

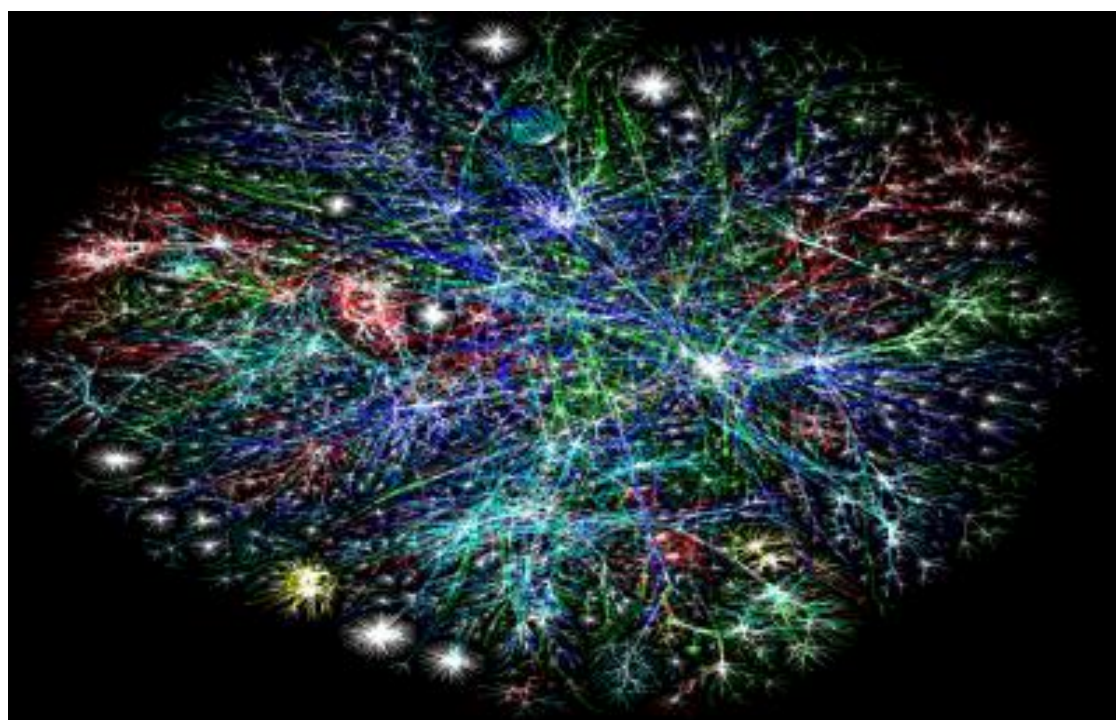


ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ (Τ.Ε.Ι.) ΔΥΤΙΚΗΣ
ΕΛΛΑΔΑΣ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΑΣ

Τμήμα Διοίκησης Επιχειρήσεων (Πάτρα)

Εισαγωγή στον Σημασιολογικό Παγκόσμιο Ιστό και Ανάλυση
οντολογιών OWL



ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ των:

Ελένη Κατσουλάκου

Αλέξανδρος Μαΐστρος

Στυλιάνα Τσουτσούρη

Επιβλέπων: Χαλκιάπουλος Κωνσταντίνος

ΠΑΤΡΑ, Ιούνιος 2015

ΕΙΣΑΓΩΓΗ ΣΤΟΝ ΣΗΜΑΣΙΟΛΟΓΙΚΟ ΠΑΓΚΟΣΜΙΟ ΙΣΤΟ
ΚΑΙ ΑΝΑΛΥΣΗ ΟΝΤΟΛΟΓΙΩΝ OWL

ΠΕΡΙΛΗΨΗ

Στόχος της εργασίας είναι η παρουσίαση της εξελεγκτικής πορείας του σημασιολογικού ιστού και των τάσεων που επικρατούν σχετικά με την εξέλιξη του.

Στο πρώτο κεφάλαιο γίνεται εκτενής αναφορά στον Σημασιολογικό Ιστό. Πιο συγκεκριμένα, επισημαίνονται τα στάδια εξέλιξής του, η δομή του και ο στόχος του. Ακολούθως, πραγματοποιείται η ανάλυση των Οντολογιών OWL, με επίκεντρο τις γλώσσες περιγραφής Οντολογίας και τις προγενέστερες αυτών. Το τρίτο κεφάλαιο εμπεριέχει τη διαχείριση των Οντολογιών σε συνδυασμό με τα Εργαλεία Ανάπτυξης και τις Τεχνολογίες Εξόρυξης της γνώσης των Οντολογιών. Η έρευνα ολοκληρώνεται με τα αποτελέσματα της μελέτης που πραγματοποιήθηκαν.

Λέξεις κλειδιά: σημασιολογικός ιστός, οντολογίες, OWL, εργαλεία ,συλλογιστική οντολογιών

ABSTRACT

The aim of this paper is to present the evolutionary course of the semantic web and trends on the development of.

The first chapter is a comprehensive reference to the Semantic Web. In particular, they marked the stages of evolution, the structure and the aim of. Next, place the analysis of Ontologies OWL, focusing on Ontology description language and the earlier these. The third chapter includes the management of Ontologies in combination with development tools and Mining Technologies knowledge of Ontology. The investigation is completed with the results of the study performed.

Key words: semantic web, ontologies, OWL, tools, reasoners

Περιεχόμενα

1.	ΕΙΣΑΓΩΓΗ	10
1.1	Αντικείμενο Μελέτης.....	10
1.2	Στόχος.....	11
1.3	Μεθοδολογία	11
2.	ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ ΙΣΤΟΣ	13
2.1	Εξελικτική Πορεία	13
2.2	Δομή του Σημασιολογικού Ιστού.....	16
2.3	Στόχος του Σημασιολογικού Ιστού	18
2.4	Σημασιολογικά Εργαλεία	19
2.5	Μεταδεδομένα.....	21
2.6	Οντολογίες.....	23
2.7	Συστατικά Οντολογιών	25
2.8	Κατηγορίες Οντολογιών.....	27
2.8.1	Βαθμός Τυπικότητας.....	27
2.8.2	Εσωτερική Δομή	28
2.8.3	Γενικότητα της Σύλληψης.....	29
2.9	Πράκτορες Σημασιολογικού Ιστού	30

2.10	Οι Μηχανές Αναζήτησης στο Σημασιολογικό Ιστό	31
2.11	Σημασιολογικός Ιστός και Ψηφιακές Βιβλιοθήκες	33
2.12	Εργαλεία Ανάπτυξης Οντολογιών.....	33
3.	WEB ONTOLOGY LANGUAGE (OWL)	35
3.1	Γλώσσες Περιγραφής Οντολογίας	35
3.2	Γλώσσα Οντολογίας Ιστού (OWL).....	37
3.3	Προγενέστερες Γλώσσες (XML, RDF).....	39
3.4	Χαρακτηριστικά της Γλώσσας Οντολογίας OWL	43
3.5	Υπο-Γλώσσες OWL	46
3.6	Σύγκριση της OWL με Προγενέστερες Γλώσσες (XML, RDF και RDFS)..	48
3.7	Έγγραφα OWL	49
3.6.1	Κεφαλίδα (Header)	50
3.6.2	Κλάσεις και στιγμιότυπα	50
3.6.3	Ιδιότητες.....	51
4.	ΔΙΑΧΕΙΡΙΣΗ ΟΝΤΟΛΟΓΙΩΝ	52
4.1	Εργαλεία Ανάπτυξης & Διαχείρισης Οντολογιών	52
4.2	Η Αρχιτεκτονική του Protege.....	54
4.2.1	Protege – Frames Editor.....	56
4.2.2	Protege – OWL Editor	56
4.3	Τεχνολογίες Εξόρυξης Οντολογικής Γνώσης	56

4.3.1	Γλώσσες Επερωτήσεων	57
4.3.2	Συστήματα μετατροπής σχημάτων γνώσης	58
4.3.3	Προγραμματιστικές Διεπαφές.....	59
4.3.4	Συστήματα Διαχείρισης Βάσεων Γνώσης.....	62
4.4	Συλλογιστική Οντολογιών (reasoners).....	64
4.4.1	Racer	65
4.4.2	FaCT++	66
4.4.3	Pellet	67
4.4.4	KAON2	67
4.5	Διαχείριση Οντολογικής Γνώσης με Υπηρεσίες Ιστού.....	68
5.	ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΛΕΤΗΣ.....	70
6.	ΒΙΒΛΙΟΓΡΑΦΙΑ	72

