες XML και Java με στόχο τη μέγιστη ευελιξία, διαλειτουργικότητα και ανεξαρτησία από υπολογιστική πλατφόρμα. Το SymOntoX μπορεί να διαχειριστεί διαφόρων ειδών οντολογίες, τύπους χρηστών και χρήσεων. Ένας χρήστης μπορεί να εγγραφεί στο σύστημα ως απλός *χρήστης* με δικαιώματά ανάγνωσης μόνο, ως *υπερ-χρήστης* με δικαίωμα υποβολής προτάσεων για νέες έννοιες ή ως *διαχειριστής οντολογιών* με πλήρη δικαιώματα διαχείρισης της οντολογίας και αποδοχής ή απόρριψης των προτάσεων που κάνουν οι υπερ-χρήστες. Το σύστημα μπορεί να χρησιμοποιηθεί ως λεξικό όρων (είναι ορατά μόνο το όνομα και η περιγραφή σε φυσική γλώσσα των εννοιών), θησαυρός (είναι επιπλέον ορατά η ιεραρχία και οι σχέσεις ομοιότητας), σύστημα οντολογιών (απεικονίζονται όλες οι σχέσεις) ή βάση γνώσης (□ αίνονται τα στιγμιότυπα των εννοιών).

Το SymOntoX έχει ένα γραφικό περιβάλλον με □ όρμεγια την επεξεργασία και την επισκόπηση των οντολογιών και τη δυνατότητα απεικόνισης διαγραμμάτων για περιήγηση στο περιεχόμενο των οντολογιών. Τέλος ένα Java API παρέχει την διαλειτουργικότητα και ενσωμάτωση σε άλλα συστήματα, ενώ ένας ελεγκτής διασφαλίζει τη συνέπεια των οντολογιών με βάση τα αξιώματα του OPAL. Όλα τα δεδομένα του συστήματος αποθηκεύονται σε μια XML βάση δεδομένων.

## 2.13.10 WebODE

The screenshot shows the WebODE 2.0 interface. At the top, there is a navigation bar with buttons: 'Show Term Properties', 'Relations With Other Ontologies', 'Graphical Taxonomy Edition', 'Intermediate Representations', 'Inference Engine', 'Instances', and 'Back'. A search bar contains the text 'knowe>'. Below the navigation bar, the title 'Instance Attributes for Term Travel' is displayed. On the left, there is a 'Properties Groups' sidebar with a tree view of ontologies. The 'Travel' ontology is selected. The main area contains a table of instance attributes for the term 'Travel' and a form for editing these attributes. The table has columns: 'Instance Attribute Name', 'Description', 'Value Type', 'Cardinality', 'Measurement Unit', 'Precision', and 'Value Interval'. The form has fields for 'Term Name', 'Instance Attribute Name', 'Description', 'Value Type', 'Minimum-Maximum Cardinality', 'Measurement Unit', 'Precision', 'Minimum Value', and 'Maximum Value'. Red annotations highlight the 'Clipboard' icon, the 'Browsing area' (the ontology tree), and the 'Edition area' (the attribute form).

Instance Attribute Name	Description	Value Type	Cardinality	Measurement Unit	Precision	Value Interval
arrival_date	Date of arrival of the trip	Date	(0, 1)			
company_name	Transportation company or companies in charge of a trip.	String	(0, N)			
departure_date	Date of departure of the trip	Date	(0, 1)			
single_fare	Fare of a single ticket	Float	(0, 1)			

εικόνα 2.13.10: Στιγμιότυπο οθόνης του WebODE

Το WebODE (<http://webode.dia.fi.upm.es/>)[55] είναι ένα εργαλείο μηχανικής οντολογιών που παρέχει διάφορες υπηρεσίες σχετικές με την ανάπτυξη, διαχείριση και χρήση οντολογιών. Το WebODE χρησιμοποιεί ένα εκφραστικό μοντέλο γνώσης για την αναπαράσταση των οντολογιών που βασίζεται στη μεθοδολογία αναπαράστασης οντολογιών METHONTOLOGY. Η μεθοδολογία METHONTOLOGY περιέχει δομικά στοιχεία οντολογιών, όπως έννοιες, διαμερίσεις, ταξονομικές και άλλες προκαθορισμένες σχέσεις, στιγμιότυπα, αξιώματα, κανόνες, σταθερές και βιβλιογραφικές αναφορές. Επίσης επιτρέπει τη χρήση όρων από άλλες εξωτερικές οντολογίες.

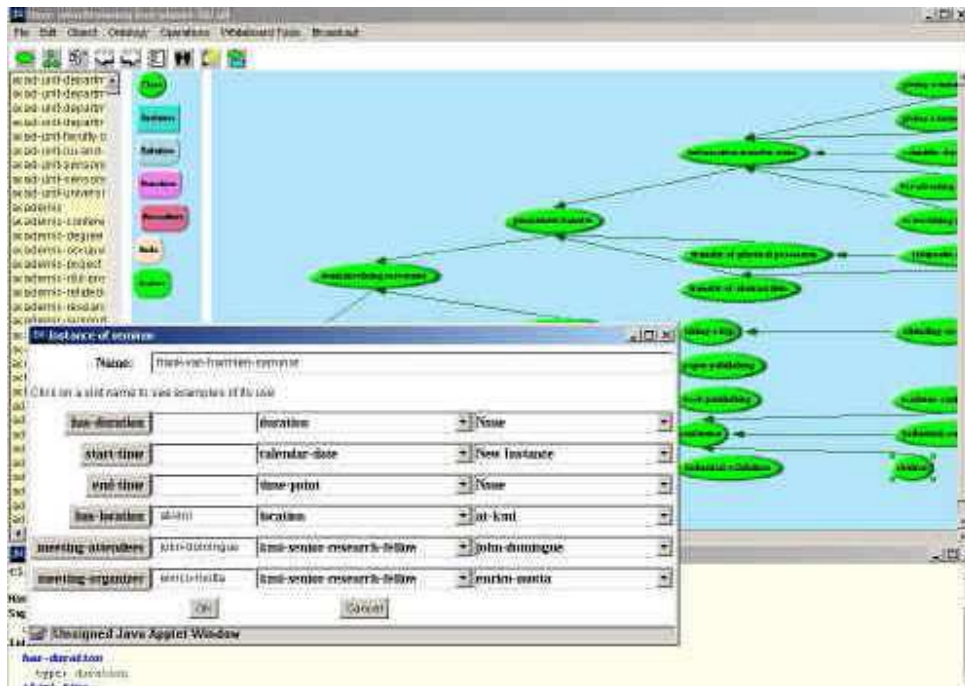
Οι οντολογίες στο WebODE αποθηκεύονται σε μια σχεσιακή βάση δεδομένων. Το WebODE προσφέρει επίσης ένα καλά ορισμένο API για πρόσβαση στις οντολογίες που



καθιστά εύκολη την ενσωμάτωσή τους σε άλλα συστήματα. Οι οντολογίες που κατασκευάζονται με το WebODE μπορούν εύκολα να ενσωματωθούν σε άλλα συστήματα και με τη χρήση της λειτουργίας εξαγωγής και εισαγωγής από και προς XML, από και προς διάφορες γλώσσες αναπαράστασης οντολογιών (RDF(S), OIL, DAML+OIL, CARIN, FLogic) και σε άλλες γλώσσες προγραμματισμού και συστήματα όπως Java και Jess. Η επεξεργασία οντολογιών στο WebODE πραγματοποιείται μέσω φορμών και ενός γραφικού περιβάλλοντος, ενός ελεγκτή συνέπειας, μιας μηχανής εξαγωγής συμπερασμάτων, ενός συστήματος κατασκευής αξιωμάτων και μιας υπηρεσίας συγγραφής τεκμηρίωσής. Δυο καινοτόμα χαρακτηριστικά του WebODE είναι τα σύνολα στιγμιοτύπων που επιτρέπουν την παράλληλη δημιουργία στιγμιοτύπων του ίδιου εννοιολογικού μοντέλου για διαφορετικά σενάρια, και εννοιολογικών οπτικοποιήσεων για το ίδιο μοντέλο. Τα σύνολα στιγμιοτύπων επιτρέπουν τη δημιουργία και αποθήκευση διάφορων τμημάτων της οντολογίας και την παραμετροποίηση της οπτικοποίησης της οντολογίας για κάθε χρήστη.

### **2.13.11 WebOnto**

Το WebOnto είναι ένα εργαλείο 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. Τα βασικά χαρακτηριστικά του είναι : διαχείριση οντολογιών μέσω γραφικού περιβάλλοντος (εικόνα 2.13.11), αυτόματη δημιουργία στιγμιοτύπων μέσω  ορμών 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.

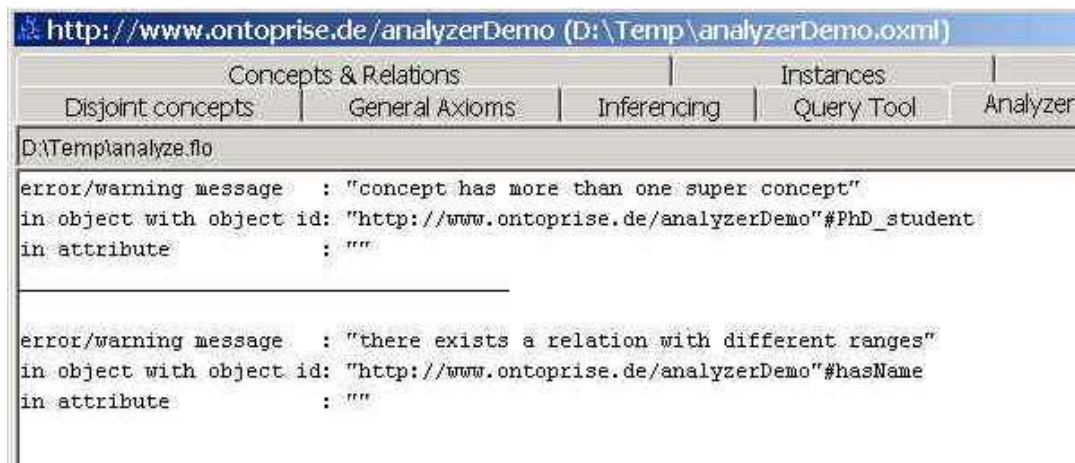


εικόνα 2.13.11: Στιγμιότυπο οθόνης του WebOnto

## 2.14 Εργαλεία αξιολόγησης οντολογιών

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 και δυνατότητας διαβάθμισης (scalability). 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 :

### 2.14.1 OntoAnalyser



εικόνα 2.14.1: Στιγμιότυπο οθόνης του OntoAnalyser

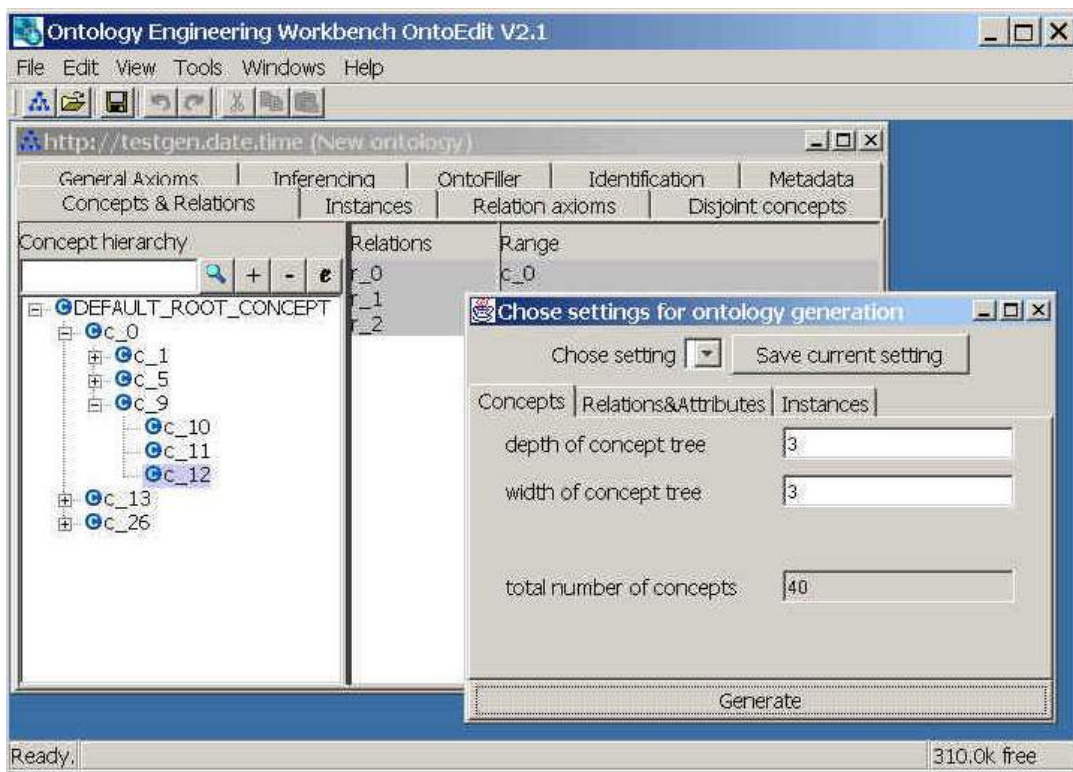
Το OntoAnalyser (<http://www.ontoprise.de/>)[56] έχει υλοποιηθεί ως πρόσθετο (plugin) για την εφαρμογή OntoEdit που αποτελεί ένα γραφικό περιβάλλον για ανάπτυξη οντολογιών. Το OntoAnalyser εστιάζει την αξιολόγηση στη συμμόρφωση με τους κανόνες της γλώσσας στην οποία είναι εκφρασμένη η οντολογία και στην συνέπεια (consistency).

Για την αναπαράσταση των κανόνων σύμφωνα με τους οποίους γίνεται η αξιολόγηση χρησιμοποιείται η γλώσσα λογικής F-logic. Το OntoAnalyser παίρνει σαν είσοδο κανόνες

αξιολόγησης, τους εκτελεί στη μηχανή εξαγωγής συμπερασμάτων OntoBroker και παρουσιάζει στο χρήστη τα αποτελέσματα του ελέγχου.

Ως πρόσθετο του OntoEdit, το OntoAnalyser, είναι πλήρως ενσωματωμένο στο περιβάλλον ανάπτυξης. Ο χρήστης μπορεί να αξιολογεί μια οντολογία καθώς την αναπτύσσει στο OntoEdit. Έτσι δεν χρειάζεται να εισάγει, να εξάγει και να μετατρέψει την οντολογία, προκειμένου να την αξιολογήσει.

## 2.14.2 OntoGenerator



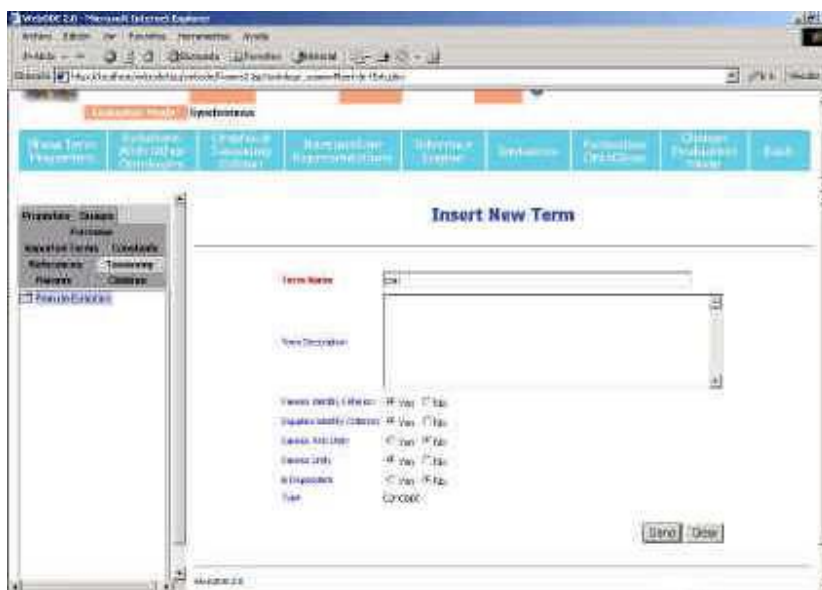
εικόνα 2.14.2: Στιγμιότυπο οθόνης του OntoGenerator

Το OntoGenerator (<http://www.ontoprise.de/>)[57] είναι κατάλληλο για την διενέργεια δοκιμών απόδοσης (stress tests) για εφαρμογές που βασίζονται σε οντολογίες. Κατασκευάζει 'συνθετικές' οντολογίες που δεν προορίζονται για την περιγραφή κάποιας θεματικής περιοχής, αλλά είναι κατασκευασμένες έτσι ώστε να πληρούν κάποιες τεχνικές προδιαγραφές, όπως συγκεκριμένο αριθμό εννοιών και στιγμιότυπων ή κανόνες

συγκεκριμένου είδους. Προκειμένου να αποτιμηθεί η απόδοση εργαλείων που βασίζονται στη χρήση οντολογιών είναι απαραίτητη η δοκιμή τους με διάφορων ειδών οντολογίες. Έτσι το OntoGenerator σχεδιάστηκε για την παραγωγή συνθετικών οντολογιών με προκαθορισμένες παραμέτρους. Ο χρήστης μπορεί να καθορίσει τις ακόλουθες παραμέτρους για τις παραγόμενες οντολογίες :

- το βάθος του εννοιολογικού δένδρου,
- το πλάτος του εννοιολογικού δένδρου,
- το πλήθος σχέσεων,
- το πλήθος ιδιοτήτων και
- το πλήθος στιγμιοτύπων.