Ευρετήριο Εικόνων

Εικόνα 1 Τα στάδια της επιστημονικής έρευνας	12
Εικόνα 2 Χρονολογική εξέλιξη του web	15
Εικόνα 3 Τυπικότητα οντολογιών. Παραδείγματα τυπικών δομών.....	28
Εικόνα 4 Είδη οντολογιών	28
Εικόνα 5 Ταξινόμηση οντολογιών με βάση το επίπεδο της γενικότητας τους.....	30
Εικόνα 6 Γλώσσες περιγραφής οντολογίας	37
Εικόνα 7 Πρόγονοι της OWL (Πετράκης, 2011).....	38
Εικόνα 8 HTML vs. XML	40
Εικόνα 9 Γράφοι αναπαράστασης πληροφορίας σύμφωνα με το μοντέλο RDF.....	43
Εικόνα 10 (Α) Γράφος αναπαράστασης πληροφορίας - (Β) RDF/XML σύνταξη	43
Εικόνα 11 Η δημιουργία της γλώσσας OWL	45
Εικόνα 12 Η αρχιτεκτονική των συστημάτων του Protege	55
Εικόνα 13 Αρχιτεκτονική ενσωμάτωσης εργαλείων Jena και Protege.....	61
Εικόνα 14 Ρόλοι και Λειτουργίες Υπηρεσιών Ιστού	70

Ευρετήριο Πινάκων

Πίνακας 1 Συστατικά οντολογιών	25
Πίνακας 2 Βαθμοί τυπικότητας	27
Πίνακας 3 Σύστημα ταξινόμησης των Οντολογιών με βάση το επίπεδο γενικότητας	29
Πίνακας 4 Γλώσσες αναπαράστασης οντολογιών	35
Πίνακας 5 Υπο-γλώσσες OWL.....	46

1. ΕΙΣΑΓΩΓΗ

1.1 Αντικείμενο Μελέτης

Ο Σημασιολογικός Ιστός (Semantic Web, Web 3.0) αποτελεί την επέκταση του Παγκόσμιου Ιστού (World Wide Web, Web 2.0) που προσφέρει τη δυνατότητα στους ανθρώπους να μοιράζονται περιεχόμενο πέρα από τα όρια των εφαρμογών των ιστοσελίδων. Αυτό εξυπηρετεί τα αιτήματα των χρηστών από τη χρήση του περιεχομένου του Ιστού. Ωστόσο, οι σελίδες του Παγκόσμιου Ιστού αφορούν αποκλειστικά και μόνο τους ανθρώπους, με την έννοια ότι η πληροφορία που περιέχουν είναι αναγνώσιμη και κατανοητή μόνον από αυτούς.

Χωρίς τις μηχανές αναζήτησης (π.χ. Google, Yahoo κλπ.), η ανάπτυξη του Παγκόσμιου Ιστού δεν θα ήταν τόσο καταλυτική.

Ωστόσο, υπάρχουν ορισμένα προβλήματα που αφορούν τη χρήση τους. Ένα από αυτά είναι το γεγονός ότι τα αποτελέσματα που επιστρέφουν είναι απλές ιστοσελίδες. Στην περίπτωση που είναι απαραίτητα δεδομένα και πληροφορίες που είναι κατανεμημένες σε ξεχωριστά έγγραφα θα πρέπει να εκτελεστούν ποικίλες αναζητήσεις με διαφορετικές λέξεις-κλειδιά προκειμένου να ανακτήσουμε τις σχετικές πληροφορίες. Συνάμα, η μεγάλη «ευαισθησία» στο λεξιλόγιο των μηχανών αναζήτησης αποτελεί ένα σημαντικό πρόβλημα αυτών. Πολλές φορές τα έγγραφα σχετίζονται με τις λέξεις-κλειδιά δεν επιστρέφονται, διότι εφαρμόζουν διαφορετική ορολογία. Αυτό δεν είναι αποτελεσματικό, γιατί οι σημασιολογικά όμοιες αναζητήσεις θα έπρεπε να επιστρέφουν και ίδια αποτελέσματα. Τη λύση στα προβλήματα που αντιμετωπίζει ο Παγκόσμιος Ιστός έρχεται να δώσει ο Σημασιολογικός Ιστός. Ο Σημασιολογικός Ιστός στοχεύει στην εξέλιξη του σημερινού Διαδικτύου, ώστε τα δεδομένα που υπάρχουν και διακινούνται σε αυτό να είναι κατανοητές, και αυτόματα επεξεργάσιμες από τους υπολογιστές. Ο Σημασιολογικός Ιστός δεν αποτελεί νέο Παγκόσμιο Ιστό και είναι μια επέκταση και βελτίωση του σημερινού Ιστού στην κατεύθυνση, κυρίως, της δόμησης της πληροφορίας έτσι ώστε να είναι προσπελάσιμη από προγράμματα υπολογιστών. Η σημερινή αναπαράσταση των κειμένων στις σελίδες του Ιστού, που

προορίζεται για χρήση από ανθρώπους θα αντικατασταθεί από αναπαράσταση κατανοητή στους υπολογιστές.

1.2 Στόχος

Από το 1995 γίνεται μια έντονη ερευνητική προσπάθεια στον τομέα της αναπαράστασης της γνώσης που εκφράζεται με σημασιολογικό τρόπο, ώστε να είναι αναγνώσιμη και ανταλλάξιμη από μηχανές με τη χρήση έξυπνων εφαρμογών λογισμικού. **Στόχος της εργασίας είναι η παρουσίαση της εξελεγκτικής πορείας του σημασιολογικού ιστού και των τάσεων που επικρατούν σχετικά με την εξέλιξη του.**

1.3 Μεθοδολογία

Στο πρώτο κεφάλαιο γίνεται εκτενής αναφορά στον Σημασιολογικό Ιστό. Πιο συγκεκριμένα, επισημαίνονται τα στάδια εξέλιξής του, η δομή του και ο στόχος του. Ακολούθως, πραγματοποιείται η ανάλυση των Οντολογιών OWL, με επίκεντρο τις γλώσσες περιγραφής Οντολογίας και τις προγενέστερες αυτών. Το τρίτο κεφάλαιο εμπεριέχει τη διαχείριση των Οντολογιών σε συνδυασμό με τα Εργαλεία Ανάπτυξης και τις Τεχνολογίες Εξόρυξης της γνώσης των Οντολογιών. Η έρευνα ολοκληρώνεται με τα αποτελέσματα της μελέτης που πραγματοποιήθηκαν.

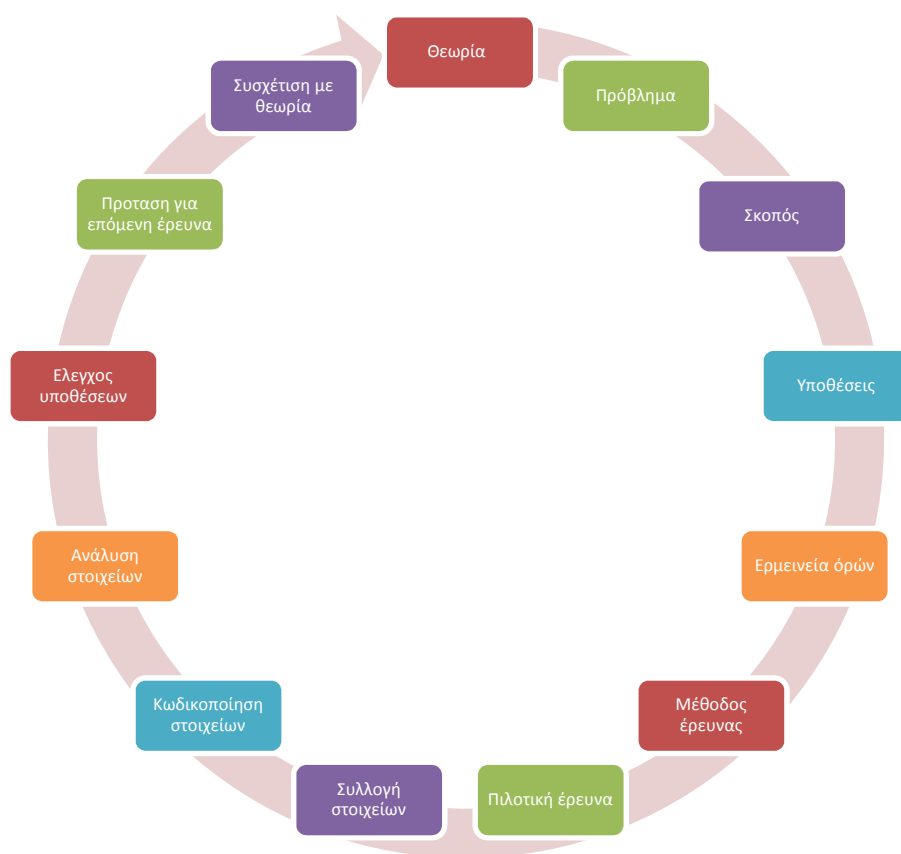
Αναφερόμενοι στη μεθοδολογία της παρούσης έρευνας, θα πρέπει να σημειωθεί πως σκοπός της έρευνας είναι η διευκόλυνση κατανόησης των φαινομένων, η πρόβλεψή τους και η δυνατότητα για τον έλεγχο τους. Σύμφωνα με τον Mouly (1970), έρευνα είναι μια διαδικασία που οδηγεί μέσα από προγραμματισμένη συστηματική συλλογή, ανάλυση κι ερμηνεία δεδομένων, στην αξιόπιστη λύση προβλημάτων.