### 2.14.3 OntoClean



εικόνα 2.14.3: Στιγμιότυπο οθόνης του OntoClean

Η ομάδα οντολογιών του National Research Council (CNR) στην Padova της Ιταλίας έχει αναπτύξει μια σειρά από κριτήρια αξιολόγησης οντολογιών βασισμένα σε  $\square$  ιλοσοφικές έννοιες, όπως: αυστηρότητα (rigidity), ομοιότητα (identity), ενότητα (unity), εξάρτηση (dependence), και άλλες. Αυτά τα κριτήρια χρησιμοποιούνται στη μεθοδολογία

αξιολόγησης οντολογιών OntoClean που έχει προταθεί από την ίδια ομάδα για τον καθαρισμό πεπλεγμένων οντολογιών.

Παράλληλα, η ομάδα οντολογιών του Laboratorio de Inteligencia Artificial του Univesidad Politecnica de Madrid (UPM) έχει αναπτύξει μια μεθοδολογία καλούμενη METHONTOLOGY για την ανάπτυξη, μετατροπή, επαναχρησιμοποίηση και αξιολόγηση οντολογιών. Η ομάδα του UPM έχει επίσης κατασκευάσει το WebODE, ένα εργαλείο που υποστηρίζει τα βασικά βήματα (σχεδιασμός, ανάπτυξη, τροποποίηση, επαναχρησιμοποίηση και αξιολόγηση) της μεθοδολογίας METHONTOLOGY. Εφόσον οι εργασίες των δυο παραπάνω ομάδων αλληλοσυμπληρώνονται, δημιουργήθηκε μια συνδυασμένη μεθοδολογία που ενσωματώνει την μεθοδολογία OntoClean στη φάση σχεδιασμού της METHONTOLOGY. Επιπλέον, ένα εργαλείο ονομαζόμενο OntoClean έχει υλοποιηθεί ως πρόσθετο (plugin) του WebODE που υποστηρίζει και τις δύο μεθοδολογίες. Το πρόσθετο OntoClean παρέχει τις παρακάτω λειτουργίες :

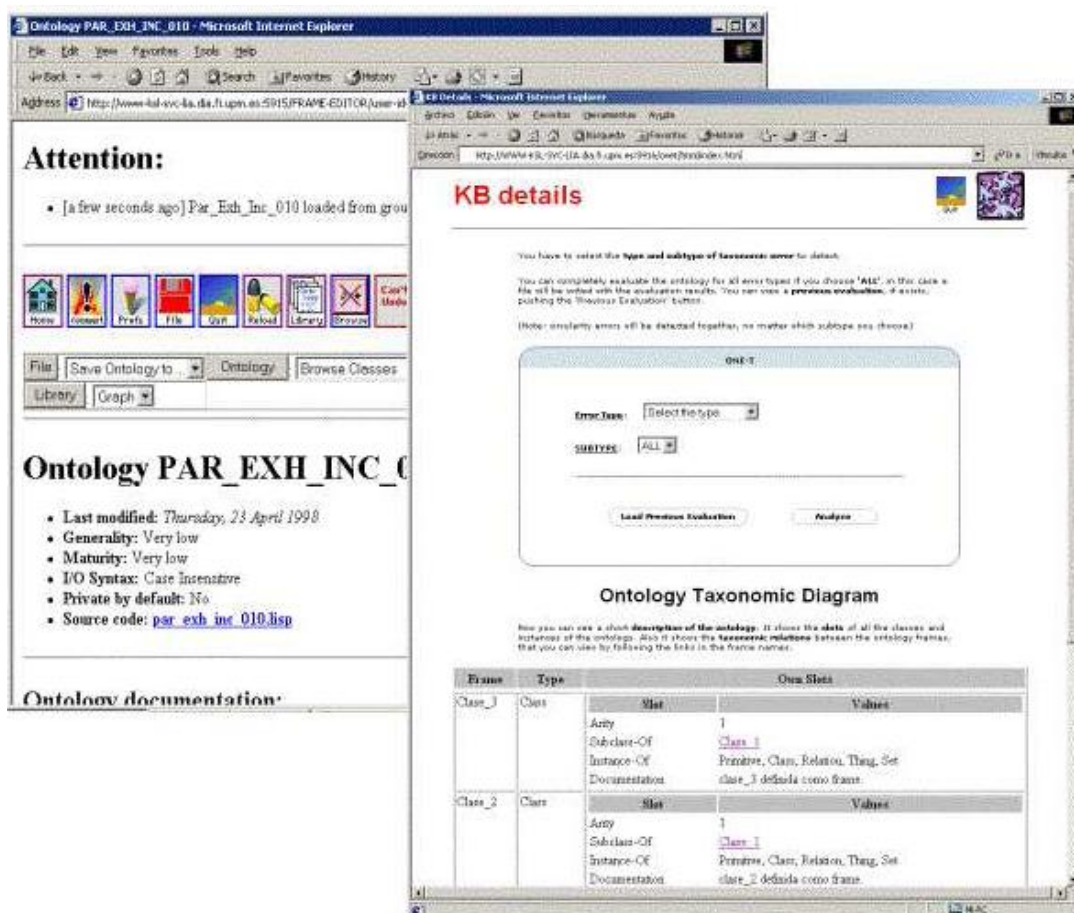
1. *Καθορισμός της μεθόδου αξιολόγησης.* Ο χρήστης μπορεί να επιλέξει εάν το σύστημα θα δείχνει τα σφάλματα αυτόματα αμέσως μόλις εντοπιστεί κάποια παραβίαση των αξιωμάτων του OntoClean ή εάν θα δείχνει τα σφάλματα μόνο όταν ο χρήστης το ηγήσει.
2. *Ανάθεση μετα-ιδιοτήτων στις έννοιες.* Ο χρήστης μπορεί να θέσει μετα-ιδιότητες που να αφορούν την ομοιότητα, ενότητα, εξάρτηση ή την αυστηρότητα των εννοιών.
3. *Εστίαση (ή μη-εστίαση) σε μη αυστηρά ορισμένες ιδιότητες.* Ο χρήστης μπορεί να επιλέξει οι μη αυστηρά ορισμένες ιδιότητες να είναι λιγότερο εμφανείς κατά την παρουσίαση της οντολογίας. Αυτή η δυνατότητα είναι σημαντική, μια και το OntoClean εστιάζει κυρίως στις αυστηρά ορισμένες ιδιότητες.
4. *Αξιολόγηση σύμφωνα με ταξονομικούς περιορισμούς.* Το σύστημα αξιολογεί την οντολογία με βάση τις αρχές της ομάδας του CNR. Ο χρήστης μπορεί κάνει την αξιολόγηση λιγότερο ή περισσότερο αυστηρή απλά επιλέγοντας λιγότερα ή περισσότερα κριτήρια.

Το βασικό πλεονέκτημα του OntoClean είναι ότι τα κριτήρια που χρησιμοποιεί για την αξιολόγηση είναι εκφρασμένα σε WebODE. Το OntoClean χρησιμοποιεί τη μηχανή συμπερασμού του WebODE που είναι υλοποιημένη σε Prolog για τον εντοπισμό

ασυνεπειών στην ταξινόμηση, επιτρέποντας στο χρήστη να καθαρίσει την οντολογία με βάση ταξονομικούς περιορισμούς. Μετά τον καθαρισμό, η οντολογία μπορεί να εξαχθεί σε άλλες γλώσσες υλοποίησης, μέσω μεταφραστών του WebODE.

## 2.14.4 ONE-T

Το ONE-T (Ontology Evaluation Tool) έχει αναπτυχθεί από την ομάδα οντολογιών του UPM. Το ONE-T είναι μια εφαρμογή Java που επιτρέπει την αξιολόγηση οντολογιών αποθηκευμένων σε οποιονδήποτε Ontolingua Server. Έχει χρησιμοποιηθεί για την αξιολόγηση των οντολογιών του Ontolingua Server 5.0 και Ontolingua Server 6.0. Είναι εύκολα επεκτάσιμο, μια και βασίζεται στο πρότυπο OKBC. Το ONE-T εντοπίζει τα ακόλουθα ήδη σφαλμάτων ασυνέπειας στις ταξονομίες εννοιών.



εικόνα 2.14.4: Στιγμιότυπο οθόνης του ONE-T



1. *Σφάλματα κυκλικότητας*. Εμφανίζονται όταν μια κλάση ορίζεται ως γενίκευση ή εξειδίκευση του εαυτού της. Με βάση το πλήθος των σχέσεων που εμπλέκονται σε έναν τέτοιο κύκλο, τα σφάλματα κυκλικότητας μπορούν να διακριθούν σε : σφάλματα κυκλικότητας περιμέτρου μηδέν (μια κλάση που ορίζεται με βάση τον εαυτό της), σφάλματα κυκλικότητας περιμέτρου 1 και σφάλματα κυκλικότητας περιμέτρου  $n$ .
2. *Σφάλματα διαμέρισης*. Μέσω των διαμερίσεων οι έννοιες μπορούν να διαχωριστούν σε ξένες μεταξύ τους ή πλήρης κλάσεις. Καθώς οι πλήρεις διαμερίσεις εισάγουν τον περιορισμό της πληρότητας στα παραγόμενα υποσύνολα μπορούμε να έχουμε σφάλματα διαμέρισης και σφάλματα πλήρους διαμέρισης.

Σφάλματα διαμέρισης :

- ∅ *Διαμέριση με κοινά στιγμιότυπα*. Παρατηρείται όταν ένα ή περισσότερα στιγμιότυπα ανήκουν σε περισσότερες από μια υπο-κλάσεις της διαμέρισης. Π. χ. εάν οι 'σκύλοι' και οι 'γάτες' αποτελούν μια διαμέριση του συνόλου των 'θηλαστικών', ένα σφάλμα τέτοιου τύπου θα αποτελούσε ο ορισμός του 'Πλούτο' ως στιγμιότυπου και των δυο αυτών κλάσεων. Ο συγγραφέας της οντολογίας θα πρέπει να αφαιρέσει την λανθασμένη σχέση για να λύσει αυτό το πρόβλημα.
- ∅ *Διαμέριση με κοινές κλάσεις*. Προκύπτει όταν μια ή περισσότερες κλάσεις αποτελούν ταυτόχρονα υπο-κλάσεις περισσότερων από μιας κλάσεων της διαμέρισης. Πχ. εάν οι 'σκύλοι' και οι 'γάτες' αποτελούν μια διαμέριση του συνόλου των 'θηλαστικών' ένα τέτοιο σφάλμα θα προέκυπτε αν ορίζαμε την κλάση 'Ντόπερμαν' ως υποκλάση και των δυο κλάσεων. Για να επιλυθεί το πρόβλημα πρέπει να αφαιρεθεί η εσφαλμένη σχέση.

Σφάλματα πλήρους διαμέρισης :

- ∅ *Πλήρης διαμέριση με εξωτερικά στιγμιότυπα*. Ένα τέτοιο σφάλμα προκύπτει όταν έχει οριστεί μια πλήρης διαμέριση των κλάσεων της οντολογίας και υπάρχουν στιγμιότυπα που δεν ανήκουν σε καμιά κλάση της διαμέρισης. Πχ. Αν οι 'αριθμοί' έχουν διαμεριστεί πλήρως σε 'άρτιους' και 'περιττούς' και ο αριθμός 'τέσσερα' έχει οριστεί ως στιγμιότυπο της



κλάσης 'αριθμός' (αντί της κλάσης 'άρτιος'), θα αντιμετωπίζαμε ένα τέτοιο σφάλμα.

- ∅ *Πλήρης διαμέριση με κοινά στιγμιότυπα.* Προκύπτει όταν ένα ή περισσότερα στιγμιότυπα ανήκουν σε περισσότερες από μια κλάσεις της πλήρους διαμέρισης. Στο παραπάνω παράδειγμα, ο ορισμός του αριθμού 'τέσσερα' ως στιγμιότυπου και των δυο κλάσεων 'άρτιοί' και 'περιπτοί' αποτελεί ένα τέτοιο σφάλμα.
- ∅ *Πλήρης διαμέριση με κοινές κλάσεις.* Προκύπτει όταν μια ή περισσότερες κλάσεις της οντολογίας είναι ορισμένες ως υπο-κλάσεις περισσότερων από μιας κλάσεων της πλήρους διαμέρισης. Πχ. Στο παράδειγμα με τους αριθμούς, αν οι αριθμοί έχουν διαμεριστεί πλήρως στις κλάσεις 'άρτιοί' και 'περιπτοί' και η κλάση 'πρώτος αριθμός' οριστεί ως υποκλάση και των 'άρτιων' και των 'περιπτών' αριθμών, το αποτέλεσμα είναι ασυνεπές, μια και ένα στιγμιότυπο 'πρώτου αριθμού' θα είναι ταυτόχρονα στιγμιότυπο και των 'άρτιων' και των 'περιπτών' αριθμών.

### 3. Πλεονασμός:

*Πλεονασμός σχέσεων subclass-of.* Προκύπτει όταν επαναλαμβάνονται σχέσεις subclass-of μεταξύ των κλάσεων. Μπορούμε να διακρίνουμε δύο ήδη επανάληψης : την άμεση και την έμμεση.

- ∅ *Άμεση επανάληψη* έχουμε όταν υπάρχουν δύο ή περισσότερες σχέσεις subclass-of μεταξύ των ίδιων δύο κλάσεων. Π. χ. όταν στην ΟΝΤΟΛΟΓΙΑ συμπεριλαμβάνεται δύο □ ορέξη σχέση subclass-of μεταξύ των κλάσεων 'σκύλος' και 'θηλαστικό'.
- ∅ *Έμμεση επανάληψη* έχουμε αν πχ. ορίσουμε την κλάση 'σκύλος' ως υποκλάση της 'κατοικίδιο', την 'κατοικίδιο' ως υποκλάση της 'ζώο' και επίσης την 'σκύλος' ως υποκλάση της 'ζώο'.

*Πλεονασμός σχέσεων instance-of.* Προκύπτει όταν επαναλαμβάνονται σχέσεις instance-of μεταξύ ενός στιγμιότυπου και μιας κλάσης. Κι εδώ μπορεί να έχουμε άμεση ή έμμεση επανάληψη.

### 4. Σφάλματα γραμματικής

*Σε σχέσεις subclass-of*, όταν τα ορίσματα της σχέση subclass-of δεν είναι κλάσεις.

*Σε σχέσεις instance-of* , όταν η σχέση instance-of δεν συσχετίζει ένα στιγμιότυπο με μια κλάση.

*Σε διαμερίσεις*, όταν ένα στιγμιότυπο αποτελεί μέρος μιας διαμέρισης.

*Πανομοιότυποι ορισμοί κλάσεων*. Προκύπτει όταν δύο ή περισσότερες κλάσεις έχουν πανομοιότυπους ορισμούς, με μόνη διαφορά τα διαφορετικά ονόματά τους.

## **2.15 Γλώσσες αναπαράστασης οντολογιών:**

### **2.15.1 OIL (Ontology Inference Layer)**

Η OIL είναι μια πρόταση τυποποίησης μιας διαδικασίας δημιουργίας μιας πολυεπίπεδης αρχιτεκτονικής προσδιορισμού οντολογιών. Βασίζεται πάνω σε γνωστές προτάσεις προτύπων του WWW Consortium όπως το XML και το RDF, εμπλουτίζοντας τα με νέα χαρακτηριστικά απαραίτητα για οντολογίες. Αποτελεί το πρώτο σημαντικό αποτέλεσμα του έργου On-To-Knowledge. Προέκυψε από την ανάγκη δημιουργίας μιας γλώσσας που θα μπορούσε να υποστηρίξει και να εκφράσει οντολογίες. Η δημιουργία της γλώσσας αυτής βασίστηκε σε τρεις παραδοχές:

1. Θα πρέπει να είναι απόλυτα κατανοητή στο χρήστη, και να εκμεταλλεύεται την επιτυχία των αντικειμενοστραφών και Frame-based συστημάτων
2. Θα πρέπει να έχει μια καλά ορισμένη τυπική αναπαράσταση εννοιών και να υποστηρίξει την έκφραση λογικής έτσι ώστε να υπάρχει πληρότητα, ορθότητα και αποδοτικότητα στην οντολογία
3. Θα πρέπει να βασίζεται σε υπάρχοντα πρότυπα όπως το XML τόσο για την εξασφάλιση της διαλειτουργικότητας όσο και για την μεγαλύτερη αποδοχή της.

## 2.15.2 LOOM

Η LOOM είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού αλλά και ένα περιβάλλον ανάπτυξης για τη δημιουργία ευφυών εφαρμογών. Η LOOM χρησιμοποιείται για τη δημιουργία και συντήρηση εφαρμογών βασισμένων σε μοντέλα. Η LOOM επιτρέπει τον προσδιορισμό μοντέλων ενώ διαθέτει επίσης μηχανισμούς που μπορούν να χρησιμοποιηθούν για να χειριστούν και να αναζητήσουν πληροφορίες στα μοντέλα της. Αναπτύχθηκε το 1986, από το Πανεπιστήμιο της Νότιας Καλιφόρνια.

Η LOOM θεωρείται μοναδική γλώσσα από διάφορες απόψεις, αφού υποστηρίζει γλώσσα περιγραφής μοντέλων, αντικειμένων και σχέσεων, υποστηρίζει επίσης διαδικαστικό προγραμματισμό (Procedural Programming), χρησιμοποιώντας μεθόδους που είναι βασισμένες σε πρότυπα (Pattern-directed), ενώ διαθέτει δυνατότητες εξαγωγής πληροφορίας ενσωματωμένες στους μηχανισμούς παραγωγής και ταξινόμησης, που υλοποιούν παραγωγικό συλλογισμό (deductive reasoning).

## 2.15.3 KIF (Knowledge Interchange Format)

Αποτελεί μια γλώσσα για την ανταλλαγή γνώσης μεταξύ διαφορετικών εφαρμογών, αναπτύχθηκε για να λύσει το πρόβλημα της ανομοιογένειας μεταξύ γλωσσών που χρησιμοποιούνται για αναπαράσταση γνώσης. Περιλαμβάνει επεξηγηματικές δηλώσεις (που σημαίνει ότι είναι δυνατόν να «καταλάβει» το νόημα των εκφράσεων που δηλώνονται χωρίς να χρειάζεται διερμηνέα για να τις χειριστεί). Επιτρέπει τον ορισμό αντικειμένων, συναρτήσεων και σχέσεων. Έχει προταθεί ως πρότυπο για την περιγραφική λογική.

Η KIF αποτελεί μια συγκεκριμένη γλώσσα λογικής η οποία προτάθηκε σαν πρότυπο για να περιγράψει αντικείμενα σε ευφυείς εφαρμογές. Επιπλέον, έχει ειδικά σχεδιαστεί για να χρησιμοποιηθεί ως interlingua, μια γλώσσα μεσάζουσα στη διαδικασία μεταφράσεων από άλλες γλώσσες.

Η περιγραφή της KIF περιλαμβάνει τις προδιαγραφές για την σύνταξη της και τις έννοιες της. Με την KIF είναι δυνατόν να οριστούν απλά δεδομένα, όπως για παράδειγμα οι

παρακάτω εκφράσεις, που αποτελούν δείγμα της βάσης προσωπικού (όπου έχουμε το ID number, τμήμα και μισθό):

(salary 015-46-3946 widgets 72000)

(salary 026-40-9152 grommets 36000)

(salary 415-32-4707 fidgets 42000)

Η KIF περιλαμβάνει μια σειρά από λογικούς τελεστές που χρησιμοποιούνται για την κωδικοποίηση λογικής πληροφορίας ( όπως άρνηση, κανόνες, ποσοτικοί τύποι, κλπ.). Παράδειγμα περισσότερο πολύπλοκης έκφρασης αποτελεί η παρακάτω, η οποία δηλώνει πως ένα chip είναι μεγαλύτερο από κάποιο άλλο:

(> (\* (width chip1) (length chip1))

(\* (width chip2) (length chip2)))

Ακόμα, η KIF επιτρέπει την κωδικοποίηση γνώσης για τη γνώση, χρησιμοποιώντας κατάλληλους τελεστές (όπως (^) και (,)).

Στην έκφραση "(interested Joe `(salary ,?x ,?y ,?z))" δηλώνεται ότι ο Joe ενδιαφέρεται για τα instances της φράσης (τη γνώση που αναπαρίσταται από την δομή) (salary ?x ?y ?z) και όχι για την ίδια την φράση κατά γράμμα. Τέλος, η KIF μπορεί να χρησιμοποιηθεί για να περιγράψει διαδικασίες,

#### **2.15.4. OWL (Ontology Web Language)**

Η OWL [25] δημιουργήθηκε για να ικανοποιήσει την ανάγκη για μια γλώσσα οντολογίας Ιστού και αποτελεί μια από τις προτάσεις του W3C που είναι σχετικές με το Semantic Web. Η OWL έχει σχεδιαστεί για τις εφαρμογές που πρέπει να επεξεργαστούν το περιεχόμενο των πληροφοριών αντί απλά να παρουσιάσουν τις πληροφορίες στους ανθρώπους. Σύμφωνα με το W3C, ο σκοπός της OWL είναι να παρασχεθεί ένα τυποποιημένο σχήμα που είναι συμβατό με την αρχιτεκτονική του World Wide Web και του Semantic Web.[ Η τυποποίηση των οντολογιών σε γλώσσα OWL θα κάνει τα δεδομένα στο Web περισσότερο επεξεργάσιμα από μηχανές και επαναχρησιμοποιή-

σιμα στις εφαρμογές. Έτσι, η επεκτασιμότητα, η μετατρεψιμότητα και η διαλειτουργικότητα κατέχουν υψηλή προτεραιότητα στο σχεδιασμό της γλώσσας. Συγχρόνως, η OWL προσπαθεί να επιτύχει μια καλή ανταλλαγή μεταξύ της εξελισιμότητας και της εκφραστικής δύναμης.

Μπορεί να χρησιμοποιηθεί για να αντιπροσωπεύσει σαφώς την έννοια όρων σε λεξιλόγια και τις σχέσεις μεταξύ αυτών των όρων. Αυτή η αναπαράσταση των όρων και των αλληλεξαρτήσεών τους καλείται οντολογία. Υπερβαίνει τις γλώσσες XML, XML Schema, RDF και RDF Schema αφού υποστηρίζει μεγαλύτερη διασαφήνιση του περιεχομένου του Ιστού από τις μηχανές, παρέχοντας πρόσθετο λεξιλόγιο μαζί με μια επίσημη σημασιολογία. Πιο συγκεκριμένα:

- Η XML παρέχει μια σύνταξη επιφάνειας για τα δομημένα έγγραφα, αλλά δεν επιβάλλει κανέναν σημασιολογικό περιορισμό στην έννοια αυτών των εγγράφων.
- Το XML Schema είναι μια γλώσσα για τον περιορισμό της δομής των εγγράφων XML και επεκτείνει επίσης την XML με τα datatypes.
- Η RDF είναι ένα μοντέλο δεδομένων για τα αντικείμενα και τις μεταξύ τους σχέσεις, παρέχει μια απλή σημασιολογία για αυτό το μοντέλο και αυτά τα μοντέλα μπορούν να αναπαρασταθούν σε μια σύνταξη XML.
- Το RDF Schema είναι ένα λεξιλόγιο για την περιγραφή των ιδιοτήτων και των κλάσεων RDF αντικειμένων, με μια σημασιολογία για τις γενικευμένες ιεραρχίες τέτοιων ιδιοτήτων και κλάσεων.
- Η OWL προσθέτει περισσότερο λεξιλόγιο για την περιγραφή των ιδιοτήτων και των κλάσεων, όπως για παράδειγμα: σχέσεις μεταξύ των κλάσεων (π.χ. disjointness), αριθμός στοιχείων συνόλου (π.χ. "ακριβώς ένας"), λογικοί συνδυασμοί κλάσεων (π.χ. ένωση, τομή, συμπλήρωμα), ισότητα, πλουσιότεροι τύποι ιδιοτήτων, χαρακτηριστικά ιδιοτήτων (π.χ. συμμετρία, μοναδικότητα, μεταβατικότητα), και απαριθμημένες κλάσεις.

Ένα πλεονέκτημα των OWL οντολογιών θα είναι η διαθεσιμότητα των εργαλείων που μπορούν να εξαγάγουν συλλογισμούς για αυτές. Τα εργαλεία θα παράσχουν γενική υποστήριξη που δεν είναι συγκεκριμένη για μία ιδιαίτερη περιοχή, το οποίο θα συνέβαινε

εάν επρόκειτο να δημιουργηθεί ένα σύστημα για την εξαγωγή συλλογισμών για ένα συγκεκριμένο βιομηχανικό πρότυπο XML σχήματος. Η οικοδόμηση ενός υγιούς και χρήσιμου συστήματος συλλογισμού δεν είναι μια απλή προσπάθεια. Η κατασκευή μιας οντολογίας είναι πολύ πιο βολική.

Η γλώσσα οντολογίας Ιστού OWL είναι μια σημασιολογική γλώσσα σήμανσης για την έκδοση και τη διανομή των οντολογιών στο World Wide Web. Αναπτύσσεται ως επέκταση του λεξιλογίου RDF, προέρχεται από τη γλώσσα οντολογίας Ιστού DAML+OIL και είναι γραμμένη σε XML ώστε να μπορεί εύκολα να είναι ανεξάρτητη του λειτουργικού συστήματος και των γλωσσών εφαρμογής ενός υπολογιστή. Επιπλέον, το Semantic Web, χρησιμοποιώντας οντολογίες βασισμένες στην XML κι όχι παραδοσιακές δομές δέντρων, θα καταστήσει ευκολότερη για τις εφαρμογές υπηρεσιών Web την επεξεργασία στοιχείων από πολλαπλές πηγές. Η γλώσσα παρέχει έναν τυποποιημένο τρόπο να καθοριστούν οι βασισμένες στο Web οντολογίες έτσι ώστε τα στοιχεία να μπορούν να περιγραφούν ως αυτό που είναι, αντί ως έγγραφο σε μια δομή δέντρων ή άλλη αφαίρεση βάσεων δεδομένων.

Μια OWL οντολογία μπορεί να περιλάβει τις περιγραφές των κλάσεων, των ιδιοτήτων και των στιγμιοτύπων τους. Λαμβάνοντας υπόψη μια τέτοια οντολογία, η επίσημη OWL σημασιολογία διευκρινίζει πώς να παραγάγει τα λογικά συμπεράσματα της, δηλαδή γεγονότα όχι κυριολεκτικά παρόντα στην οντολογία, αλλά συνεπαγόμενα από τη σημασιολογία. Αυτές οι συνεπαγωγές μπορούν να βασιστούν σε ένα ενιαίο έγγραφο ή σε πολλαπλά κατανεμημένα έγγραφα που έχουν συνδυαστεί χρησιμοποιώντας καθορισμένους OWL μηχανισμούς.

Η OWL είναι ένα συστατικό του Semantic Web που στοχεύει να καταστήσει τους Web πόρους ευκολότερα προσιτούς στις αυτοματοποιημένες διαδικασίες, με την προσθήκη πληροφοριών για τους πόρους που περιγράφουν ή παρέχουν Web περιεχόμενο. Δεδομένου ότι το Semantic Web είναι κατανεμημένο, η OWL πρέπει να επιτρέψει να συγκεντρώνονται πληροφορίες από κατανεμημένες πηγές. Αυτό γίνεται εν μέρει επιτρέποντας στις οντολογίες να είναι συσχετισμένες, συμπεριλαμβάνοντας ρητά την εισαγωγή πληροφοριών από άλλες οντολογίες.

Επιπλέον, η OWL κάνει μια ανοικτή παγκόσμια υπόθεση. Δηλαδή, οι περιγραφές των πόρων δεν είναι περιορισμένες σε ένα ενιαίο αρχείο ή πεδίο. Ενώ η κλάση C1 μπορεί να καθοριστεί αρχικά στην οντολογία O1, μπορεί να επεκταθεί σε άλλες οντολογίες. Οι συνέπειες αυτών των πρόσθετων προτάσεων για τη C1 είναι μονοτονικές. Οι νέες πληροφορίες δεν μπορούν να αποσύρουν τις προηγούμενες πληροφορίες. Οι νέες πληροφορίες μπορούν να είναι αντιφατικές, αλλά τα γεγονότα και οι συνεπαγωγές μπορούν μόνο να προστεθούν, ποτέ να διαγραφούν. Η δυνατότητα τέτοιων αντιφάσεων είναι κάτι που ο σχεδιαστής μιας οντολογίας πρέπει να λάβει υπόψη.

Οι πρακτικές εφαρμογές της OWL περιλαμβάνουν τα Web portals, όπου μπορεί να χρησιμοποιηθεί για να δημιουργήσει τους κανόνες κατηγοριοποίησης προκειμένου να βελτιωθεί η αναζήτηση, τις συλλογές πολυμέσων, όπου μπορεί να χρησιμοποιηθεί για να επιτρέψει βασισμένες σε περιεχόμενο αναζητήσεις των μέσων και τις υπηρεσίες Ιστού (Web services), όπου μπορεί να χρησιμοποιηθεί για την ανακάλυψη και τη σύνθεση υπηρεσιών Ιστού καθώς επίσης και για τη διαχείριση δικαιωμάτων και τον έλεγχο πρόσβασης.

### **Οι υπο-γλώσσες της OWL**

Η OWL έχει τρεις όλο και περισσότερο-εκφραστικές υπο-γλώσσες: OWL Lite, OWL DL, και OWL Full.

- Η *OWL Lite* υποστηρίζει εκείνους τους χρήστες που χρειάζονται κυρίως μια ιεραρχία ταξινόμησης και απλούς περιορισμούς. Χρησιμοποιεί μόνο μερικά από τα γλωσσικά χαρακτηριστικά της OWL και έχει περισσότερους περιορισμούς στη χρήση των χαρακτηριστικών αυτών από την OWL DL ή την OWL Full. Παραδείγματος χάριν, στην OWL Lite οι κλάσεις μπορούν μόνο να οριστούν σε σχέση με κατονομασμένες (named) υπερκλάσεις (οι υπερκλάσεις δε μπορεί να είναι αυθαίρετες εκφράσεις), και μόνο ορισμένα είδη περιορισμών κλάσεων μπορούν να χρησιμοποιηθούν. Η ισοδυναμία μεταξύ κλάσεων και σχέσεων υποκλάσεων επιτρέπεται μόνο μεταξύ named κλάσεων, και όχι μεταξύ αυθαίρετων εκφράσεων κλάσης. Όμοια, οι περιορισμοί στην OWL Lite χρησιμοποιούν μόνο named κλάσεις. Η OWL Lite έχει επίσης μια περιορισμένη

έννοια του αριθμού στοιχείων συνόλου - τα μόνα cardinalities που επιτρέπονται για να δηλωθούν ρητά είναι 0 ή 1. Η OWL Lite έχει χαμηλότερη πολυπλοκότητα από την OWL DL. Η OWL Lite έχει την ιδιότητα να μπορεί να γίνεται αντικείμενο αποδοτικής επεξεργασίας από αλγόριθμους εξαγωγής συμπερασμάτων (inference algorithms).

- Η *OWL DL* υποστηρίζει εκείνους τους χρήστες που θέλουν τη μέγιστη εκφραστικότητα διατηρώντας την υπολογιστική πληρότητα (όλα τα συμπεράσματα είναι εγγυημένα να είναι υπολογίσιμα) και decidability (όλοι οι υπολογισμοί θα τελειώσουν σε πεπερασμένο χρόνο). Η OWL DL περιλαμβάνει όλα τα OWL γλωσσικά κατασκευάσματα, αλλά μπορούν να χρησιμοποιηθούν μόνο κάτω από ορισμένους περιορισμούς (παραδείγματος χάριν, ενώ μια κλάση μπορεί να είναι υποκλάση πολλών κλάσεων, μια κλάση δεν μπορεί να είναι αντικείμενο instance μιας άλλης κλάσης). Η OWL DL ονομάζεται έτσι λόγω της αντιστοιχίας της με την περιγραφική λογική, έναν τομέα της έρευνας που έχει μελετήσει τη λογική που αποτελεί τη βάση της OWL. Η OWL DL και η OWL Full χρησιμοποιούν το ίδιο λεξιλόγιο αν και η OWL DL υπόκειται σε μερικούς περιορισμούς. Κατά προσέγγιση, η OWL DL απαιτεί χωρισμό τύπων (μια κλάση δεν μπορεί να είναι επίσης άτομο ή ιδιότητα, μια ιδιότητα δεν μπορεί να είναι επίσης άτομο ή κλάση). Αυτό υπονοεί ότι οι περιορισμοί δεν μπορούν να εφαρμοστούν στα γλωσσικά στοιχεία της OWL DL (κάτι που επιτρέπεται στην OWL Full). Επιπλέον, η OWL DL απαιτεί οι ιδιότητες να είναι είτε ObjectProperties είτε DatatypeProperties: Οι DatatypeProperties είναι σχέσεις μεταξύ instances κλάσεων και RDF κυριολεκτικών (literals) και XML Schema τύπων δεδομένων, ενώ οι ObjectProperties είναι σχέσεις μεταξύ instances δύο κλάσεων. Η OWL DL έχει την ιδιότητα να μπορεί να γίνεται αντικείμενο αποδοτικής επεξεργασίας από έναν description logic reasoner.
- Η *OWL Full* απευθύνεται στους χρήστες που θέλουν τη μέγιστη εκφραστικότητα και τη συντακτική ελευθερία της RDF χωρίς υπολογιστικές εγγυήσεις. Παραδείγματος χάριν, στην OWL Full μια κλάση μπορεί να αντιμετωπιστεί ταυτόχρονα ως συλλογή ατόμων και ως άτομο σαν αποτέλεσμα ιδιοτήτων. Η OWL Full επιτρέπει σε μια οντολογία να αυξήσει την έννοια του προκαθορισμένου



(RDF ή OWL) λεξιλογίου. Είναι απίθανο ότι οποιοδήποτε λογισμικό συλλογισμού θα είναι σε θέση να υποστηρίξει τον πλήρη συλλογισμό για κάθε χαρακτηριστικό γνώρισμα της OWL Full.