Έτσι, το γενικό πλαίσιο της έρευνας περιλαμβάνει τη συλλογή δευτερογενών δεδομένων, ήτοι δεδομένων που έχουν «δημιουργηθεί» από κάποιον άλλον πλην του ερευνητή, χαρακτηριστικό παράδειγμα των οποίων είναι η βιβλιογραφία, καθώς και από τη συλλογή πρωτογενών δεδομένων, ήτοι δεδομένων που έχουν «δημιουργηθεί» από τον ερευνητή.

Η επιστημονική έρευνα, περιλαμβάνει συγκεκριμένα στάδια που ο ερευνητής πρέπει να ακολουθήσει για να καταλήξει σε έγκυρα και επιστημονικά ευρήματα. Τα στάδια αυτά περιγράφονται αναλυτικά παρακάτω στο σχήμα που ακολουθεί.

1. Καθορισμός προβλήματος
2. Επισκόπηση της βιβλιογραφίας
3. Διατύπωση Υποθέσεων
4. Επιλογή ερευνητικής μεθόδου
5. Έρευνα πεδίου
6. Ερμηνεία αποτελεσμάτων
7. Συγγραφή τελικής έκθεσης

Σε μία πιο αναλυτική επισκόπηση, τα βήματα που καλείται να ακολουθήσει ένας ερευνητής είναι αυτά που διακρίνονται στο παρακάτω σχήμα.



Εικόνα 1 Τα στάδια της επιστημονικής έρευνας

2. ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ ΙΣΤΟΣ

2.1 Εξελικτική Πορεία

Την πλειοψηφία της χρήσης του Διαδικτύου ως εργαλείο για να διευκολύνουν την επικοινωνία μεταξύ των ατόμων αποτέλεσαν για πολλά έτη τα email και τα usenet news¹. Ακολούθως, υπάρχουν πολυάριθμες καινοτόμες εφαρμογές σε αυτή την περιοχή που είναι γνωστές ως social software². Ο κύριος καταλύτης για την ανάπτυξη του Web 2.0 είναι το social software, μέσω του οποίου επιτράπηκε η επικοινωνία και η συνεργασία μεταξύ δύο ή περισσότερων υπολογιστών είτε αυτοί βρίσκονται στο ίδιο κτίριο είτε σε διαφορετική Ήπειρο. (Παπαταξιάρχης, 2008)

Σε ευρύτερο επίπεδο, ο όρος Web 2.0 επισημαίνεται σε μια δεύτερη γενιά των υπηρεσιών διαθέσιμων στο Internet (World Wide Web) που έδινε τη δυνατότητα στους ανθρώπους να συνεργαστούν και να μοιραστούν τις πληροφορίες on-line. Το Web 2.0, σε αντίθεση με την πρώτη γενιά, δίνει στους χρήστες μια εμπειρία πιο κοντά στις εφαρμογές υπολογιστών γραφείου από τις παραδοσιακές στατικές ιστοσελίδες. Συχνά, ο όρος εφαρμόζεται για να γίνει αναφορά σε μια τρέχουσα μετάβαση του World Wide Web από μια συλλογή ιστοσελίδων και ιστοχώρων σε μια ολοκληρωμένη υπολογιστική πλατφόρμα με εφαρμογές εξυπηρετώντας τους χρήστες του διαδικτύου. (Κάσσε, 2011)

Στην προσπάθεια να εντάξουμε τα δεδομένα και τις πληροφορίες σε μια σειρά, είναι χρήσιμο να σημειωθεί να ότι η θεματική και σχεδιαστική έναρξη του Web έγινε με το Web 1.0. Είναι δύσκολο να καθοριστεί η έννοια του web 1.0 από έναν γενικό ορισμό,

¹ με τον όρο usenet εννοούμε ένα σύστημα για την ανταλλαγή μηνυμάτων με τη μορφή άρθρων ταξινομημένα σύμφωνα με συγκεκριμένες κατηγορίες που ονομάζονται newsgroups

² Ο όρος Social Software ισχύει για οποιοδήποτε εργαλείο το οποίο επιτρέπει σε δύο ή περισσότερα άτομα να συνεργαστούν όντας κάθε πρόσωπο σε μια διαφορετική θέση. Η συνεργασία μπορεί να εμφανιστεί σε πραγματικό χρόνο ή σε διαφορετικούς χρόνους, ενώ οι θέσεις μπορούν να αφορούν ηπείρους ή απλά ένα κτήριο γραφείων. Ανεξαρτήτως της απόστασης ή της μορφής συνεργασίας, ο απώτερος στόχος του social software είναι η διαμόρφωση κοινοτήτων γνώσης στην οποία οι συμμετέχοντες δίνουν και λαμβάνουν συνεχώς πολύτιμες πληροφορίες.

αφού αποδόθηκε μετά την εισαγωγή του όρου web 2.0 από τον O'Reilly (2004), για να δηλώσει την κατάσταση του διαδικτύου πριν από αυτήν την περίοδο. Εξαιτίας της αναδρομικής του ισχύος, δεν αναφέρεται σε ένα γενικό σύνολο τεχνολογικής προόδου και σχεδιαστικού στυλ των ιστοσελίδων. Σε αντίθεση με το web 2.0 που αναφέρεται σε ένα σύνολο τεχνικών και τεχνολογικών εργαλείων για την υλοποίηση των ιστοσελίδων.

Πέραν αυτού, κάποιες από τις τεχνολογίες που χρησιμοποιούνται στο web 2.0 είναι αυτές που έχουν χρησιμοποιηθεί και στο web 1.0. Το web 1.0 και 2.0 είναι δύο έννοιες εξαρτημένες μεταξύ τους, και έτσι είναι δύσκολο να οροθετηθεί το χρονικό διάστημα στο οποίο αναφέρεται το web 1.0.

Εάν θεωρήσουμε πως το web 1.0 πρόκειται για μια δικτύωση αρχείων τότε το web 2.0 πρόκειται για μια δικτύωση ανθρώπων. Στο Web 1.0 δεν υπήρχε συγκεκριμένος χώρος για τον ερασιτέχνη user, ενώ για τον επαγγελματία, τον hacker και τον τελικό χρήστη οι ρόλοι ήταν πιο καθορισμένοι. Δεν υπήρχε κάποια διάδραση με τις σελίδες. Υπήρχε ο δημιουργός/admin/webmaster, ο οποίος διαχειριζόταν το περιεχόμενο, και οι χρήστες/clients οι οποίοι απλά διάβαζαν, χωρίς να τους δίνεται η δυνατότητα να δράσουν μαζί του. (Κουτσομητρόπουλος, 2004)

Το WEB 2.0 αντιπροσωπεύει μια τεχνολογική και κοινωνική επανάσταση, διότι πλέον ο αρχικός παθητικός του ρόλος παρουσίασης πληροφοριών μεταλλάσσεται. Έννοιες όπως διαδραστικότητα, δυναμικό περιεχόμενο, συνεργασία, συνεισφορά και κοινότητα, social computing διαδραματίζουν πλέον πρωταγωνιστικό ρόλο.