Κάθε μια από αυτές τις υπο-γλώσσες είναι μια επέκταση του απλούστερου προκατόχου της, όσον αφορά αυτό που μπορεί να εκφραστεί έγκυρα αλλά και αυτό που μπορεί να προκύψει εγκύρως ως συμπέρασμα. Ισχύουν οι παρακάτω σχέσεις:

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

Η επιλογή της κατάλληλης υπο-γλώσσας της OWL για την ανάπτυξη της οντολογίας εξαρτάται από τις ανάγκες της. Η επιλογή μεταξύ της OWL Lite και της OWL DL εξαρτάται από το βαθμό στον οποίο οι χρήστες απαιτούν τα περισσότερο-εκφραστικά κατασκευάσματα που παρέχονται από την OWL DL. Η επιλογή μεταξύ της OWL DL και της OWL Full εξαρτάται κυρίως από το βαθμό στον οποίο οι χρήστες απαιτούν τις ευκολίες μετα-διαμόρφωσης του σχήματος RDF (π.χ. καθορισμός κλάσεων άλλων κλάσεων, ή σύνδεση των ιδιοτήτων με τις κλάσεις). Κατά χρησιμοποίηση της OWL Full σε σύγκριση με την OWL DL, η υποστήριξη συλλογισμού είναι λιγότερο προβλέψιμη αφού πλήρεις OWL Full εφαρμογές δεν υπάρχουν.

Η OWL Full μπορεί να αντιμετωπισθεί ως επέκταση της RDF, ενώ η OWL Lite και η OWL DL μπορούν να αντιμετωπισθούν ως επεκτάσεις μιας περιορισμένης άποψης της RDF. Κάθε OWL έγγραφο (Lite, DL, Full) είναι ένα έγγραφο RDF, και κάθε έγγραφο RDF είναι ένα OWL Full έγγραφο, αλλά μόνο μερικά έγγραφα RDF θα είναι ένα νόμιμο έγγραφο OWL Lite ή OWL DL. Λόγω αυτού, κάποια προσοχή πρέπει να ληφθεί όταν θέλει ένας χρήστης να μετατρέψει ένα έγγραφο RDF σε OWL. Όταν η εκφραστικότητα της OWL DL ή της OWL Lite κρίνεται κατάλληλη, μερικές προφυλάξεις πρέπει να ληφθούν για να εξασφαλίσουν ότι το αρχικό έγγραφο RDF συμμορφώνεται με τους πρόσθετους περιορισμούς που επιβάλλονται από την OWL DL και την OWL Lite. Μεταξύ των άλλων, κάθε URI που χρησιμοποιείται όπως ένα όνομα κλάσης πρέπει να δηλωθεί σαφώς ότι είναι του τύπου `owl:Class` (και ομοίως για τις ιδιότητες), κάθε άτομο

πρέπει να δηλωθεί ότι ανήκει τουλάχιστον σε μια κλάση (ακόμα και μόνο στην owl:Thing), τα URI που χρησιμοποιούνται για τις κλάσεις, ιδιότητες και άτομα πρέπει να είναι διαχωρισμένα.

### **2.15.5 RDF(Resource Description Framework)**

Ο Σημασιολογικός Ιστός (Semantic Web)[[3] είναι (όπως αναφέρθηκε και παραπάνω) το επόμενο σημαντικό βήμα του Παγκόσμιου Ιστού. Εφαρμογές του Σημασιολογικού Ιστού όπως είναι οι πύλες γνώσης (Knowledge portals) και οι ηλεκτρονικές αγορές (E-Marketplaces) απαιτούν τη διαχείριση τεράστιων όγκων μεταδεδομένων, δηλαδή πληροφορίας που περιγράφει τα δεδομένα. Έτσι, με τις εφαρμογές αυτές τεράστιες ποσότητες πόρων ιστού (web resources), όπως δεδομένα, έγγραφα, προγράμματα, θα είναι διαθέσιμες μαζί με διαφόρων ειδών μεταπληροφορία. Η καλύτερη γνώση για το νόημα, τη χρησιμότητα, την προσπελασιμότητα ή την ποιότητα των πηγών αυτών θα διευκολύνει την αυτόματη επεξεργασία των διαθέσιμων υπηρεσιών στον Ιστό.

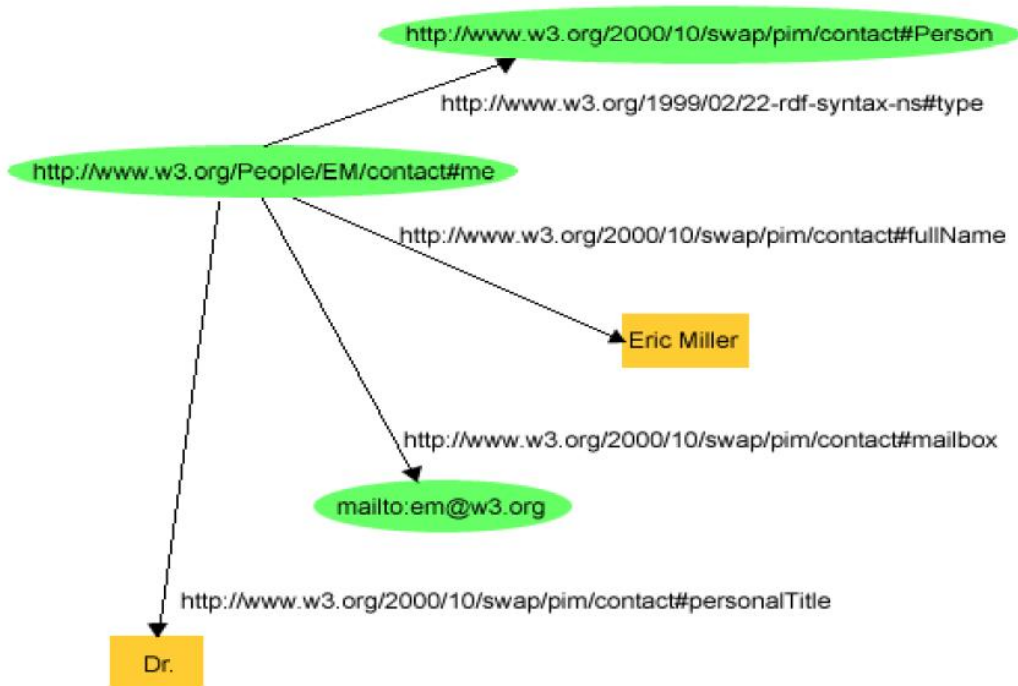
Το πλαίσιο RDF[2] (Resource Description Framework) επιτρέπει τη δημιουργία και την ανταλλαγή μεταδεδομένων των πόρων όπως και οποιαδήποτε άλλα δεδομένα στον Ιστό. Είναι ιδιαίτερα προοριζόμενο για την αντιπροσώπευση μεταδεδομένων για τους πόρους του Ιστού, όπως τίτλος, συντάκτης και ημερομηνία τροποποίησης μιας ιστοσελίδας, πνευματικά δικαιώματα και πληροφορία χορήγησης αδειών για ένα έγγραφο του Ιστού, ή πρόγραμμα διαθεσιμότητας για κάποιο κοινό πόρο. Παρόλα αυτά, με τη γενίκευση της έννοιας "πόρος Ιστού", το RDF μπορεί επίσης να χρησιμοποιηθεί για να αντιπροσωπεύσει πληροφορίες για πράγματα που μπορούν να προσδιοριστούν στον Ιστό ακόμα και όταν δεν μπορούν να ανακτηθούν άμεσα από αυτόν. Παραδείγματα περιλαμβάνουν πληροφορίες για αντικείμενα διαθέσιμα από ηλεκτρονικές αγορές (πχ., πληροφορίες για τις προδιαγραφές, τις τιμές και τη διαθεσιμότητα) ή περιγραφή των προτιμήσεων ενός χρήστη του Ιστού για την παράδοση πληροφορίας.

Το μοντέλο RDF στοχεύει σε καταστάσεις στις οποίες αυτή η πληροφορία πρέπει να υποβληθεί σε επεξεργασία από εφαρμογές και όχι μόνο για να επιδειχθεί στους

ανθρώπους. Επίσης, παρέχει ένα κοινό πλαίσιο για την πληροφορία αυτή έτσι ώστε να μπορεί να ανταλλαχθεί μεταξύ των εφαρμογών χωρίς απώλεια της σημασίας. Δεδομένου αυτού, οι σχεδιαστές εφαρμογών μπορούν να εκμεταλλευτούν τη διαθεσιμότητα κοινών RDF parsers και εργαλείων επεξεργασίας. Η δυνατότητα να ανταλλαχθεί πληροφορία μεταξύ διαφορετικών εφαρμογών σημαίνει ότι πληροφορία μπορεί να τεθεί στη διάθεση και άλλων εφαρμογών εκτός από εκείνες για τις οποίες δημιουργήθηκε αρχικά.

Το RDF είναι βασισμένο στην ιδέα του προσδιορισμού των πραγμάτων χρησιμοποιώντας προσδιοριστικά του Ιστού (web identifiers, αποκαλούμενα Uniform Resource Identifiers ή URIs) και της περιγραφής των πόρων από την άποψη απλών ιδιοτήτων και τιμών των ιδιοτήτων αυτών (properties and property values). Αυτό επιτρέπει στα RDF να αντιπροσωπεύουν απλές δηλώσεις για πόρους ως ένα γράφο από κόμβους και τόξα που αντιπροσωπεύουν τους πόρους, τις ιδιότητες και τις τιμές τους. Για να είμαστε πιο συγκεκριμένοι, η φράση «υπάρχει ένας άνθρωπος προσδιορισμένος από το <http://www.w3.org/People/EM/contact#me> του οποίου το όνομα είναι Eric Miller, η ηλεκτρονική διεύθυνση [em@w3.org](mailto:em@w3.org) και ο τίτλος του είναι "Dr" » μπορεί να παρουσιαστεί σε ένα γράφο RDF όπως φαίνεται στο Σχήμα 2.15.5.1

Το σχήμα αυτό δείχνει ότι το RDF χρησιμοποιεί URIs για να προσδιορίσει άτομα όπως ο Eric Miller που προσδιορίζεται από το <http://www.w3.org/People/EM/contact#me>, είδη πραγμάτων, π.χ. Person που προσδιορίζεται από <http://www.w3.org/2000/10/swap/pim/contact#Person>, ιδιότητες εκείνων των πραγμάτων, π.χ. ταχυδρομική θυρίδα που προσδιορίζεται από <http://www.w3.org/2000/10/swap/pim/contact#mailbox>, τιμές εκείνων των ιδιοτήτων, π.χ. <mailto:em@w3.org> ως τιμή της ιδιότητας ταχυδρομικής θυρίδας (το RDF χρησιμοποιεί επίσης συμβολοσειρές όπως "Eric Miller" και τιμές από άλλα τύπους δεδομένων (datatypes) όπως ακέραιοι αριθμοί και ημερομηνίες, ως τιμές των ιδιοτήτων).



### Σχήμα 2.15.5.1

Πηγή: RDF Primer, W3C Recommendation 10 February 2007, <http://www.w3.org/TR/rdf-primer>[22]

Επίσης, το RDF παρέχει μια σύνταξη σε XML μορφή (RDF/XML) για την καταγραφή και την ανταλλαγή τέτοιων γράφων. Το παρακάτω παράδειγμα δείχνει το παραπάνω σχήμα στη μορφή αυτή:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:contact="http://www.w3.org/2000/10/swap/pim/contact#">
  <contact:Person rdf:about="http://www.w3.org/People/EM/contact#me">
    <contact:fullName>Eric Miller</contact:fullName>
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
    <contact:personalTitle>Dr.</contact:personalTitle>
  </contact:Person>
```

</rdf:RDF>

Όπως στην HTML, η RDF/XML είναι μηχανικά επεξεργάσιμη και χρησιμοποιώντας URIs μπορεί να συνδέσει κομμάτια πληροφοριών στον Ιστό. Εντούτοις, αντίθετα από το συμβατικό υπερκείμενο (hypertext), τα URIs μπορούν να αναφερθούν σε οποιοδήποτε ευπροσδιόριστο πράγμα, συμπεριλαμβανομένων και εκείνων που μπορεί να μην είναι άμεσα ανακτήσιμα από τον Ιστό (όπως το πρόσωπο Eric Miller). Το αποτέλεσμα είναι ότι εκτός από την περιγραφή πραγμάτων όπως ιστοσελίδων, τα RDF μπορούν επίσης να περιγράψουν αυτοκίνητα, επιχειρήσεις, ανθρώπους, γεγονότα ειδήσεων κλπ. Επιπλέον, οι ιδιότητες στα RDF έχουν URIs, για να προσδιορίσουν ακριβώς τις σχέσεις που υπάρχουν μεταξύ των συνδεδεμένων στοιχείων.

Γενικότερα, το RDF είναι βασισμένο στην ιδέα ότι τα πράγματα που περιγράφονται έχουν ιδιότητες (properties) που έχουν τιμές και ότι οι πόροι μπορούν να περιγραφούν με εκφράσεις οι οποίες διευκρινίζουν τις ιδιότητες αυτές και τις τιμές τους. Το RDF χρησιμοποιεί μια ιδιαίτερη ορολογία για τα διάφορα μέρη των εκφράσεων. Συγκεκριμένα, το μέρος που προσδιορίζει το πράγμα για το οποίο είναι η πρόταση (μια ιστοσελίδα για παράδειγμα) καλείται υποκείμενο (subject). Το μέρος που προσδιορίζει την ιδιότητα ή το χαρακτηριστικό του υποκειμένου που η πρόταση διευκρινίζει (δημιουργός, δημιουργία-ημερομηνία, ή γλώσσα σε αυτά τα παραδείγματα) καλείται κατηγορημα (predicate) και το μέρος που προσδιορίζει την τιμή εκείνης της ιδιότητας καλείται αντικείμενο (object). Έτσι, παίρνοντας την αγγλική πρόταση: <http://www.example.org/index.html> has a creator whose value is John Smith θα έχουμε:

subject: <http://www.example.org/index.html>

predicate: <http://purl.org/dc/elements/1.1/creator>

object: <http://www.example.org/staffid/85740>

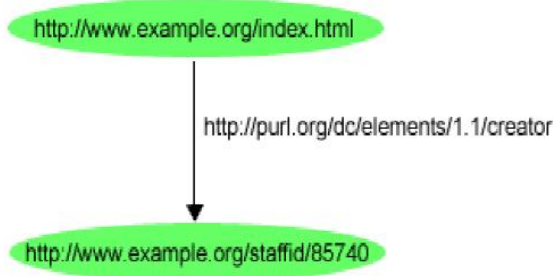
Το RDF μοντελοποιεί τέτοιες εκφράσεις ως κόμβους και τόξα σε ένα γράφο. Με αυτήν την μοντελοποίηση μια πρόταση αντιπροσωπεύεται από:

ένα κόμβο (node) για το υποκείμενο

ένα κόμβο (node) για το αντικείμενο

ένα τόξο (arc) για το κατηγορημα κατευθυνόμενο από το υποκείμενο στο αντικείμενο

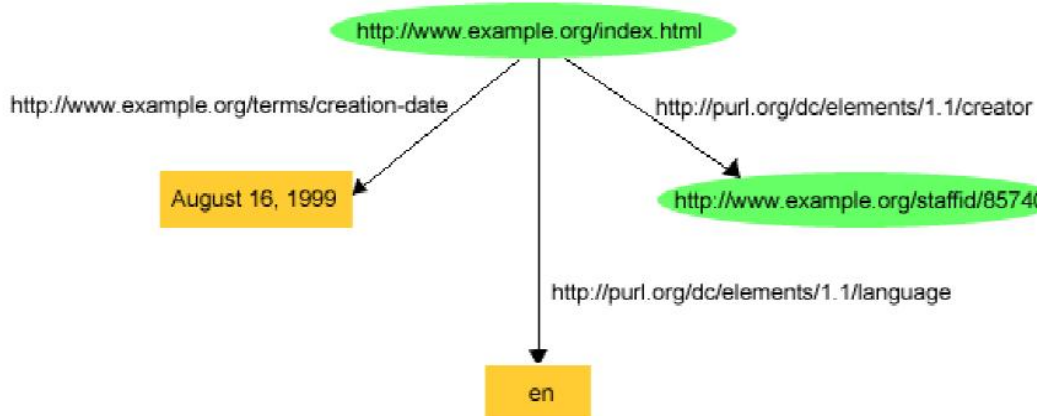
Έτσι, η παραπάνω πρόταση θα μπορούσε να αναπαρασταθεί με τον γράφο που φαίνεται στο Σχήμα 2.15.5.2



**Σχήμα 2.15.5.2**

[Πηγή: RDF Primer, W3C Recommendation 10 February 2007, http://www.w3.org/TR/rdf-primer \[22\]](#)

Στο παράδειγμα του Σχήματος 2.15.5.3 μπορούμε να διαπιστώσουμε ότι τα αντικείμενα του RDF μπορούν να είναι είτε URIs είτε σταθερές τιμές που λέγονται κυριολεκτικά (literals) και αναπαριστούνται από συμβολοσειρές προκειμένου να αντιπροσω-πεύσουν ορισμένα είδη ιδιοτήτων. Τα κυριολεκτικά δεν μπορούν να χρησιμο-ποιηθούν ως υποκείμενα ή ως κατηγορήματα σε RDF εκφράσεις.



**Σχήμα 2.15.5.3** [Πηγή: RDF Primer, W3C Recommendation 10 February 2007, http://www.w3.org/TR/rdf-primer \[22\]](#)

Ένας εναλλακτικός τρόπος να αναπαραστήσουμε ένα RDF είναι γράφοντας κάθε έκφραση που προκύπτει, δηλαδή σε μορφή triples. Έτσι, το παραπάνω σχήμα θα μπορούσε να γραφεί ως εξής:

```
<http://www.example.org/index.html>  
<http://purl.org/dc/elements/1.1/creator>  
<http://www.example.org/staffid/85740>.  
  
<http://www.example.org/index.html>  
<http://www.example.org/terms/creation-date> "August 16, 1999".  
  
<http://www.example.org/index.html>  
<http://purl.org/dc/elements/1.1/language> "en".
```

Για μεγαλύτερη ευκολία, αντί να γράφονται πλήρως τα URIs το RDF δίνει τη δυνατότητα για χρήση των namespaces. Έτσι, έχοντας ορίσει τα namespaces

```
ex: http://www.example.org/  
dc: http://purl.org/dc/elements/1.1/  
exterms: http://www.example.org/terms/  
exstaff: http://www.example.org/staffid/
```

Το παράδειγμα αυτό θα μπορούσε να γραφτεί ως εξής:

```
ex:index.html dc:creator exstaff:85740.  
ex:index.html exterms:creation-date "August 16, 1999".  
ex:index.html dc:language "en".
```

Τέλος, το παραπάνω RDF θα μπορούσε να γραφτεί σε μορφή RDF/XML, όπως προαναφέραμε, ως εξής:

```
<?xml version="1.0"?>  
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"  
xmlns:exterms="http://www.example.org/terms/">
```

```
<rdf:Description rdf:about="http://www.example.org/index.html">
    <exterms:creation-date>August 16, 1999</exterms:creation-date>
</rdf:Description>
</rdf:RDF>
```

## 2.15.6 RDFS

Το πλαίσιο RDF παρέχει έναν τρόπο να εκφραστούν απλές εκφράσεις για πόρους, χρησιμοποιώντας ονομασμένες ιδιότητες και τιμές. Παρόλα αυτά, οι κοινότητες χρηστών RDF χρειάζονται επίσης τη δυνατότητα να καθορίσουν το λεξιλόγιο (όρους) που αυτοί σκοπεύουν να χρησιμοποιήσουν σε εκείνες τις εκφράσεις, και συγκεκριμένα, να δείξουν ότι περιγράφουν συγκεκριμένα είδη ή κατηγορίες πόρων και γι' αυτό θα χρησιμοποιήσουν συγκεκριμένες ιδιότητες στην περιγραφή εκείνων των πόρων. Για παράδειγμα, άνθρωποι ενδιαφερόμενοι για την περιγραφή βιβλιογραφικών πόρων θα ήθελαν να περιγράψουν κατηγορίες όπως `ex:Book` ή `ex:MagazineArticle` και να χρησιμοποιήσουν ιδιότητες όπως `ex:author`, `ex:title` και `ex:subject` για να τις περιγράψουν. Το ίδιο το RDF δεν παρέχει κανένα μέσο για τον προσδιορισμό κλάσεων και ιδιοτήτων. Αντί αυτού, τέτοιες κλάσεις και ιδιότητες περιγράφονται ως ένας RDF λεξιλόγιο χρησιμοποιώντας επεκτάσεις του RDF που είναι το RDF Schema ή RDFS.[14]

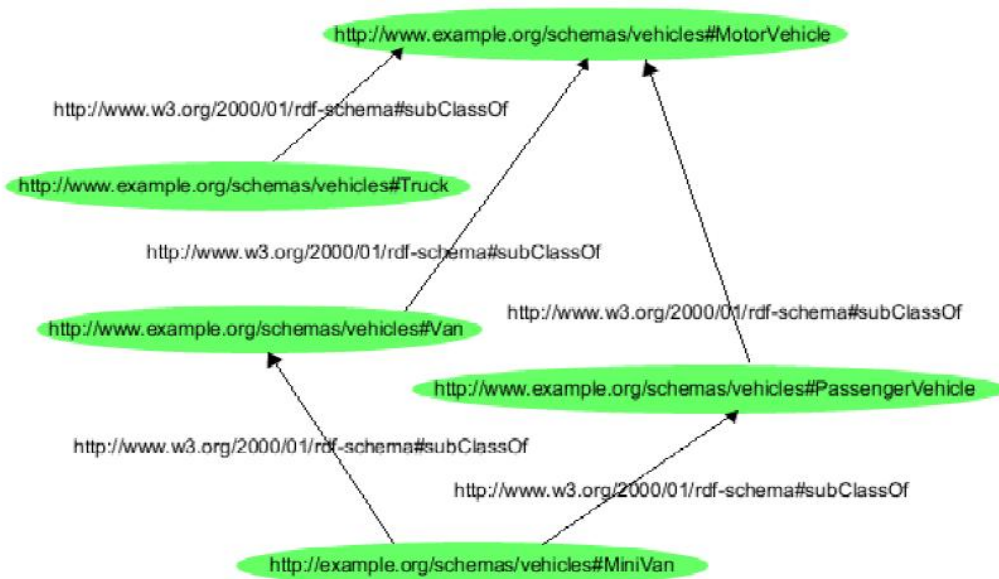
Το RDFS παρέχει τις προϋποθέσεις που απαιτούνται για να περιγράψει κανείς τέτοιες κλάσεις και ιδιότητες και για να προσδιορίσει ποιες κλάσεις και ιδιότητες αναμένονται να χρησιμοποιηθούν μαζί. Με άλλα λόγια, το RDF σχήμα παρέχει ένα σύστημα τύπων για RDF. Το σύστημα αυτό για το RDFS είναι παρόμοιο κατά κάποιον τρόπο με τα συστήματα τύπων των αντικειμενοστρεφών γλωσσών προγραμματισμού όπως η Java. Παραδείγματος χάριν, το RDF σχήμα επιτρέπει στους πόρους να είναι ορισμένοι ως στιγμιότυπα μιας ή περισσότερων κλάσεων. Επιπλέον, επιτρέπει στις κλάσεις να είναι οργανωμένες με έναν ιεραρχικό τρόπο. Για παράδειγμα, μια κλάση `ex:Dog` μπορεί να οριστεί ως υποκλάση της `ex:Mammal` που είναι υποκλάση της `ex:Animal`, σημαίνοντας ότι οποιοσδήποτε πόρος που είναι στιγμιότυπο της κλάσης `ex:Dog` είναι επίσης έμμεσα



στιγμιότυπο της κλάσης `ex:Animal`. Εντούτοις, οι κλάσεις και οι ιδιότητες του RDF είναι κατά κάποιον τρόπο πολύ διαφορετικές από τους τύπους των γλωσσών προγραμματισμού. Η περιγραφή των RDF κλάσεων και ιδιοτήτων δεν δημιουργούν ένα καλούπι στο οποίο οι πληροφορίες πρέπει να ανήκουν, αλλά παρέχουν πρόσθετη πληροφορία για τους πόρους RDF που περιγράφουν.

Ένα βασικό βήμα σε οποιοδήποτε είδος διαδικασίας περιγραφής είναι ο προσδιορισμός των διαφόρων πραγμάτων που περιγράφονται. Το RDF σχήμα αναφέρεται σε αυτά τα "πράγματα" ως κλάσεις. Μια κλάση (class) αντιστοιχεί στη γενική έννοια ενός τύπου ή μιας κατηγορίας, όπως την έννοια μιας κλάσης στις αντικειμενοστρεφείς γλώσσες προγραμματισμού. Οι κλάσεις μπορούν να χρησιμοποιηθούν για να αντιπροσωπεύσουν σχεδόν οποιαδήποτε κατηγορία πράγματος, όπως ιστοσελίδες, ανθρώπους, τύπους εγγράφων, βάσεις δεδομένων ή αφηρημένες έννοιες. Οι κλάσεις στο RDFS περιγράφονται χρησιμοποιώντας τους πόρους `rdfs:Class` και `rdfs:Resource` και τις ιδιότητες `rdf:type` και `rdfs:subClassOf`.

Ας πάρουμε για παράδειγμα μια εταιρία που έχει οχήματα (vehicles). Σε αυτό το παράδειγμα μπορούμε να ορίσουμε μια ιεραρχία όπως φαίνεται στο παρακάτω σχήμα:



### Σχήμα 2.15.6

Πηγή: RDF Primer, W3C Recommendation 10 February 2007, <http://www.w3.org/TR/rdf-primer>[22]

Αυτή η ιεραρχία μπορεί να γραφτεί σε RDF/XML μορφή με τη χρήση namespaces ως εξής:

```

<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd http://www.w3.org/2001/XMLSchema#>]>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

  <rdfs:Class rdf:ID="MotorVehicle"/>

  <rdfs:Class rdf:ID="PassengerVehicle">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Truck">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Van">
    <rdfs:subClassOf rdf:resource="#Truck"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="MiniVan">
    <rdfs:subClassOf rdf:resource="#Van"/>
    <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
  </rdfs:Class>
</rdf:RDF>
  
```

```
</rdfs:Class>
```

```
<rdfs:Class rdf:ID="Truck">
```

```
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
```

```
</rdfs:Class>
```

```
<rdfs:Class rdf:ID="Van">
```

```
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
```

```
</rdfs:Class>
```

```
<rdfs:Class rdf:ID="MiniVan">
```

```
  <rdfs:subClassOf rdf:resource="#Van"/>
```

```
<rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
```

```
</rdfs:Class>
```

```
</rdf:RDF>
```

Το RDF/XML στο παράδειγμα αυτό εισάγει ονόματα, όπως το `MotorVehicle`, για τους πόρους (κλάσεις) που περιγράφει χρησιμοποιώντας το `rdf:ID`. Αυτό είναι χρήσιμο εδώ επειδή συντομεύει το `URIref` και παρέχει έναν πρόσθετο έλεγχο ότι η τιμή του `rdf:ID` είναι μοναδική σε όλη την τρέχουσα βάση URI (συνήθως το έγγραφο URI). Αυτό βοηθά να πάρει τις επαναλαμβανόμενες τιμές `rdf:ID` κατά την καθορισμό των ονομάτων των κλάσεων και των ιδιοτήτων στα RDF σχήματα. Σχετικά `URIrefs` βασισμένα σε αυτά τα ονόματα μπορούν έπειτα να χρησιμοποιηθούν σε άλλους ορισμούς κλάσεων μέσα στο ίδιο σχήμα (πχ. όπως το `#MotorVehicle` χρησιμοποιείται στη περιγραφή των άλλων κλάσεων). Το πλήρες `URIref` αυτής της κλάσης, που υποθέτει ότι το ίδιο το σχήμα ήταν ο πόρος <http://example.org/schemas/vehicles>, θα ήταν <http://example.org/schemas/vehicles#MotorVehicle>. Για να εξασφαλιστεί ότι οι αναφορές σε αυτές τις κατηγορίες σχημάτων θα διατηρηθούν με συνέπεια ακόμα κι αν το σχήμα μετατοπιστεί ή αντιγραφτεί (ή για να ορίσει απλά μια βάση `URIref` για τις κλάσεις σχημάτων χωρίς να υποθέσουμε ότι όλες υπάρχουν σε μια μοναδική θέση), οι περιγραφές των κλάσεων μπορούν επίσης να περιλάβουν τη δήλωση `xml:base="http://example.org/schemas/vehicles"`.

Εκτός από την περιγραφή συγκεκριμένων κλάσεων πραγμάτων που θέλουν να περιγράψουν, οι κοινότητες χρηστών πρέπει επίσης να είναι σε θέση να περιγράψουν συγκεκριμένες ιδιότητες που χαρακτηρίζουν εκείνες τις κλάσεις (όπως rearSeatLegRoom για να περιγράψει ένα επιβατικό όχημα (passenger vehicle)). Στο RDF σχήμα, οι ιδιότητες (properties) περιγράφονται χρησιμοποιώντας την RDF κλάση `rdf:Property` και τις ιδιότητες σχημάτων RDF `rdfs:domain`, `rdfs:range`, και `rdfs:subPropertyOf`. Επίσης, παρέχει λεξιλόγιο για να περιγράψει πώς οι ιδιότητες και οι κλάσεις προορίζονται για να χρησιμοποιηθούν μαζί στα RDF δεδομένα. Η σημαντικότερη πληροφορία αυτού του είδους παρέχεται με τη χρησιμοποίηση των ιδιοτήτων `rdfs:range` και `rdfs:domain` για να περιγράψουν περαιτέρω τις ιδιότητες που έχουν οριστεί από εφαρμογή.

Το `rdfs:range` χρησιμοποιείται για να δείξει ότι οι τιμές μιας συγκεκριμένης ιδιότητας είναι στιγμιότυπα μιας οριζόμενης κλάσης. Παραδείγματος χάριν, εάν θέλαμε να δείξουμε ότι η ιδιότητα `ex:author` έχει τιμές που είναι στιγμιότυπα της κλάσης `ex:Person`, θα γράφαμε τα RDF triples:

<code>ex:Person</code>	<code>rdf:type</code>	<code>rdfs:Class</code> .
<code>ex:author</code>	<code>rdf:type</code>	<code>rdf:Property</code> .
<code>ex:author</code>	<code>rdfs:range</code>	<code>ex:Person</code> .

Μια ιδιότητα (property), για παράδειγμα το `ex:hasMother`, μπορεί να έχει ένα ή περισσότερα από ένα range properties. Εάν το `ex:hasMother` δεν έχει κανένα range τότε τίποτα δεν λέγεται για τις τιμές της ιδιότητας `ex:hasMother`. Εάν το `ex:hasMother` έχει ένα range, για παράδειγμα το `ex:Person`, αυτό λέει ότι οι τιμές της ιδιότητας `ex:hasMother` είναι στιγμιότυπα της κλάσης `ex:Person`. Εάν το `ex:hasMother` έχει περισσότερα από ένα ranges, για παράδειγμα το `ex:Person` και το `ex:Female`, αυτό λει ότι οι τιμές του `ex:hasMother` είναι πόροι που είναι στιγμιότυπα όλων των κλάσεων που είναι ranges δηλ., ότι οποιαδήποτε τιμή του `ex:hasMother` είναι και `ex:Female` και `ex:Person`. Τέλος, το `rdfs:range` μπορεί επίσης να χρησιμοποιηθεί για να δείξει ότι η τιμή μιας ιδιότητας μπορεί να δίνεται και από κυριολεκτικό (literal).

Το `rdfs:domain` χρησιμοποιείται για να δείξει ότι μια συγκεκριμένη ιδιότητα ισχύει για μια οριζόμενη κλάση. Παραδείγματος χάριν, εάν θέλαμε να δείξουμε ότι η ιδιότητα `ex:author` ισχύει για τα στιγμιότυπα της κλάσης `ex:Book`, θα γράφαμε τα RDF triples:

```
ex:Book      rdf:type      rdfs:Class.
ex:author    rdf:type      rdf:Property.
ex:author    rdfs:domain  ex:Book.
```

Μια δεδομένη ιδιότητα, για παράδειγμα `exterms:weight`, μπορεί να έχει ένα ή περισσότερα από ένα `domain`. Εάν το `exterms:weight` δεν έχει κανένα `domain`, τότε τίποτα δεν λέγεται για τους πόρους με τους οποίους το `exterms:weight` μπορεί να χρησιμοποιηθεί. Εάν το `exterms:weight` έχει ένα `domain`, για παράδειγμα το `ex:Book`, αυτό λέει ότι η ιδιότητα `exterms:weight` ισχύει για τα στιγμιότυπα της κλάσης `ex:Book`. Εάν το `exterms:weight` έχει περισσότερα από ένα `domain`, για παράδειγμα το `ex:Book` και το `ex:MotorVehicle`, αυτό λέει ότι οποιοσδήποτε πόρος που έχει μια ιδιότητα `exterms:weight` είναι στιγμιότυπο όλων των κλάσεων που είναι `domain` του `property` αυτού, δηλ. οποιοσδήποτε πόρος που έχει ιδιότητα `exterms:weight` είναι και ένα `ex:Book` και ένα `ex:MotorVehicle`.

Το RDF σχήμα παρέχει έναν τρόπο να ειδικευτούν οι ιδιότητες, όπως και οι κλάσεις. Αυτή η σχέση ειδίκευσης μεταξύ δύο ιδιοτήτων περιγράφεται χρησιμοποιώντας την προκαθορισμένη ιδιότητα `rdfs:subPropertyOf`. Παραδείγματος χάριν, εάν τα `ex:primaryDriver` και `ex:driver` είναι και οι δύο ιδιότητες, θα μπορούσαμε να περιγράψουμε αυτές τις ιδιότητες, και το γεγονός ότι το `ex:primaryDriver` είναι μια ειδίκευση του `ex:driver`, με το γράψιμο των RDF δηλώσεων:

```
ex:driver      rdf:type      rdf:Property.
ex:primaryDriver  rdf:type      rdf:Property.
ex:primaryDriver  rdfs:subPropertyOf  ex:driver.
```

Μια ιδιότητα (`property`) μπορεί να είναι ένα υπο-ιδιότητα (`subproperty`) καμίας, μίας ή περισσότερων ιδιοτήτων. Τα `rdfs:range` και `rdfs:domain` του RDFS που ισχύουν για μια ιδιότητα ισχύουν επίσης για κάθε μία από τις υπο-ιδιότητές της. Έτσι, στο παραπάνω παράδειγμα, το σχήμα RDF καθορίζει το `ex:primaryDriver` επίσης να έχει `rdfs:domain` το `ex:MotorVehicle`, λόγω της σχέσης `subproperty` με το `ex:driver`.

Παρακάτω φαίνεται ένα ολοκληρωμένο παράδειγμα σε μορφή RDF/XML:

```
<?xml version="1.0"?>
<!DOCTYPE rdf:RDF [<!ENTITY xsd "http://www.w3.org/2001/XMLSchema#">]>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://example.org/schemas/vehicles">

  <rdfs:Class rdf:ID="MotorVehicle"/>
  <rdfs:Class rdf:ID="PassengerVehicle">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Truck">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Van">
    <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="MiniVan">
    <rdfs:subClassOf rdf:resource="#Van"/>
    <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Person"/>

  <rdfs:Datatype rdf:about="xsd:integer"/>

  <rdf:Property rdf:ID="registeredTo">
```

```

    <rdfs:domain rdf:resource="#MotorVehicle"/>
    <rdfs:range rdf:resource="#Person"/>
</rdf:Property>

<rdf:Property rdf:ID="rearSeatLegRoom">
    <rdfs:domain rdf:resource="#PassengerVehicle"/>
    <rdfs:range rdf:resource="&xsd;integer"/>
</rdf:Property>

<rdf:Property rdf:ID="driver">
    <rdfs:domain rdf:resource="#MotorVehicle"/>
</rdf:Property>

<rdf:Property rdf:ID="primaryDriver">
    <rdfs:subPropertyOf rdf:resource="#driver"/>
</rdf:Property>

</rdf:RDF>

```

Στο σημείο αυτό, πρέπει να πούμε ότι παρόλο που το σύστημα τύπων του RDFS είναι παρόμοιο με αυτό των αντικειμενοστραφών γλωσσών προγραμματισμού, υπάρχουν σημαντικές διαφορές σε αρκετά σημεία. Πρώτον, το RDF Schema, αντί να περιγράφει τις κλάσεις σαν να έχουν κάποιες συγκεκριμένες ιδιότητες, περιγράφει τις ιδιότητες σαν να εφαρμόζονται σε συγκεκριμένους πόρους κλάσεις ως πεδίο ορισμού ή πεδίο τιμών. Για παράδειγμα, αν είχαμε μια κλάση Book με ένα χαρακτηριστικό author και μια άλλη κλάση SoftwareModule που επίσης είχε χαρακτηριστικό author, αυτά τα δύο χαρακτηριστικά θα θεωρούνταν διαφορετικά στις γλώσσες προγραμματισμού. Με άλλα λόγια, το πεδίο δράσης ενός χαρακτηριστικού καθορίζεται μέσα στην κλάση που το ορίζει. Αντίθετα, στο RDFS οι ιδιότητες είναι ανεξάρτητες και έχουν καθολικό πεδίο δράσης (εκτός αν οριστούν να εφαρμόζονται σε συγκεκριμένες κλάσεις χρησιμοποιώντας το πεδίο ορισμού).

Επιπλέον, οι ιδιότητες στο RDF μπορούν να κληρονομηθούν σε αντίθεση με τον αντικειμενοστρεφή προγραμματισμό που κάτι τέτοιο δεν ισχύει. Δηλαδή οι ιδιότητες μπορούν να κληρονομούν χαρακτηριστικά από άλλες ιδιότητες δημιουργώντας έτσι μια ιεραρχία ιδιοτήτων, όπως και οι κλάσεις.

Τέλος, μια εξίσου σημαντική διαφορά είναι ότι στο RDF τα στιγμιότυπα μιας ιδιότητας είναι προαιρετικά. Δηλαδή είναι δυνατό να υπάρχει ένα στιγμιότυπο μιας κλάσης χωρίς να υπάρχει στιγμιότυπο της ιδιότητας που εφαρμόζεται ως πεδίο ορισμού. Για παράδειγμα, στον προγραμματισμό, ένα στιγμιότυπο της κλάσης Book δεν επιτρέπεται να μην έχει το χαρακτηριστικό author, και επίσης, εάν το χαρακτηριστικό author έπαιρνε τιμές που είναι στιγμιότυπα της κλάσης Person, δε θα επιτρεπόταν να μην έχει τιμή. Αντίθετα, το RDF παρέχει πληροφορία για το σχήμα ως κάτι επιπρόσθετο χωρίς να επιβάλλει περιορισμούς στο πώς να χρησιμοποιηθεί η πληροφορία αυτή.

### 2.15.7 RDF Data Query Language (RDQL)

Η RDQL[16] είναι μια γλώσσα διατύπωσης ερωτήσεων για την RDF στα μοντέλα της Jena. Η ιδέα είναι να παρασχεθεί ένα αντικειμενοστρεφές μοντέλο ερώτησης έτσι ώστε να υπάρχει μια πιο δηλωτική προσέγγιση για να συμπληρώσει το Jena API.[]

Ένα αρχείο RDF μπορεί να αναπαρασταθεί μέσω ενός γράφου. Στη Jena, ο γράφος ονομάζεται μοντέλο και αναπαρίσταται από την ομώνυμη διασύνδεση (interface). Βασικές έννοιες της RDF αποτελούν οι πόροι (resources), οι ιδιότητες (properties), τα κυριολεκτικά (literals) και οι δηλώσεις (statements).

Κάθε δήλωση σε ένα RDF μοντέλο δηλώνει κάποιο γεγονός για ένα resource και έχει τρία μέρη

- το υποκείμενο, για το οποίο δηλώνεται κάτι
- το κατηγορημα (predicate), που αποτελεί μία ιδιότητα
- το αντικείμενο, που αποτελεί τον πόρο ή το literal που είναι η τιμή της ιδιότητας.



Μία δήλωση ονομάζεται και triple, λόγω των τριών μερών της. Ένα RDF μοντέλο αναπαρίσταται ως ένα σύνολο από δηλώσεις.

Οι ερωτήσεις είναι "δεδομενοστρεφείς" αφού ρωτούν μόνο τις πληροφορίες που φυλάσσονται στα μοντέλα. Δεν παράγεται κανένα συμπέρασμα. Φυσικά, το μοντέλο της Jena μπορεί να είναι "έξυπνο" στο ότι δίνει την εντύπωση ότι ορισμένες τριπλέτες υπάρχουν με τη δημιουργία τους κατόπιν παραγγελίας. Εντούτοις, το σύστημα RDQL δεν κάνει τίποτα άλλο από το να παίρνει την περιγραφή αυτού που η εφαρμογή θέλει, υπό μορφή ερώτησης, και επιστρέφει εκείνες τις πληροφορίες, υπό μορφή συνόλου συνδέσεων.

Η RDQL είναι μια εφαρμογή της SquishQL RDF γλώσσας διατύπωσης ερωτήσεων, η οποία προέρχεται από την rdfDB. Αυτή η κατηγορία γλωσσών διατύπωσης ερωτήσεων θεωρεί την RDF ως τριπλέτες δεδομένων, χωρίς το σχήμα ή πληροφορίες οντολογίας εκτός αν έχει ρητά συμπεριληφθεί στην πηγή της RDF.

Η RDF αναπαρίσταται από ένα γράφο με κατευθυνόμενες ακμές - οι κόμβοι είναι πόροι ή κυριολεκτικά. Η RDQL παρέχει έναν τρόπο καθορισμού ενός μοτίβου γράφου που συγκρίνεται με το γράφο ώστε να παραγάγει ένα σύνολο αντιστοιχιών. Επιστρέφει έναν κατάλογο συνδέσεων όπου κάθε σύνδεση είναι ένα σύνολο ζευγών ονόματος-αξίας για τις τιμές των μεταβλητών. Όλες οι μεταβλητές είναι συνδεδεμένες (δεν υπάρχει κανένας διαχωρισμός στην ερώτηση).

### **2.15.8 RDQL Syntax**

Η RDQL έχει παρόμοια σύνταξη με την SQL για το μοντέλο ερώτησης που προέρχεται από την SquishQL και την rdfDB. Στην SQL, μια βάση δεδομένων είναι ένας κλειστός κόσμος. Ο όρος "FROM" προσδιορίζει τους πίνακες στη βάση δεδομένων. Ο όρος "WHERE" προσδιορίζει τους περιορισμούς και μπορεί να επεκταθεί με "AND". Αναλογικά, το Web είναι η βάση δεδομένων και ο όρος "FROM" προσδιορίζει τα μοντέλα RDF. Οι μεταβλητές εισάγονται μετά από το σύμβολο "?" και τα URIs

αναφέρονται με "< >".URIs χωρίς "< >" μπορούν να χρησιμοποιηθούν όπου δεν υπάρχει καμία ασάφεια.

### **SELECT Clause**

Προσδιορίζει τις μεταβλητές που επιστρέφονται στην εφαρμογή. Εάν δεν απαιτούνται όλες οι μεταβλητές από την εφαρμογή, τότε η διευκρίνιση των απαραίτητων αποτελεσμάτων μπορεί να μειώσει το ποσό μνήμης που απαιτείται για τα αποτελέσματα καθώς επίσης και να παρέχει πληροφορίες σε έναν query optimizer.

### **FROM Clause**

Διευκρινίζει το μοντέλο μέσω ενός URI.

### **WHERE Clause**

Διευκρινίζει το μοτίβο του γράφου ως λίστα μοτίβων τριπλετών.

### **AND Clause**

Διευκρινίζει τις λογικές εκφράσεις.

### **USING Clause**

Ένας τρόπος να μειωθεί το μήκος των URIs. Καθώς η SquishQL είναι πιθανό να γραφτεί από τους ανθρώπους, αυτός ο μηχανισμός βοηθά να γίνει η σύνταξη πιο κατανοητή για τους ανθρώπους. Αυτό δεν είναι ένας μηχανισμός namespace. Αντ' αυτού είναι απλά ένας μηχανισμός συντομογραφιών για μακριά URIs με τον καθορισμό ενός προθέματος γραμματοσειράς.

Η ενημερωμένη γραμματική διανέμεται ως τμήμα του Jena toolkit στη διεύθυνση: [http://jena.sourceforge.net/RDQL/rdql\\_grammar.html](http://jena.sourceforge.net/RDQL/rdql_grammar.html). [14]

Η RDQL μπορεί να χρησιμοποιηθεί τόσο από το command line όσο και μέσα σε προγράμματα Java.

Για παραδείγματα και περισσότερες λεπτομέρειες σχετικά με τη σύνταξη και χρησιμοποίηση της RDQL σε command line ή Java, δείτε το tutorial της Jena στη διεύθυνση: <http://jena.sourceforge.net/tutorial/RDQL>. [14]

Όπως αναφέρθηκε και παραπάνω το Semantic Web θα φέρει δομή στο σημασιολογικό περιεχόμενο του Web, δημιουργώντας ένα περιβάλλον όπου οι πράκτορες λογισμικού που περιπλανώνται από σελίδα σε σελίδα θα μπορούν εύκολα να εκτελέσουν περίπλοκους στόχους για τους χρήστες. Το Semantic Web θα στηριχτεί στη δυνατότητα της XML να καθορίζει προσαρμοσμένα σχήματα και στην ευελιξία της RDF (Resource Description Framework) να αναπαριστά δεδομένα. Το επόμενο στοιχείο που απαιτείται για το Semantic Web είναι μια γλώσσα οντολογίας Ιστού (Web Ontology Language) που μπορεί τυπικά να περιγράψει τη σημασιολογία των κλάσεων και των ιδιοτήτων που χρησιμοποιούνται στα έγγραφα του Web. Η γλώσσα αυτή πρέπει να υπερβεί τη βασική σημασιολογία του σχήματος RDF, προκειμένου οι μηχανές να εκτελούν χρήσιμους συλλογισμούς σε αυτά τα έγγραφα.

### **2.15.9 XML**

Η XML [11] (eXtensible Markup Language) είναι ένα απλό και ιδιαίτερα ευέλικτο πρότυπο κειμένου που προήλθε από την Standard Generalized Markup Language (SGML). Αρχικά αναπτύχθηκε για καλύψει τις ανάγκες ευρείας κλίμακας δημοσίευσης δεδομένων αλλά παίζει όλο και μεγαλύτερο ρόλο στην ανταλλαγή ενός μεγάλου εύρους δεδομένων στο διαδίκτυο αλλά και αλλού. Η XML είναι ένας τρόπος για περιγραφή δεδομένων και ένα αρχείο XML μπορεί να περιέχει και τα δεδομένα που περιγράφει με λογική παρόμοια αυτής των βάσεων δεδομένων. Γλώσσες βασιζόμενες στην XML επιτρέπουν σε προγράμματα να τροποποιούν και να ελέγχουν κείμενα γραμμένα στις γλώσσες αυτές χωρίς να έχουν γνώση της μορφής τους.

Η XML προσφέρει τρόπους βασιζόμενους σε απλό κείμενο για την περιγραφή και εφαρμογή μιας δενδρικής μορφής στα δεδομένα. Είναι μια δηλωτική γλώσσα, δηλαδή τα δεδομένα συνοδεύουν δηλώσεις ή μεταδεδομένα που τα αφορούν. Μια δήλωση έχει την εξής μορφή:

```
<name attribute="value">content</name>
```

Αυτή η στοιχειώδης δήλωση ονομάζεται element. Ένα element έχει ένα όνομα (name), πιθανώς κάποια χαρακτηριστικά (attributes) και ένα περιεχόμενο (content). Το περιεχόμενο περικλείεται από δύο ετικέτες (tags) που έχουν συγκεκριμένη δομή. Κύριο χαρακτηριστικό είναι η δυνατότητα δόμησης δενδρικών δομών με την εξής μορφή:

```
<element>  
  <child1/>  
  <child2/>  
</element>
```

Η δυνατότητες της επισύναψης μεταδεδομένων και της δημιουργίας δενδρικών δομών σε συνάρτηση με το γεγονός ότι μπορούν να περιγραφούν οποιαδήποτε δεδομένα κάνει την XML ένα δυνατό εργαλείο για τη δόμηση πιο εκφραστικών δηλωτικών γλωσσών. Ακριβώς αυτά τα χαρακτηριστικά εκμεταλλεύεται και η RDF που είναι το υπόστρωμα της OWL καθώς και ένα χρήσιμο συμπλήρωμά της.

## 2.16 Γλώσσες Περιεχομένου και Οντολογίες στη JADE

Μέσα στο ACL μήνυμα, οι πληροφορίες αναπαρίστανται ως έκφραση περιεχομένου σύμφωνη με μια κατάλληλη γλώσσα περιεχομένου, π.χ. SL, και κωδικοποιούνται σε μία

κατάλληλη μορφή, π.χ. string. Οι πράκτορες έχουν το δικό τους (ενδεχομένως διαφορετικό) τρόπο να αναπαριστούν εσωτερικά τις πληροφορίες αυτές. Ο τρόπος αυτός πρέπει να επιτρέπει τον εύκολο χειρισμό των πληροφοριών, οπότε και είναι αρκετά σαφές ότι η αναπαράσταση που χρησιμοποιείται σε μια ACL έκφραση περιεχομένου δεν είναι κατάλληλη για το εσωτερικό ενός πράκτορα. Επίσης, η αποθήκευση αυτών των πληροφοριών μέσα σε έναν πράκτορα απλά ως μεταβλητή string δεν είναι κατάλληλη για τον χειρισμό των πληροφοριών. Στους πράκτορες λογισμικού που γράφονται σε Java (όπως οι πράκτορες JADE), οι πληροφορίες μπορούν βολικά να αναπαρασταθούν μέσα σε έναν πράκτορα ως Java αντικείμενα.

Η JADE υποστηρίζει τις γλώσσες περιεχομένου και τις οντολογίες προκειμένου οι μετατροπές από τη μία αναπαράσταση στην άλλη και οι λειτουργίες ελέγχου των πληροφοριών να εκτελούνται αυτόματα, επιτρέποντας στους υπεύθυνους ανάπτυξης εφαρμογών το χειρισμό των πληροφοριών μέσα στους πράκτορες ως Java αντικείμενα. Πιο συγκεκριμένα, η οντολογία επικυρώνει σημασιολογικά τις πληροφορίες, ενώ ο codec (codec για μια γλώσσα περιεχομένου είναι ένα Java αντικείμενο ικανό να διαχειριστεί εκφράσεις περιεχομένου γραμμένες στην γλώσσα αυτή) εκτελεί τη μετάφραση σε strings (ή ακολουθίες bytes) σύμφωνα με τους συντακτικούς κανόνες της σχετικής γλώσσας περιεχομένου. Ο χρήστης δεν χρειάζεται να φροντίσει για αυτές τις διαδικασίες, εκτός αν πρέπει να χρησιμοποιήσει μερικά προηγμένα χαρακτηριστικά, όπως να θέσει ερωτήσεις.

Για να εκτελέσει η JADE τους κατάλληλους σημασιολογικούς ελέγχους σε μια δεδομένη έκφραση περιεχομένου είναι απαραίτητο να ταξινομηθούν, σύμφωνα με τα γενικά σημασιολογικά χαρακτηριστικά τους, όλα τα πιθανά στοιχεία που μπορούν να εμφανιστούν μέσα σε μια έγκυρη πρόταση που στέλνεται από έναν πράκτορα σε έναν άλλον, ως περιεχόμενο ενός ACL μηνύματος. Λεπτομερέστερα στο πρώτο επίπεδο κάνουμε τη διάκριση μεταξύ των predicates (κατηγορημάτων) και των terms (όρων).

**Predicates:** Είναι εκφράσεις που λένε κάτι για την κατάσταση του κόσμου και μπορούν να είναι true ή false.

**Terms:** Είναι εκφράσεις που προσδιορίζουν τις οντότητες (αφηρημένες ή συγκεκριμένες) που “υπάρχουν” στον κόσμο και για τις οποίες οι πράκτορες μιλούν και εξάγουν συλλογισμούς. Είναι περαιτέρω ταξινομημένοι σε:

**Concepts:** Εκφράσεις που υποδεικνύουν οντότητες με μια σύνθετη δομή που μπορεί να καθοριστεί σε σχέση με τα slots. Τα concepts (έννοιες) συνήθως δεν έχουν κανένα νόημα εάν χρησιμοποιούνται άμεσα ως περιεχόμενο ενός ACL μηνύματος. Γενικά αναφέρονται μέσα σε κατηγορήματα και σε άλλα concepts.

**Agent actions:** Ειδικές έννοιες που υποδεικνύουν ενέργειες που μπορούν να εκτελεσθούν από κάποιους agents. Οι επικοινωνιακές πράξεις (δηλ. μηνύματα ACL) είναι κι αυτές ενέργειες πρακτόρων.

**Primitives:** Εκφράσεις που υποδεικνύουν ατομικές οντότητες όπως strings και ακέραιοι αριθμοί

**Aggregates:** Εκφράσεις που υποδεικνύουν οντότητες που είναι ομάδες άλλων οντοτήτων

**Identifying Referential Expressions (IRE):** Εκφράσεις που προσδιορίζουν την οντότητα (ή τις οντότητες) για την οποία ισχύει ένα δεδομένο κατηγορημα

**Variables:** Εκφράσεις (που χρησιμοποιούνται συνήθως στις ερωτήσεις) που υποδεικνύουν ένα γενικό στοιχείο άγνωστο αρχικά.

Μια πλήρως εκφραστική γλώσσα περιεχομένου πρέπει να είναι σε θέση να αναπαραστήσει και να διακρίνει τα στοιχεία όλων των ανωτέρω τύπων. Η JADE προσφέρει codecs για δύο γλώσσες περιεχομένου (την SL και την LEAP γλώσσα).

Η γλώσσα SL είναι μια γλώσσα περιεχομένου κατανοήσιμη από τον άνθρωπο, κωδικοποιημένη σε string. Είναι πιθανώς η περισσότερο διαδεδομένη γλώσσα περιεχομένου στην επιστημονική κοινότητα που ασχολείται με τους ευφυείς πράκτορες.

Η γλώσσα LEAP είναι μια γλώσσα περιεχομένου μη κατανοήσιμη από τον άνθρωπο, κωδικοποιημένη σε bytes που έχει οριστεί ως πρόσθετο της JADE μέσα στο πρόγραμμα LEAP. Είναι επομένως σαφές ότι μόνο οι JADE πράκτορες θα είναι σε θέση

να επικοινωνήσουν με τη γλώσσα LEAP. Όταν υπάρχουν ισχυροί περιορισμοί μνήμης η γλώσσα αυτή είναι προτιμητέα.

Μια οντολογία για μια δεδομένη θεματική περιοχή είναι ένα σύνολο σχημάτων που καθορίζουν τη δομή των κατηγορημάτων, των ενεργειών πρακτόρων και των εννοιών (βασικά τα ονόματά τους και τα slots τους) που σχετίζονται με εκείνη την περιοχή.

## 2.17 Σημασιολογικός ιστός και ηλεκτρονικές επιχειρήσεις

Οι υπηρεσίες Ιστού (Web Services) είναι εφαρμογές που προσπελούνται μέσω του Διαδικτύου. Για παράδειγμα, η κράτηση και έκδοση ενός εισιτηρίου μέσω κάποιας ιστοσελίδας είναι μία υπηρεσία Ιστού. Ένα πεδίο εφαρμογής του Σημασιολογικού Ιστού είναι οι Υπηρεσίες Ιστού, οι οποίες απαιτούν γλώσσες προγραμματισμού που θα τις περιγράφουν. Μέσω αυτών των γλωσσών, οι πράκτορες λογισμικού θα βρίσκουν αυτόματα τις κατάλληλες υπηρεσίες Ιστού, θα καταλαβαίνουν τις δυνατότητες και τους περιορισμούς τους, και θα τις χρησιμοποιούν. Στο επιχειρηματικό Intranet, η ακριβής αναζήτηση της πληροφορίας έχει τεράστια σημασία, ενώ η διαχείριση της γνώσης ασχολείται με τις τεχνικές απόκτησης, διατήρησης και πρόσβασης στη γνώση. Τα σημερινά συστήματα διαχείρισης γνώσης έχουν σημαντικές αδυναμίες που εντοπίζονται στα εξής σημεία:

α) Στην εύρεση-αναζήτηση των πληροφοριών: Συνήθως, η αναζήτηση που βασίζεται σε μία λέξη-κλειδί (keyword) ανακτά άσχετες πληροφορίες που ενσωματώνουν τη συγκεκριμένη λέξη, με διαφορετικό όμως νόημα από το επιδιωκόμενο. Άλλες φορές πάλι, ο χρήστης-ερευνητής χάνει χρήσιμες σχετικές πληροφορίες, διότι αυτές χρησιμοποιούν διαφορετικές λέξεις (για το περιεχόμενο) από τη λέξη-κλειδί.

β) Στην εξαγωγή πληροφοριών: Η ανθρώπινη αναζήτηση και ανάγνωση είναι απαραίτητες για την εξαγωγή σχετικών πληροφοριών από τις πηγές πληροφορίας. Οι περισσότεροι αυτόματοι πράκτορες λογισμικού στερούνται της κοινής λογικής που απαιτείται για την απόσπαση τέτοιων πληροφοριών.

γ) Στη διατήρηση των πληροφοριών: Η διατήρηση δομημένων πηγών πληροφορίας είναι δύσκολη και χρονοβόρα διαδικασία, καθώς οι πηγές αυξάνουν σημαντικά.

δ) Στην αυτόματη δημιουργία εγγράφων στο Web: Οι δικτυακές τοποθεσίες που προσαρμόζουν δυναμικά το περιεχόμενό τους, σύμφωνα με το προφίλ του χρήστη, είναι πολύ χρήσιμες. Ωστόσο, η δημιουργία ημιδομημένων παρουσιάσεων πληροφορίας απαιτεί μία σημασιολογική αναπαράσταση των αντίστοιχων πηγών πληροφορίας, και μια τέτοια αναπαράσταση δεν υπάρχει σήμερα. Ο Σημασιολογικός Ιστός θα δώσει τη δυνατότητα για δομικούς και σημασιολογικούς ορισμούς των εγγράφων, παρέχοντας έτσι νέες δυνατότητες όπως: 1) έξυπνη αναζήτηση αντί της τωρινής αναζήτησης, 2) απαντήσεις σε υποβληθείσες ερωτήσεις αντί για ανάκτηση των πληροφοριών, 3) αυτόματες ανταλλαγές εγγράφων μεταξύ των διαφορετικών τμημάτων μιας επιχείρησης και 4) προσαρμοσμένες προβολές των εγγράφων.

ε) Στην πλοήγηση του Ιστού: Οι απλές τεχνικές πλοήγησης παρεμποδίζουν την μελλοντική ανάπτυξή του, καθώς η πλοήγηση υποστηρίζεται μόνο από προκαθορισμένους συνδέσμους (links). Η τωρινή τεχνολογία πλοήγησης δεν υποστηρίζει σελίδες-συνδέσμους που βασίζονται στη σημασιολογική ομοιότητα. Επιπρόσθετες αδυναμίες που συναντάμε στο e-Επιχειρείν είναι οι ακόλουθες.

### **Αδυναμίες στο ηλεκτρονικό εμπόριο B2C (Business to Consumer)**

Σκοπός κάθε αγοραστή είναι να βρει ένα κατάστημα που να πουλά το προϊόν που ακριβώς θέλει, στην επιθυμητή ποιότητα και ποσότητα, και φυσικά να πληρώσει όσο το δυνατόν λιγότερο. Η υπάρχουσα τεχνολογία εφαρμογών ηλεκτρονικού εμπορίου B2C (Business to Consumer) έχει δύο σημαντικούς περιορισμούς:

1. Απαιτείται μεγάλη προσπάθεια από τον αγοραστή. Συγκεκριμένα, ο υποψήφιος αγοραστής πρέπει να βρει (και να χρησιμοποιήσει) τις διευθύνσεις των ηλεκτρονικών



καταστημάτων που παρέχουν το προϊόν που επιθυμεί. Τέλος, δεν είναι σίγουρο πως οι πληροφορίες των προϊόντων είναι έγκυρες και ότι οι διευθύνσεις των ηλεκτρονικών καταστημάτων τους είναι αναβαθμισμένες.

2. Ποιότητα Πληροφορίας: Οι πληροφορίες που παίρνει ο δυνητικός αγοραστής για το προϊόν είναι περιορισμένες. Πολλές φορές, οι πληροφορίες είναι ελλιπείς και περιορίζονται κυρίως στην τιμή. Για παράδειγμα, ο αγοραστής ίσως βρει το προϊόν σε μια συγκεκριμένη τιμή, αλλά σε αυτήν να έχει παραλειφθεί το έμμεσο κόστος της αποστολής του προϊόντος.

Με την εφαρμογή του Σημασιολογικού Ιστού αυτή η κατάσταση θα αλλάξει δραματικά, καθώς θα παρέχεται μία ακριβή μηχανή αναζήτησης με όλα τα διαθέσιμα στοιχεία του προϊόντος. Επιπλέον, θα κατασκευαστούν πράκτορες λογισμικού που θα «καταλαβαίνουν» όλες τις πληροφορίες του προϊόντος που θα παρέχονται στο Web

### **Αδυναμίες στο ηλεκτρονικό εμπόριο B2B (Business to Business)**

Οι εφαρμογές B2B (Business to Business) ανταλλάσσουν μεταξύ των επιχειρήσεων δεδομένα που αφορούν επιχειρησιακές διαδικασίες. Η χρήση της γλώσσας ebXML (<http://www.ebxml.org>) [58] δεν παρέχει συγκεκριμένες δομές πληροφορίας, αλλά και ορολογίες που θα περιγράφουν σημασιολογικά τις επιχειρησιακές διαδικασίες. Προτεινόμενες γλώσσες με συγκεκριμένα μοντέλα επιχειρηματικών πληροφοριών θα πρέπει να αποδείξουν πώς όντως υποστηρίζουν τον καθορισμό, και την ανταλλαγή της πληροφορίας των προϊόντων. Τέλος, θα απαιτηθούν αποτελεσματικές υπηρεσίες μετάφρασης σε θεματικές περιοχές στις οποίες δεν υπάρχουν συγκεκριμένες οντολογίες. *Για αυτό ο Σημασιολογικός Ιστός θα παίξει έναν σημαντικό ρόλο στις εφαρμογές ηλεκτρονικού εμπορίου B2B.*

## 2.18 Συμπερασματα

Ο Σημασιολογικός Ιστός (Semantic Web) είναι το επόμενο βήμα του Παγκόσμιου Ιστού, όπου η πληροφορία αποκτά δομή και σημασιολογία, έτσι ώστε να υποστηριχθεί η αποδοτική αναζήτηση, επεξεργασία και ενοποίηση των δεδομένων. είναι ένα όραμα και μια πρόταση για την μετεξέλιξη του Διαδικτύου και ειδικότερα του Παγκόσμιου Ιστού (World Wide Web).

Ο όρος Semantic Web, καθώς και η αρχιτεκτονική για την υλοποίησή του, προτάθηκαν από τον Tim Berners-Lee, τον εφευρέτη του σημερινού Παγκόσμιου Ιστού.

Ο Σημασιολογικός Ιστός είναι μία επέκταση και βελτίωση του σημερινού Web στην κατεύθυνση, κυρίως της δόμησης της πληροφορίας, έτσι ώστε αυτή να είναι προσπελάσιμη από εφαρμογές υπολογιστών, με τελικό στόχο την αυτοματοποίηση πολλών λειτουργιών στο Διαδίκτυο. Η σημερινή αναπαράσταση της πληροφορίας που προορίζεται για χρήση από ανθρώπους θα αντικατασταθεί από μία αναπαράσταση κατανοητή από υπολογιστές. Ο Σημασιολογικός Ιστός προσπαθεί να επιλύσει το πρόβλημα της αναπαράστασης της γνώσης από τους υπολογιστές. Βασικό συστατικό του Σημασιολογικού Ιστού είναι ο μηχανισμός επεξεργασίας της γνώσης που διαχειρίζεται λογικά τις πληροφορίες με σκοπό την εξαγωγή συμπερασμάτων, τη δημιουργία νέας γνώσης, την υποστήριξη στη λήψη αποφάσεων, και τέλος την αυτόματη εκτέλεση ενεργειών.

Οι βασικές αρχές του Σημασιολογικού Ιστού είναι οι παρακάτω :

Η διατήρηση του κατανεμημένου περιεχομένου του Διαδικτύου.

Η αναπαράσταση και ανάκτηση της πληροφορίας, καθώς οι εφαρμογές των υπολογιστών προσπελαίνουν δομημένες πηγές πληροφορίας και κανόνες, οι οποίοι χρησιμοποιούνται για να αιτιολογούν τις σχέσεις μεταξύ των πληροφοριών.

Η αναπαράσταση των εννοιών μιας θεματικής περιοχής επιτυγχάνεται με τη χρήση των οντολογιών.

Η ύπαρξη πρακτόρων λογισμικού (software agents), δηλαδή προγραμμάτων που θα αναλαμβάνουν για λογαριασμό του χρήστη να κινούνται στο Διαδίκτυο και να συλλέγουν την πληροφορία από διάφορες πηγές που διαθέτουν σημασιολογικό περιεχόμενο.

Ο Σημασιολογικός ιστός χρησιμοποιεί ορισμένες τεχνολογίες οι βασικότερες από αυτές είναι :

### **A ) XML**

Η XML (eXtensible Markup Language - επεκτάσιμη γλώσσα σήμανσης) είναι η επικρατέστερη γλώσσα για την περιγραφή και ανταλλαγή δεδομένων και κειμένων στο Διαδίκτυο. Η XML παρέχει τη δυνατότητα δημιουργίας κειμένων με απεριόριστα πολύπλοκη δομή και συντακτικό. Έτσι, μπορούν να δομηθούν οι πληροφορίες που περιέχονται στα κείμενα για να επεξεργάζονται πιο εύκολα από τους υπολογιστές.

### **B) RDF και RDF Schema**

Η γλώσσα RDF (Resource Description Framework - Περιβάλλον Περιγραφής Πόρων) είναι το πρότυπο που υιοθετήθηκε από το W3C για την περιγραφή πληροφοριακών πόρων και γενικότερα για την αναπαράσταση της γνώσης στο περιβάλλον του Διαδικτύου. Μέσω του RDF είναι δυνατή η μετατροπή της πληροφορίας σε σημασιολογική. Η γλώσσα RDF είναι ένα απλό μοντέλο δεδομένων, στο οποίο όλες οι προτάσεις αποτελούν μια τριπλέτα της μορφής: **P(O,V)** όπου **Property** (ιδιότητα), **Object** (αντικείμενο), **Value**(τιμή).