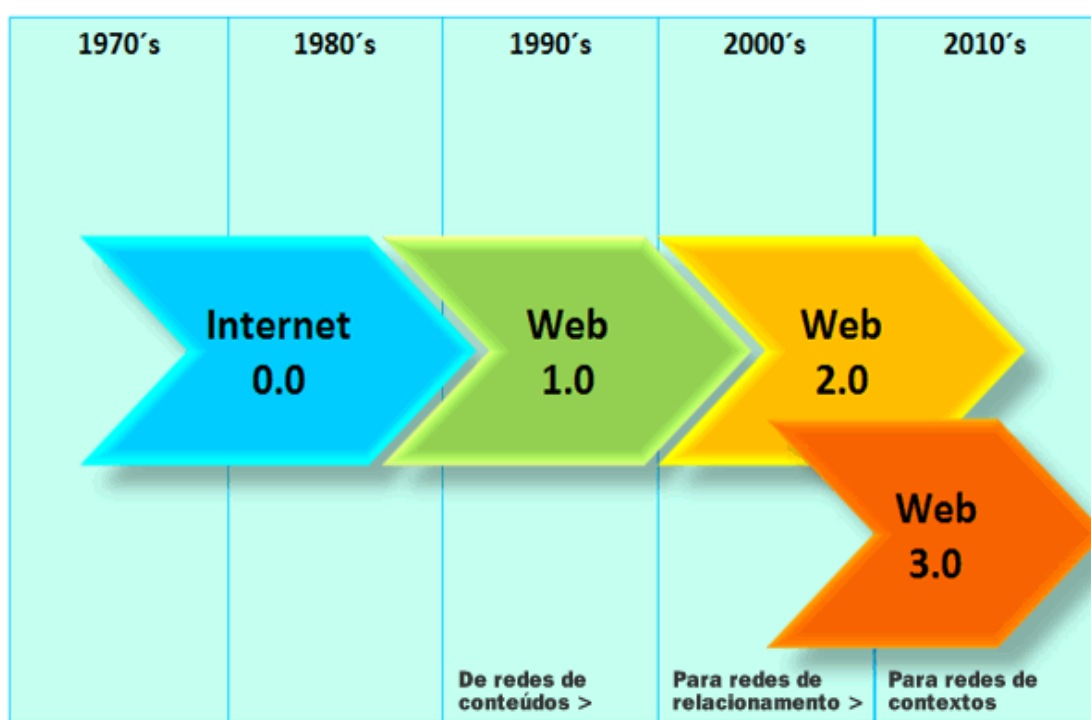
Ο όρος Web 2.0 αναφέρεται γενικά σε μια δεύτερη γενιά των υπηρεσιών διαθέσιμων στο Internet (World Wide Web) που αφήνει τους χρήστες να συνεργαστούν και να μοιραστούν τις πληροφορίες on-line. Σε αντίθεση με την πρώτη γενιά, το Web 2.0 δίνει στους χρήστες μια εμπειρία πιο κοντά στις εφαρμογές υπολογιστών γραφείου από τις παραδοσιακές στατικές ιστοσελίδες. (Berners-Lee, 2008)

Αν και ο όρος WEB 2.0 δίνει την αίσθηση ότι αποτελεί μια νέα έκδοση του WEB, ουσιαστικά δεν πρόκειται για κάποιο καινούργιο πρωτόκολλο του αλλά για αλλαγές στον τρόπο που χρησιμοποιούνται ήδη υπάρχουσες τεχνολογίες και στον τρόπο που οι σχεδιαστές πληροφοριακών συστημάτων και οι χρήστες χρησιμοποιούν το διαδίκτυο. Η κυρίαρχη τάση είναι να χρησιμοποιείται το Web 2.0 ως το μέσο

(πλατφόρμα) που εκτείνεται σε όλες τις συνδεδεμένες με αυτό συσκευές και τρέχουν οι εφαρμογές και υπηρεσίες, πολλές από τις οποίες μέχρι τώρα έτρεχαν τοπικά στους Η/Υ. (Πρέντζα, 2010)

Εφαρμογές Web 2.0 είναι εκείνες που εκμεταλλεύονται πλήρως τα εγγενή πλεονεκτήματα της πλατφόρμας αυτής. Οι εφαρμογές αυτές ενημερώνονται συνεχώς και όσο περισσότεροι άνθρωποι τις χρησιμοποιούν τόσο καλύτερες γίνονται χρησιμοποιώντας και συνδυάζοντας δεδομένα από πολλαπλές πηγές, συμπεριλαμβανομένων των μεμονωμένων χρηστών, παρέχοντας τα στοιχεία και τις υπηρεσίες τους σε μια μορφή που επιτρέπει αναδιαμόρφωση. Ως αποτέλεσμα, έχουμε μια διαφορετική εμπειρία από την πλευρά των χρηστών, συγκριτικά με το web 1.0 και την απλή μεταβίβαση ιστοσελίδων, μέσω της «αρχιτεκτονικής της συμμετοχής».

Όπως έχει ήδη τονιστεί το Web 2.0 επικεντρώνεται στη δημιουργία γνώσης και περιεχομένου από την πλευρά των χρηστών καθώς και η διαμοίρασή της με εργαλεία όπως τα Blogs και τα Wikis. Η Τρίτη γενιά το Web 3.0 επικεντρώνεται στην οργάνωση αυτής της γνώσης και την ενσωμάτωση αυτού του δυναμικού περιεχομένου (dynamic content) στον Παγκόσμιο Ιστό. (Berners-Lee, 2008)



Εικόνα 2 Χρονολογική εξέλιξη του web

Ένας σπουδαίος και φημισμένος φυσικός, ο Tim-Berners-Lee, ο οποίος εργαζόταν στο Κέντρο Πυρηνικών Ερευνών και Μελετών και υπήρξε ο πρωταρχικός εφευρέτης του World Wide Web (1989), είχε αναφέρει :

«Αυτό που προσπαθούμε να κάνουμε με το Sematic Web είναι να χτίσουμε μια στέρεα βάση συσχετισμών μεταξύ των δεδομένων που αντιλαμβάνονται οι μηχανές, κάτι που αντιστοιχεί στην ευφυΐα του αριστερού ημισφαιρίου του εγκεφάλου. Όταν λοιπόν καταφέρουμε να φτιάξουμε έναν τεράστιο όγκο δεδομένων, συνδυασμένων με έναν τρόπο ακριβή και αξιόπιστο – και θέλουμε τα δεδομένα αυτά να είναι οργανωμένα με έναν τέτοιο τρόπο -, θα έχουμε μια πολύ στέρεα βάση. Από εκεί και πέρα, θα είναι πολύ ενδιαφέρον και θα αποτελέσει μεγάλη πρόκληση, να διαμορφωθούν οι ευρετικές μέθοδοι που θα επιτρέψουν να κάνουμε δημιουργικές αναζητήσεις μέσα στα δεδομένα αυτά. Ωστόσο, δεν είναι από που προσπαθούμε να κάνουμε αυτή τη στιγμή. Προς το παρόν, μέλημά μας είναι η δημιουργία των σχέσεων μεταξύ των δεδομένων».

Όπως έχει ήδη επισημανθεί, ο Σημασιολογικός Ιστός δεν αποτελεί έναν νέο Παγκόσμιο Ιστό, αλλά πρόκειται για μια επέκταση και έναν ανασχηματισμό του σημερινού Ιστού, κυρίως της δόμησης της πληροφορίας, με απώτερο στόχο να είναι προσιτή από προγράμματα υπολογιστών. Η σημερινή αναπαράσταση των κειμένων στις σελίδες του Ιστού, που προορίζεται για εφαρμογή από τους ανθρώπους θα αντικατασταθεί από σαφή και προφανή αναπαράσταση στους υπολογιστές.

2.2 Δομή του Σημασιολογικού Ιστού

Ο Σημασιολογικός Ιστός είναι δομημένος σε μια σειρά επιπέδων. Τα επίπεδα αυτά αποτελούν περισσότερο τεχνολογικά επίπεδα λειτουργικότητας και δεν αντιστοιχούν αυστηρά στην έννοια της αρχιτεκτονικής λογισμικού. Με ένα ιεραρχικό σύστημα επιπέδων που το καθένα βρίσκεται πάνω από ένα άλλο και το επεκτείνει, συμβαδίζει η ανάπτυξη του Σημασιολογικού Ιστού. Το κάθε επίπεδο σχετίζεται με έναν

περιορισμένο τομέα του συνόλου των δραστηριοτήτων του Σημασιολογικού Ιστού και καθιστά παραγωγική τη συνεργασία μεταξύ των εκάστοτε επιστημονικών ομάδων. Εάν ο Σημασιολογικός Ιστός δεν ακολουθούσε αυτή την αρχιτεκτονική και ως ένα ενιαίο και ευρύ ερευνητικό πεδίο, πιθανότατα να υπάρχουν πολλές αποκλίσεις ανάμεσα στις απόψεις των επιστημόνων που συμμετέχουν στην προσπάθεια δημιουργίας των αναγκαίων προτύπων. (Herman, 2007)

Αποτελεί την πιθανότερη και επικρατέστερη εκδοχή της διαστρωμάτωσης, έτσι ακριβώς όπως την έχει περιγράψει ο ίδιος ο Tim Berners-Lee. Όπως αποδεικνύεται, ο Σημασιολογικός Ιστός παγιώνεται στην ήδη υπάρχουσα υποδομή του Παγκόσμιου Ιστού :

- Στο Πρωτόκολλο HTTP για τη μεταφορά των δεδομένων
- Στα URIs (Universal Resource Identifier), συμβολοσειρές που ταυτοποιούν μοναδικά κάθε οντότητα για την ονοματολογία
- Στην κωδικοποίηση Unicode για την πραγμάτωση της συνολικής προσπέλασης των δεδομένων.
- Στη γλώσσα σήμανσης XML (eXtensible Markup Language) ως κοινή συντακτική μορφή. Το πρότυπο XML συμφωνεί με την αναπαράσταση δομημένων εγγράφων με τη χρήση λεξιλογίων που ορίζονται από τους χρήστες και αποτελεί κατάλληλη λύση για τη μεταφορά αρχείων και εγγράφων στο χώρο του διαδικτύου.

Σχετικά με τις τεχνολογίες αυτές, ο Σημασιολογικός Ιστός παρουσιάζει κι άλλες κατανεμημένες σε συνολικά πέντε επίπεδα αρχιτεκτονικής, ορισμένες από τις οποίες έχουν ήδη υλοποιηθεί :

- Την τεχνολογία RDF (Resource Description Framework), ένα κύριο μοντέλο δεδομένων για τη δημιουργία απλών δηλώσεων για τα αντικείμενα του Ιστού που χρησιμοποιείται για την κοινή αναπαράσταση ημιδομημένων δεδομένων και μεταδεδομένων και την ανταλλαγή γνώσης στον χώρο του Διαδικτύου.
- Μια κοινή αναπαράσταση για τις οντολογίες, που εγκρίνουν τους όρους που χρησιμοποιούνται στο επίπεδο δεδομένων να θεσπίζονται και να συσχετίζονται μεταξύ τους (RDFS, DAML + OIL, OWL).

- Τη Λογική (Logic), που παρέχει το υπόβαθρο για την ικανότητα αυτοματοποιημένου συλλογισμού και συμπερασμών βάσει των πληροφοριών που δομούνται σε μια οντολογία. Το επίπεδο αυτό ισχυροποιείται από τη χρήση τυπικών κανόνων, βάσει των οποίων γίνεται εφικτή η ψευδο-νοήμων διαδικασία λήψης, αποφάσεων από τις υπολογιστικές μηχανές.
- Την Απόδειξη (Proof), προκειμένου τα αποτελέσματα που συμπεραίνονται από δεδομένα στον Σημασιολογικό Ιστό να μπορούν να οδηγήσουν πίσω στις υποθέσεις που τα προκάλεσαν.
- Την Εμπιστοσύνη (Trust), που συνδυαστικά με την τεχνολογία των ψηφιακών υπογραφών (digital signatures), θα παρέχει τον βαθμό στον οποίο οι πληροφορίες που διακινούνται, επεξεργάζονται και συμπεραίνονται στον Σημασιολογικό Ιστό, είναι αξιόπιστες.

2.3 Στόχος του Σημασιολογικού Ιστού

Απώτερος στόχος του σημασιολογικού διαδικτύου είναι η δυνατότητα των χρηστών και των πληροφοριακών συστημάτων να έχουν τη δυνατότητα να αντιλαμβάνονται και να επεξεργάζονται ανάλογα στοιχεία και δεδομένα. Θεωρείται δεδομένο, ότι η προοπτική αυτή θα διευρύνει τους ορίζοντες, αφού τα πληροφοριακά συστήματα θα μπορούν να επεξεργάζονται λογικά τις πληροφορίες και να εξάγουν συμπεράσματα. Οι συνέπειες της διεργασίας θα δημοσιεύονται στο διαδίκτυο, για να χρησιμοποιηθούν από τρίτους, αυξάνοντας με αυτόν τον τρόπο τη γνώση και την αξία του σημασιολογικού ιστού.

Ακολούθως, προχωρώντας ένα βήμα πιο πέρα, θα δημιουργηθούν έξυπνες υπηρεσίες, οι οποίες θα μπορούν να χρησιμοποιούνται μεταξύ τους και να συνεργάζονται, χωρίς να έχουν εξαρχής προετοιμαστεί για αυτό, (Στοϊλος, 2008) και όλα αυτά αυτόματα χωρίς τη συμμετοχή και τη παρέμβαση του ατόμου. (Κορμπάκης, 2009)

2.4 Σημασιολογικά Εργαλεία

Τα σημασιολογικά εργαλεία σχετίζονται με τα εργαλεία που επιχειρούν να καλύψουν το κενό στην επικοινωνία του ανθρώπου μηχανής. Παρόλο που οι περισσότερες από τις τρέχουσες μορφές περιεχομένου του ιστού σχεδιάζονται να είναι κατανοητές από τον άνθρωπο, αδυνατούν να είναι κατανοητές από τον Η/Υ. Σαφέστατα, ο σχεδιασμός του περιεχομένου του σημασιολογικού ιστού αποσκοπεί στο να το κατανοούν τόσο ο άνθρωπος όσο και οι Η/Υ. Η επίτευξη αυτού έγκειται στο ότι πραγματοποιείται μια προσπάθεια ανάπτυξης των εννοιών που να μπορεί να τις «κατανοήσει» ο Η/Υ. Αυτό πραγματοποιείται με την οργάνωση της επιστημονικής γνώσης και ορολογίας σε μια τέτοια που να σε συνδυασμό με ορισμένες τεχνικές να είναι κατανοητή από τον Η/Υ. Τέτοια εργαλεία είναι:

1. **Term lists:** Λίστες όρων όπως authority files, γλωσσάρια και λεξικά
2. **Metadata-like models:** όπως κατάλογοι και γεωγραφικά λεξικά
3. **Classification and categorization schemes:** Σχέδια ταξινόμησης και κατηγοριοποίησης, όπως θεματικές επικεφαλίδες, σχήματα ταξινόμησης και κατηγοριοποίησης
4. **Relationship models:** Σχεσιακά Μοντέλα όπως θησαυροί, σημασιολογικά δίκτυα, χάρτες εννοιών και οντολογίες
5. **Metadata content standards:** Πρότυπα για το περιεχόμενο των μεταδεδομένων και ειδικά αυτά που ασχολούνται με την αντιπροσώπευση της γνώσης

6. Domain-specific content mark-up languages: όπως DTD (Document Type Definition) και τα σχήματα XML

7. 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 αντικειμενοστραφείς γλώσσες διαμόρφωσης.

Πολυάριθμα από τα σημασιολογικά εργαλεία που επισημαίνονται και αναλύονται παραπάνω, ανήκουν στα συμβατικά συστήματα οργάνωσης της γνώσης, όπως οι Term lists, Classification and categorization schemes και ορισμένα από τα σχεσιακά μοντέλα. Σε μαθησιακά περιβάλλοντα έχει αξία η συμβατική αναπαράσταση εννοιών, καθώς παρέχουν πρόσβαση σε κείμενα και εφημερίδες, για τα οποία υπάρχει αναπαράσταση εννοιών που είναι επικεντρωμένη σε όρους. Πολλά από αυτά τα εργαλεία έχουν μια σημαντική αδυναμία σε περιβάλλον Ψηφιακών Βιβλιοθηκών, λόγω των σχετικά αδύναμων δομών τους. Ειδικότερα, καταλήγουν να λαμβάνουν τη μορφή γλωσσικών όρων. (Κορμπάκης, 2009)

Ωστόσο, πολλά από τα σημαντικά χαρακτηριστικά των επιστημονικών εννοιών με χαρακτηριστικά παραδείγματα την αναπαράστασή τους, τις έννοιες, τις ιδιότητες, τις σχέσεις με άλλες έννοιες δεν μπορούν να αναπαρασταθούν με τόσο απλούς γλωσσικούς όρους. Ο σημασιολογικός ιστός θα χρησιμοποιεί τα σημασιολογικά εργαλεία, όπως είναι οι οντολογίες, για να περιγράψει τους διάφορους πόρους στο δίκτυο. Συνεπώς, ο ιστός θα γίνει φορέας γνώσης, παρόλο που έως σήμερα το δίκτυο αποτελούσε έναν χώρο που περιείχε πολύ μεγάλη ποσότητα ανοργάνωτης πληροφορίας, πλέον. Αυτό θα πραγματοποιηθεί με την λογική δόμηση της πληροφορίας, του περιεχομένου δηλαδή του ιστού με ένα σημασιολογικό τρόπο, γεγονός που θα σημάνει αρκετές μεταβολές σε σχέση με τον τρόπο που λειτουργεί τώρα ο ιστός. (Πρέντζα, 2010)