Η συντριπτική πλειοψηφία της γνώσης που θέλουμε να αναπαραστήσουμε στους υπολογιστές μπορεί να αναπαρασταθεί με αυτή την μορφή. Στο περιβάλλον RDF, το αντικείμενο και η ιδιότητα δηλώνονται με ένα URI. Η τιμή μπορεί να δηλώνεται με ένα URI ή μπορεί να είναι ένα αλφαριθμητικό ή μια λέξη. Τέλος, το πρότυπο RDF ορίζει το συντακτικό XML, δηλαδή τον τρόπο με τον οποίο οι προτάσεις RDF εκφράζονται ως XML κείμενα.

Το RDF Schema είναι η οντοκεντρική επέκταση του RDF. Είναι μια γλώσσα με την οποία το μοντέλο δεδομένων του RDF εμπλουτίζεται με χαρακτηριστικά αντικειμενοστραφούς αναπαράστασης, όπου ο πόρος αντιστοιχεί σε αντικείμενο. Συγκεκριμένα, το RDF Schema ορίζει ένα λεξικό για να εκφράζονται οι κατηγορίες (κλάσεις) των πόρων, οι πόροι, οι ιδιότητές τους και οι μεταξύ τους σχέσεις.

### Γ ) OWL και Οντολογίες

Η OWL (Web Ontology Language-Γλώσσα Οντολογιών Ιστού) είναι μία γλώσσα που χρησιμοποιείται για την περιγραφή των οντολογιών που υπάρχουν στο Διαδίκτυο. Η οντολογία είναι μία αυστηρή περιγραφή των πόρων και των μεταξύ τους σχέσεων. Συγκεκριμένα, η οντολογία είναι η αποδεκτή σημασιολογικά κωδικοποίηση της πληροφορίας ενός θεματικού χώρου. Οι οντολογίες επιτρέπουν στους χρήστες να έχουν κοινή ονοματολογία και αντίληψη για τα αντικείμενα που δηλώνουν ή χρησιμοποιούν. Βοηθούν τον χρήστη να πλοηγηθεί στον θεματικό χώρο της πληροφορίας που βασίζεται σε σημασιολογικές και όχι σε λεξιλογικές έννοιες. Στις οντολογίες, η δυσκολία εντοπίζεται στο ότι οι κοινότητες χρηστών με κοινά ενδιαφέροντα θα πρέπει να συμφωνήσουν στην οντολογική περιγραφή του θεματικού χώρου ενδιαφέροντός τους. Για την περιγραφή των οντολογιών έχει αναπτυχθεί διάφορες γλώσσες όπως η KIF κλπ.....

Για να δημιουργήσουν τις οντολογίες τους, οι μηχανικοί διαχείρισης γνώσης χρησιμοποιούν επεξεργαστές οντολογιών (ontology editors), οι οποίοι υποστηρίζουν τον ορισμό της ιεραρχίας των εννοιών και παράλληλα προσφέρουν τη δυνατότητα επιθεώρησης, αναζήτησης, κωδικοποίησης και τροποποίησης των οντολογιών. Οι επεξεργαστές οντολογιών είναι χρήσιμοι όταν παρέχουν γραφικές διασυνδετικές διατάξεις (GUIs) και εναρμονίζονται με τα ήδη υπάρχοντα πρότυπα της web-based ανάπτυξης λογισμικού.

Μερικά από τα πιο γνωστά συστήματα ανάπτυξης και διαχείρισης των οντολογιών που μπορούν να χρησιμοποιηθούν είτε για την ανάπτυξη των οντολογιών από μηδενική

βάση είτε για την επεξεργασία και επαναχρησιμοποίηση οντολογιών που κατασκευάστηκαν από τρίτους είναι τα παρακάτω :

1. Apollo
2. LinKFactory
3. OilEd
4. OntoEdit
5. Ontolingua Server
6. OntoSaurus
7. OpenKnoME
8. Protege
9. SymOntoX
10. WebODE
11. WebOnto.

Η συστηματική αξιολόγηση των οντολογιών ίσως οδηγήσει σε προτυποποιημένα μέτρα σύγκρισης και αντίστοιχες πιστοποιήσεις ποιότητας .Τα πιο γνωστά από τα εργαλεία αξιολόγησης των οντολογιών είναι τα εξής :

1. OntoAnalyser
2. OntoGenerator
3. OntoClean
4. ONE-T.

#### **Δ) Πράκτορες Λογισμικού ( Software agents )**

Οι πράκτορες λογισμικού (software agents) είναι προγράμματα που εκτελούν κάποια λειτουργία και παράγουν αποτελέσματα με το πέρας της εκτέλεσης αυτής. Συνήθως, οι πράκτορες λογισμικού περιδιαβαίνουν το Διαδίκτυο και επεξεργάζονται τις πληροφορίες που βρίσκουν στις ιστοσελίδες που επισκέπτονται. Συχνά, οι πράκτορες λογισμικού χρησιμοποιούνται για λειτουργίες όπως: εύρεση, ταξινόμηση και επιλογή δεδομένων. Στο ηλεκτρονικό εμπόριο, μερικά παραδείγματα λειτουργιών τους είναι: η σύγκριση

τιμών του ίδιου προϊόντος σε πολλά ηλεκτρονικά καταστήματα, η ειδοποίηση για την εμφάνιση νέου περιεχομένου σε δικτυακούς τόπους ειδήσεων και ενημέρωσης κ.α.

Σήμερα ο Σημασιολογικός Ιστός και οι τεχνολογίες ανάπτυξης που έχουν αναπτυχθεί καλούνται να αντιμετωπίσουν και να επιλύσουν μια σειρά από ανοιχτά θέματα, τα περισσότερα από τα οποία είναι πολύ σημαντικά. Ενδεικτικά αναφέρονται τα παρακάτω :

1. Ανάπτυξη οντολογιών-προτύπων σε διάφορα πεδία γνώσης (κατασκευή οντολογιών που θα λειτουργούν ως κέντρα αναφοράς για τις υπόλοιπες).
2. Εισαγωγή νέων οντολογιών σε υφιστάμενες διαδικασίες και η εφαρμογή της έννοιας του χρόνου πάνω σε αυτές.
3. Μετάφραση των οντολογιών.
4. Συνταίριασμα των οντολογιών.
5. Εγκατάσταση μηχανισμών εμπιστοσύνης πάνω στις ζητούμενες πληροφορίες.
6. Η ανάπτυξη εργαλείων τα οποία θα επιτρέπουν στους τελικούς χρήστες να αναπτύσσουν native οντολογίες.
7. Η αντιστοίχιση των συσχετίσεων και αλληλεξαρτήσεων των οντολογιών σε διαφορετικές γλώσσες.
8. Η ενσωμάτωση εννοιών κουλτούρας σε εθνικές οντολογίες
9. Η σημασιολογία για την ανακάλυψη, ανάθεση, σύνθεση και ενορχήστρωση υπηρεσιών δικτύου

Σταδιακά, η χρήση του Σημασιολογικού Ιστού επεκτείνεται σε συσκευές (π .χ Κινητά τηλέφωνα, palmtops, συσκευές GPS, τηλεοράσεις, ταμπλό αυτοκινήτων, ψυγεία, MP3 players ) που συνδέονται στο Διαδίκτυο. Καθώς οι συσκευές αυτές θα χρησιμοποιούν τεχνολογίες Σημασιολογικού Ιστού, θα περιγράφουν τις λειτουργίες τους και θα τις ανακοινώνουν στο Διαδίκτυο. Έτσι, θα προκύψουν νέες δυνατότητες και αποδοτικότερη συνεργασία μεταξύ των συσκευών αυτών και του Web. Για παράδειγμα, στο μέλλον η τηλεόραση θα μας προτείνει προγράμματα σύμφωνα με κριτήρια που θέσαμε στον Σημασιολογικό Ιστό. Ένα σημαντικό χαρακτηριστικό των μελλοντικών συστημάτων διαχείρισης πελατειακών σχέσεων (CRM) θα είναι οι «συσκευές διεπαφής χρήστη», οι

οποίες θα μπορούν να «αναγνωρίζουν» τον χρήστη και να «κατανοούν» τη διάθεση και τα συναισθήματά του. Στη συνέχεια, θα του προτείνουν εξατομικευμένες υπηρεσίες και προϊόντα

Δεν είναι δύσκολο να προβλέψουμε ότι στο απώτερο ο Ιστός θα είναι μία τεράστια περιοχή παράλληλης επεξεργασίας, που με αρκετό προγραμματισμό τεχνητής νοημοσύνης, ενσωμάτωση σημασιολογικών συστημάτων και συνεχή εισροή περιεχομένου από δισεκατομμύρια ανθρώπους, θα έχει τη δυνατότητα να γίνει μία ενοποιημένη νοημοσύνη, μία σφαίρα πληροφοριών που θα περικλείει τον πλανήτη και θα είναι ισχυρότερη από τα συστατικά του μέρη.

Φυσικά η έννοια της ενοποιημένης υπερ-νοημοσύνης δεν είναι καινούργια και έχει εκφραστεί κυρίως μέσα από τη λογοτεχνία επιστημονικής φαντασίας. Αλλά καλό είναι να σκεφτούμε πόσα πράγματα υπήρχαν στην επιστημονική φαντασία και αργότερα έγιναν πραγματικότητα. Το Internet και ο Ιστός ειδικότερα, είναι κάτι πρωτοφανές στα χρονικά αυτού του πλανήτη. Δεν είναι μόνο η δυνατότητα να συνδέσει τον καθένα, αλλά και να επεκτείνει γεωμετρικά κάθε εγκέφαλο. Ίσως πάρει περισσότερο χρόνο, αλλά τελικά ο Ιστός του μέλλοντος θα είναι συντριπτικά διαφορετικός και πιο ισχυρός απ' ότι μπορούμε μάλλον να φανταστούμε σήμερα. [9]

## Βιβλιογραφία – Αναφορές

- [1] Stenmark Dick (2002). Designing the new Intranet, Gothenburg Studies in Informatics, Report 21, March 2002, ISSN 1400-741X
- [2] RDF Primer, W3C Recommendation , <http://www.w3.org/TR/rdf-primer>
- [3] Semantic Web, <http://www.semanticweb.org>
- [4] “Ontology”, Wikipedia, <http://en.wikipedia.org>
- [5] “Semantic Web”, Wikipedia, <http://en.wikipedia.org>
- [6] “What is ontology?”, [http://www.ontologyworks.com/what\\_is\\_ontology.php](http://www.ontologyworks.com/what_is_ontology.php)
- [7] G. Antoniou and F. van Harmelen, “Web Ontology Language: OWL”, <http://www.cs.vu.nl/~frankh/postscript/OntoHandbook03OWL.pdf>
- [8] T. Berners-Lee, J. Hendler, O. Lassila, “The Semantic Web”, Scientific American, May 2001.
- [9] C. Boulton, “W3C Wraps Up Semantic Web Standards”, <http://www.internetnews.com/infra/article.php/3310831>
- [10] B. Chandrasekaran, J. R. Josephson, V. Richard Benjamins, “What Are Ontologies, and Why Do We Need Them?” IEEE Intelligent Systems, 1999, 14 (1): pp. 20 – 26, <http://www.aaai.org/AITopics/html/ontol.html#what>
- [11] Extensible Markup Language (XML), [www.w3.org/XML/](http://www.w3.org/XML/)
- [12] T. Gruber and the Knowledge Systems Laboratory at Stanford University, “What is an Ontology?”, <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
- [13] Intelligent Software Agents Lab, <http://www-2.cs.cmu.edu/~softagents>



- [14] An Introduction to RDF and the Jena RDF API, [http://jena.sourceforge.net/tutorial/RDF\\_API](http://jena.sourceforge.net/tutorial/RDF_API)
- [15] The Jena 2 Ontology API, <http://jena.sourceforge.net/ontology>
- [16] Jena RDQL Documentation, [http://jena.sourceforge.net/RDQL/rdql\\_grammar.html](http://jena.sourceforge.net/RDQL/rdql_grammar.html)
- [17] Jena Semantic Web Framework, <http://jena.sourceforge.net/>
- [18] Jena Tutorial, A Programmer's Introduction to RDQL, <http://jena.sourceforge.net/tutorial/RDQL>
- [19] R. Mizoguchi, "Tutorial on ontological engineering - Part 2: Ontology development, tools and languages", <http://www.ei.sanken.osaka-u.ac.jp/pub/miz/Part2V3.pdf>
- [20] Naming and Addressing: URIs, URLs..., <http://www.w3.org/Addressing/>
- [21] A. Pease, "Why use OWL?", <http://www.xfront.com/why-use-owl.html>
- [22] Resource Description Framework (RDF), <http://www.w3.org/RDF-primer/>
- [23] A. Seaborne, "Jena Tutorial, A Programmer's Introduction to RDQL", <http://jena.sourceforge.net/tutorial/RDQL/index.html>
- [24] R. Seeley "The Semantic Web: The OWL has landed", ADT Magazine, <http://www.adtmag.com/article.asp?id=8144>
- [25] The World Wide Web Consortium, OWL, <http://www.w3.org/TR/2004/>
- [26] W3C Semantic Web, <http://www.w3.org/sw/>
- [27] W3Schools, [http://www.w3schools.com/rdf/rdf\\_owl.asp](http://www.w3schools.com/rdf/rdf_owl.asp)
- [28] Web Ontology Language (OWL), <http://www.w3.org/OWL/>
- [29] T. Berners-Lee, Semantic Web on XML, 2000, <http://www.w3.org/2000/Talks/1206-xml2k-tbl/slide1-0.html>
- [30] Unicode consortium, <http://www.unicode.org/>

- [31] Dubois, D. D. (1998) (Ed.) Preface. In Dubois, D. D. (Ed.), The Competency Casebook: Twelve Studies in Competency-Based Performance Improvement. Amherst, MA: HRD Press, Inc.
- [32] Marrelli, A. F. (1998). "An Introduction to Competency Analysis and Modeling." In Performance Improvement, 37: 8-17
- [33] Intagliata, J., Ulrich, D., & Smallwood, N. (2000). Leveraging leadership competencies to produce leadership brand: Creating distinctiveness by focusing on strategy and results. Human Resource Planning, 23(3), 12-23.
- [34] Dubois, D. (1993). Competency-based performance: A strategy for organizational change. Boston, MA: HRD Press
- [35] Lindgren, R., Stenmark, D. and Ljungberg, J. Rethinking Competence Systems for Knowledge-Based Organizations. Under consideration by European Journal of Information Systems. A previous version appeared in the Proceedings of ECIS 2001, Bled, Slovenia, pp. 775-786.
- [36] <http://www.w3.org/2001/sw/>
- [37] Lu Shiyong, Dong Ming, Fotouhi Farshad (2002), "The Semantic Web: opportunities and challenges for next-generation Web applications", Information Research. <http://informationr.net/ir/7-4/paper134.html#sax>
- [38] Miller Eric, "Enabling the Semantic Web for Scientific Research And Collaboration". [http://www.sis.pitt.edu/~dlwkshop/paper\\_miller.html](http://www.sis.pitt.edu/~dlwkshop/paper_miller.html)
- [39] Terence R., Zeng Marcia L. and the ADEPT Project Team, "Building Semantic Tools for Concept-based Learning Spaces: Knowledge Bases of Strongly-Structured Models for Scientific Concepts in Advanced Digital Libraries". <http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Smith/>
- [40] McGuinness Deborah L., "Ontologies Come of Age". [http://www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-\(with-citation\).htm](http://www.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-(with-citation).htm)
- [41] Lau Raymond, Hofstede Athur H.M., Chen Otto, "Logic-Based Adaptive Information Agents on the Web". [www10.org/cdrom/posters/1013.pdf](http://www10.org/cdrom/posters/1013.pdf)

- [42] **Phelps Thomas A., Wilensky Robert**, “Robust Hyperlinks and Locations”,  
<http://www.dlib.org/dlib/july00/wilensky/07wilensky.html>
- [43] UC Berkeley Digital Library Project: <http://elib.cs.berkeley.edu/>,  
<http://elib.cs.berkeley.edu/vision.html>
- [44] Gill Tony, “Building semantic bridges between museums, libraries and archives:  
The CIDOC Conceptual Reference”,  
[http://www.firstmonday.dk/issues/issue9\\_5/gill/#author](http://www.firstmonday.dk/issues/issue9_5/gill/#author) (Tε
- [45] (<http://apollo.open.ac.uk>)
- [46] (<http://www.landcglobal.com/pages/linkfactory.php>)
- [47] (<http://oiled.man.ac.uk/>)
- [48] <http://www.daml.org/2001/03/daml+oil.daml>
- [49] <http://www.research.att.com/areas/visualization/projects\ software/graphviz.php>
- [50] <http://www.ontoknowledge.org/tools/ontoedit.shtml>
- [51] <http://www.ksl.stanford.edu/software/ontolingua>
- [52] (<http://www.isi.edu/isd/ontosaurus.html>)
- [53] <http://topthing.com>
- [54] <http://protege.stanford.edu>
- [55] <http://webode.dia.fi.upm.es>
- [56] <http://www.ontoprise.de>
- [57] (<http://www.ontoprise.de/>)
- [58] <http://www.ebxml.org>