2.5 Μεταδεδομένα

Τα μεταδεδομένα είναι «δεδομένα για τα δεδομένα». Πιο συγκεκριμένα, πρόκειται για μια συστηματική - δομημένη μέθοδο περιγραφής πόρων βελτιώνοντας την προσπέλαση τους. Πρόκειται για δομημένα δεδομένα, τα οποία περιγράφουν τα χαρακτηριστικά ενός αντικειμένου. Το πρόθεμα «μετα» υποδηλώνει την περιγραφή σε κάποιο ανώτερο επίπεδο ή πιο βασικό. Τα μεταδεδομένα είναι ο συνδυασμός κρίκων μεταξύ του περιεχομένου του πόρου και της σημασίας του περιεχομένου. Τα μεταδεδομένα ενός εγγράφου του ιστού μπορεί να περιλαμβάνουν : τον τίτλο του, το δημιουργό του, ημερομηνία δημιουργίας κ.α. Καθένας θα μπορούσε να προσθέσει τα δικά του χαρακτηριστικά στα μεταδεδομένα. Τότε το αποτέλεσμα θα ήταν κάθε έγγραφο του ιστού να έχει δική του δομή και έτσι να είναι αδύνατη η επεξεργασία των μεταδεδομένων από αυτόματα λογισμικά σε μεγάλη κλίμακα. Για τον εξής λόγο, έχουν δημιουργηθεί πρότυπα μεταδεδομένων που καθορίζουν ότι μια εγγραφή μεταδεδομένων αποτελείται από ένα πλήθος προ-ορισμένων στοιχείων που αναπαριστούν συγκεκριμένες ιδιότητες ενός αντικειμένου/πηγής και καθένα στοιχείο μπορεί να λάβει μια ή περισσότερες τιμές. Αυτού του είδους τα πρότυπα καλούνται σχήματα μεταδεδομένων. (Στασινόπουλος, 2008)

Ένας περιορισμένος αριθμός στοιχείων, τα ονόματά τους και η σημασία τους εμπεριέχονται σε κάθε σχήμα μεταδεδομένων, ως χαρακτηριστικά. Πολλές φορές, στη σημασιολογική περιγραφή γίνεται αναφορά στον χώρο ύπαρξης, στα περιεχόμενα, στις φυσικές ιδιότητες, στον τύπο (π.χ. κείμενο ή εικόνα, χάρτης ή μοντέλο) και στη μορφή (π.χ. εκτυπωμένο αντίγραφο, ηλεκτρονικό αρχείο). Χαρακτηριστικά στοιχεία μεταδεδομένων που υποστηρίζουν την πρόσβαση σε έγγραφα, περιλαμβάνουν τον συγγραφέα, τον χρόνο και τον τόπο που εκδόθηκε, καθώς και τις θεματικές ενότητες που καλύπτονται. Στην περίπτωση που το υλικό των πληροφοριών είναι σε μη ψηφιακή μορφή, όπως σε εκτυπωμένη έκδοση, παρέχονται κι άλλα μεταδεδομένα που βοηθούν στον εντοπισμό του υλικού αυτού, όπως για παράδειγμα τα τηλέφωνα βιβλιοθηκών κ.α. (Κάσσε, 2011)

Το Dublin Core θεωρείται ένα σχήμα μεταδεδομένων και περιλαμβάνει 13 στοιχεία, τα οποία αργότερα έφτασαν τα 15. Προτάθηκε ως ο ελάχιστος αριθμός

μεταδεδομένων που απαιτούνται, με σκοπό να γίνει ευκολότερη και πιο προσιτή η ανακάλυψη εγγράφων στο διαδίκτυο. Ορισμένα άλλα διαδεδομένα σχήματα μεταδεδομένων αποτελούν τα AACR2 (Anglo-American Catalogin Rules), GILS (Government Information Locator Service), EAD (Encoded Archives Description), IMS (IMS Global Learning Consortium), AGLS (Australian Government Locator Service)

Είναι εφικτή η σύνδεση των μεταδεδομένων με μια πηγή δεδομένων. Αυτή πραγματοποιείται με την ενσωμάτωσή τους μέσα στην ίδια την πηγή με χρήση ειδικών META-ετικετών, με ένα ξεχωριστό έγγραφο που έχει κάποιο σύνδεσμο προς αυτήν, καθώς και με το να βρίσκονται μέσα σε μια βάση δεδομένων και από εκεί να συνδέονται με την πηγή. Ο απλούστερος τρόπος είναι ο πρώτος και συγκεκριμένα οι δημιουργοί των σελίδων να προσθέτουν HTML <meta> ετικέτες στο <head>...</head> τμήμα της ιστοσελίδας κατά τη δημιουργία της. Αντίθετα, η επινόηση μεταδεδομένων κατευθείαν σε βάση δεδομένων και η σύνδεσή τους με την πηγή πληροφοριών, προόδευε ως μια δραστηριότητα αυτάρκη από τη δημιουργία των ίδιων των πηγών. Ταυτόχρονα, υπάρχουν εργαλεία που δημιουργούν μεταδεδομένα για μια τρέχουσα ιστοσελίδα. (Στασινόπουλος, 2008)

Σχετικά με το πρόβλημα της αξιοποίησης των μεταδεδομένων, αυτό αποδεικνύεται ότι βρίσκεται σε αδιέξοδο. Αυτό συμβαίνει, διότι όσοι εκδίδουν στοιχεία και λεπτομέρειες δεν επενδύουν στην παροχή μεταδεδομένων, και επιπρόσθετα οι μηχανές αναζήτησης δεν επιχειρούν να προσθέσουν μεταδεδομένα, αφού δεν υπάρχουν αρκετά διαθέσιμα. Παράλληλα, τις περισσότερες φορές έως και πάντα, κυριαρχεί το ερώτημα της εγκυρότητας των μεταπληροφοριών που εφοδιάζονται από τους εκδότες. Οι άνθρωποι αυτοί, πάντα θα επιχειρούν να επεκτείνουν την αξιολόγηση των δεδομένων τους από τις μηχανές αναζήτησης, εφαρμόζοντας θεμιτά και αθέμιτα μέσα. Αναμφισβήτητα, παρά τις αντιξοότητες και τις δυσχέρειες, τα μεταδεδομένα μπορούν να συμβάλλουν στην αντιμετώπιση της ανικανότητας των μηχανών αναζήτησης σχετικά με τον εντοπισμό των ιδανικότερων δεδομένων και να χορηγούν καλύτερα αποτελέσματα στους χρήστες τους. (Στεφανίδης Β., 2009)

2.6 Οντολογίες

Οι οντολογίες αποτελούν το δομικό στοιχείο του Σημασιολογικού Ιστού. Με τον όρο οντολογία νοείται η ακριβής περιγραφή πραγμάτων και εννοιών, καθώς και οι σχέσεις που υπάρχουν ανάμεσά τους.

Ο πιο φημισμένος ορισμός στην επιστήμη των υπολογιστών για την οντολογία, δόθηκε από τον Gruber και είναι ο εξής :

«An ontology is a formal, explicit specification of a conceptualization».

Στον ορισμό αυτό ο όρος «conceptualization» αναφέρεται σε ένα αφηρημένο μοντέλο ενός φαινομένου στον κόσμο. Μια οντολογία είναι ο προσδιορισμός των χαρακτηριστικών ενός τέτοιου μοντέλου, καθώς δεν είναι η τελική μορφή αναπαράστασης που θα ενσωματωθεί σε ένα σύστημα λογισμικού, αλλά απεναντίας ένα αφηρημένο μοντέλο που χρησιμοποιείται στα πρώτα στάδια της διαδικασίας ανάπτυξης λογισμικού. Η οντολογία είναι σαφής καθώς οι έννοιες που χρησιμοποιούνται και οι περιορισμοί στη χρήση τους ορίζονται ξεκάθαρα και με σαφήνεια αντί να ορίζονται έμμεσα στο λογισμικό. Εν τέλει, μια οντολογία είναι τυπική, διότι πρέπει να δύναται να διαβαστεί από μηχανές.

Ο ορισμός της οντολογίας που δίνει το W3C είναι ο ακόλουθος :

«Μία οντολογία ορίζει τους όρους που χρησιμοποιούνται για να περιγράψουν και να αναπαραστήσουν μια περιοχή γνώσης»

Αναλύοντας τον παραπάνω ορισμό εξάγεται το εξής συμπέρασμα :

- Μια οντολογία δεν μπορεί να αναπαραστήσει όλη τη γνώση, αλλά ένα συγκεκριμένο τομέα γνώσης. Ένας τομέας είναι μια συγκεκριμένη θεματική περιοχή ή μια σφαίρα γνώσης όπως π.χ. ιατρική, εκπαίδευση κ.τ.λ.
- Μια οντολογία περιέχει όρους και τις σχέσεις μεταξύ τους. Οι όροι καλύπτονται και κλάσεις ή έννοιες εναλλακτικά. Οι σχέσεις μεταξύ των

κλάσεων δηλώνονται χρησιμοποιώντας ιεραρχική δομή. Οι υπερκλάσεις καλύπτουν υψηλότερου επιπέδου έννοιες, ενώ οι υποκλάσεις πιο εξειδικευμένες έννοιες οι οποίες όμως κληρονομούν όλα τα χαρακτηριστικά αυτών των υπερκλάσεων.

- Μια οντολογία περιέχει ιδιότητες που περιγράφουν τα χαρακτηριστικά των εννοιών και συσχετίζουν διαφορετικές κλάσεις μεταξύ τους. Συνεπώς, οι σχέσεις μεταξύ των κλάσεων δεν είναι αποκλειστικά αυτές της υπερκλάσης και της υποκλάσης, αλλά και σχέσεις που εκφράζονται από τις ιδιότητες. Η οντολογία καθορίζει τον τύπο των τιμών των ιδιοτήτων.

Η οντολογία είναι κατανοητή από μηχανές αναπαράστασης της ορολογίας μιας θεματικής περιοχής και των σχέσεων μεταξύ των όρων που ανήκουν στη περιοχή. Το ερέθισμα για την εξέλιξη, την εφαρμογή και τη διάδοση των οντολογιών είναι η καταγραφή γνώσης με έναν ξεκάθαρο και διακριτό τρόπο, δυνατό για επεξεργασία και ανάγνωση τόσο από ανθρώπους όσο και από μηχανές. Πρωταρχικός και κύριος σκοπός της οντολογίας είναι η κωδικοποίηση ενός συγκεκριμένου πεδίου γνώσης, κατανοητό από τις μηχανές. Οι τεχνολογίες που εφαρμόζονται για την κατασκευή υπολογιστικών συστημάτων συχνά μεταβάλλονται, σε αντίθεση με την αναπαράσταση μιας θεματικής ενότητας με τη μορφή μιας οντολογίας, η οποία είναι πιο ανεπηρέαστη στις τεχνολογικές εξελίξεις και πιο ελεύθερη από περιορισμούς υλοποίησης των διάφορων εφαρμογών. Επομένως, η χρήση οντολογίας στοχεύει στη διαχρονικότητα των δεδομένων, την ανεξαρτησία των πληροφοριών από την εκάστοτε πραγμάτωση και την ανοικοδόμησή της σε διάφορες ξεχωριστές εφαρμογές. Ορισμένα από τα πλεονεκτήματα χρήσης των οντολογιών είναι τα ακόλουθα :

- Παροχή ορισμών για συγκεκριμένους έννοιες σε ένα πεδίο ενδιαφέροντος
- Παροχή ενός τρόπου ανακύκλωσης της γνώσης του πεδίου
- Παροχή ενός τρόπου κωδικοποίησης της γνώσης ώστε να είναι κατανοητή από μηχανές

- Σε συνεργασία με τις γλώσσες περιγραφής οντολογιών καθιστούν εφικτή την αυτόματη μηχανική επεξεργασία των δεδομένων και της σημασιολογίας τους σε μεγάλη κλίμακα

Αναγκαίο βήμα για κάποιον που επιθυμεί να εφαρμόσει μια οντολογία είναι αρχικά ο έλεγχος μεταξύ των υπάρχοντων οντολογιών στον παγκόσμιο Ιστό. Έτσι μπορεί να αποφευχθεί ο κόπος σχεδιασμού μιας νέας οντολογίας και συνάμα, να υπάρξει πιθανότατα κέρδος από την ανταλλαγή γνώσεων μεταξύ των διάφορων εφαρμογών που ήδη τη χρησιμοποιούν και της νέας υπό ανάπτυξης εφαρμογής.

Σε γενικό πλαίσιο, ο βαθμός επαναχρησιμοποίησης μιας οντολογίας εξαρτάται και από το πόσο γενικευμένη είναι. Πιο συγκεκριμένα, όσο πιο γενική είναι μια έννοια που περιγράφεται, τόσο πιο πιθανή είναι η αναπροσαρμογή της από διαφορετικές εφαρμογές. Καθώς οι γενικοί όροι τοποθετούνται ψηλά στα επίπεδα της ιεραρχικής κληρονομικότητας, οι συλλογές που καταγράφουν τέτοιες έννοιες αποκαλ'πτονται ως άνω οντολογίες. Ορισμένες από τις γνωστότερες είναι οι CYC, OpenCYC, Sumo < Wordnet, με την τελευταία να είναι περισσότερο χρήσιμη στην κατασκευή οντολογιών παρά ως οντολογία για απευθείας χρήση σε κάποια εφαρμογή.

2.7 Συστατικά Οντολογιών

Οι σύγχρονες οντολογίες έχουν πολλές δομικές ομοιότητες, ανεξάρτητα από τη γλώσσα στην οποία εκφράζονται. Οι περισσότερες οντολογίες περιγράφουν άτομα-στιγμιότυπα (individuals-instances), κλάσεις (έννοιες) (classes-concepts), ιδιότητες (attributes), και σχέσεις (relations). Στη συνέχεια, περιγράφεται κάθε ένα από αυτά τα συστατικά. Τα συνήθη συστατικά των οντολογιών παρουσιάζονται στον ακόλουθο πίνακα.

Πίνακας 1 Συστατικά οντολογιών

Άτομα (Individuals)	περιπτώσεις (instances) ή αντικείμενα (τα βασικά)
---------------------	---

	ή «πρωταρχικού επιπέδου» αντικείμενα).
Κλάσεις (Classes)	σύνολα, συλλογές, έννοιες, τύποι αντικειμένων, ή είδη πραγμάτων.
Ιδιότητες (Attributes)	πτυχές, ιδιότητες, χαρακτηριστικά γνωρίσματα, ή παράμετροι που τα αντικείμενα (και οι κλάσεις) μπορούν να έχουν.
Σχέσεις (Relations)	τρόποι με τους οποίους οι κλάσεις και τα στιγμιότυπα μπορούν να αλληλοσυσχετίζονται.
Όροι λειτουργικότητας (Function terms)	σύνθετες δομές που διαμορφώνονται από συγκεκριμένες σχέσεις, και μπορούν να χρησιμοποιηθούν αντί ενός στιγμιότυπου σε μια δήλωση.
Περιορισμοί (Restrictions)	τυπικά δηλωμένες περιγραφές του τι πρέπει να ισχύει έτσι ώστε κάποιος ισχυρισμός να γίνεται αποδεκτός ως είσοδος.
Κανόνες (Rules):	δηλώσεις στη μορφή μιας «if-then» πρότασης, που περιγράφουν τα λογικά συμπεράσματα που μπορούν να προκύψουν από έναν ισχυρισμό σε μια ιδιαίτερη μορφή.
Αξιώματα (Axioms)	ισχυρισμοί (συμπεριλαμβανομένων και των κανόνων) διατυπωμένοι σε μια λογική μορφή, που στο σύνολό τους εμπεριέχουν τη γενική θεωρία που η οντολογία περιγράφει στο πεδίο εφαρμογής της.
Γεγονότα (Events)	η αλλαγή των ιδιοτήτων ή των σχέσεων

Οι οντολογίες κωδικοποιούνται συνήθως χρησιμοποιώντας γλώσσες οντολογίας (ontology languages). (όπως αναλύονται στο κεφάλαιο 3.1)

2.8 Κατηγορίες Οντολογιών

Ο Gomez-Perez (1999) παρουσιάζει ένα μικρό σύνολο των τύπων των οντολογιών που έχουν παρουσιαστεί στη βιβλιογραφία λόγω του μεγάλου αριθμού τους. Κάτι τέτοιο είναι λογικό καθώς δεν μπορεί να υπάρξει μόνο ένα είδος Οντολογιών πράγμα που φαίνεται και από το πλήθος των διαφορετικών ορισμών που υπάρχουν. Αφορμή για την ομαδοποίηση των Οντολογιών μπορεί να είναι το μέγεθος, η δομή, το αντικείμενο και ο σκοπός που καλύπτει η σύλληψη. Παραδείγματα αυτών είναι οι οντολογίες Αναπαράστασης Γνώσης, οι Οντολογίες Εφαρμογής, οι Μετά – Οντολογίες. Αυτή η ενότητα θα σταθεί σε τέσσερις κατηγορίες που έχουν να κάνουν με την τυπικότητα, τις γλώσσες, την εσωτερική δομή και την γενικότητα των Οντολογιών.

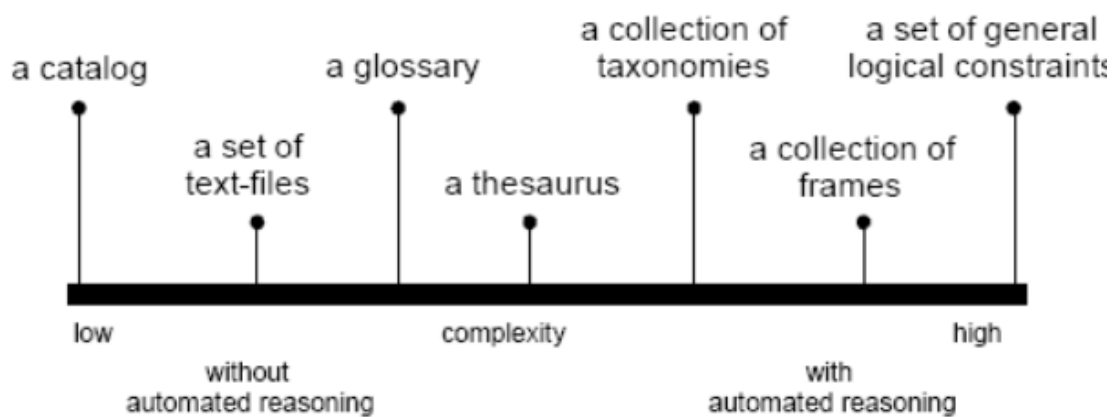
2.8.1 Βαθμός Τυπικότητας

Αναφέρεται στον βαθμό τυπικότητας με τον οποίο ένα λεξιλόγιο δημιουργείται και παίζει πολύτιμο ρόλο στην περίπτωση που οι όροι του χρειάζεται να επαναχρησιμοποιηθούν ή να επεκταθούν αλλά και στην περίπτωση συμπερασμού. Θεωρώντας την τυπικότητα ως ένα συνεχές ευθύγραμμο τμήμα τα δύο του άκρα θα ήταν υψηλή άτυπη και αυστηρά τυπική.

Πίνακας 2 Βαθμοί τυπικότητας

Υψηλά άτυπη	Εκφράζεται με μια χαλαρή φυσική γλώσσα, όπως καταλόγους, γλωσσάρια, θησαυρούς.
Ημι – άτυπη	Εκφράζεται σε μια κλειστή και δομημένη φυσική γλώσσα χαρακτηριστικό της είναι ότι πρέπει να αυξήσει τη σαφήνεια

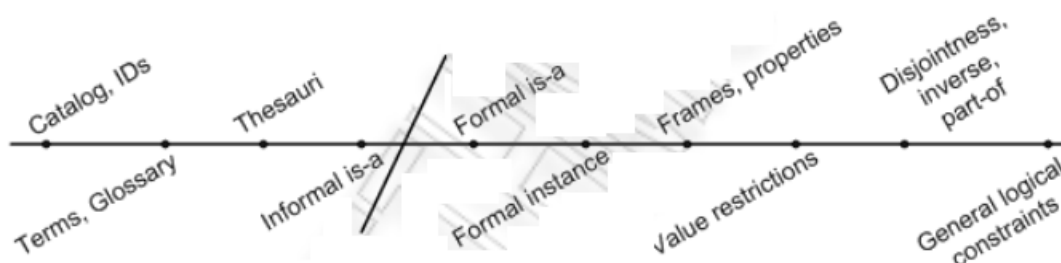
	μειώνοντας την ασάφεια.
Ημι – τυπική	Εκφράζεται μέσω μιας τεχνητής τυπικά ορισμένης γλώσσας, όπως με αυστηρές ιεραρχικές υποκλάσεις.
Αυστηρά Τυπική	Εκφράζεται με σχολιαστικό καθορισμό όρων και ιδιοτήτων, με τυπική σημασιολογία, θεωρήματα και αποδείξεις για την ορθότητα και την πληρότητα τους, όπως τον καθορισμό ιδιοτήτων των εννοιών, περιορισμούς και αξιώματα



Εικόνα 3 Τυπικότητα οντολογιών. Παραδείγματα τυπικών δομών

2.8.2 Εσωτερική Δομή

Με βάση την κατηγοριοποίηση κατά τυπικότητα οι Lassila και Swick εξετάζουν συγκεκριμένες δομές που έχουν υλοποιηθεί. Εδώ παρόλο που η κατηγοριοποίηση συνάδει πλήρως με τις παραπάνω περιγραφές, δίνεται βάρος στις διαφορετικές δομές που υπάρχουν.



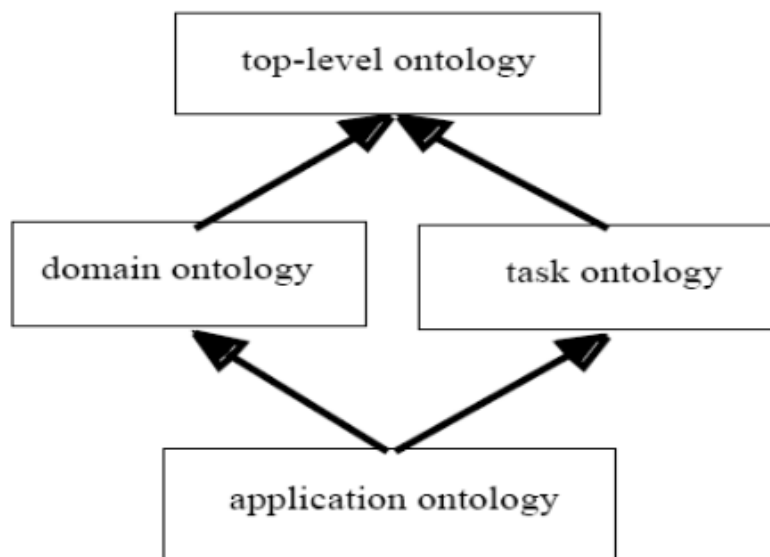
Εικόνα 4 Είδη οντολογιών

2.8.3 Γενικότητα της Σύλληψης

Ο Guarino το 1998 παρουσίασε ένα σύστημα ταξινόμησης των Οντολογιών με βάση το επίπεδο γενικότητας των αντικειμένων που διαπραγματεύονται. Κατατάσσει τρία επίπεδα, όπως αναλύονται στον ακόλουθο πίνακα :

Πίνακας 3 Σύστημα ταξινόμησης των Οντολογιών με βάση το επίπεδο γενικότητας

Οντολογίες Υψηλού Επιπέδου	Περιγράφουν πολύ γενικές έννοιες και είναι κοινές για όλα τα επίπεδα γνώσης. Τέτοιες μπορεί να είναι ο χρόνος, ο χώρος, τα γεγονότα, οι ενέργειες, τα αντικείμενα, αλλά και άλλες αφηρημένες έννοιες οι οποίες είναι ανεξάρτητες του προβλήματος ή του πεδίου εφαρμογής.
Οντολογίες Πεδίου και Οντολογίες Διεργασιών	Περιγράφουν αντίστοιχα το λεξιλόγιο ή τις διεργασίες και τις δραστηριότητες μέσα σε ένα συγκεκριμένο τομέα, οι οποίες ουσιαστικά είναι μια εξειδίκευση των εννοιών από τις Οντολογίες Υψηλότερου Επιπέδου.
Οντολογίες Εφαρμογών	Περιγράφουν έννοιες που εξαρτώνται τόσο από ένα συγκεκριμένο πεδίο αλλά και από συγκεκριμένες διαδικασίες, που ουσιαστικά είναι μια ειδικότερη περιγραφή των άλλων δύο σχετικών οντολογιών.



Εικόνα 5 Ταξινόμηση οντολογιών με βάση το επίπεδο της γενικότητας τους

2.9 Πράκτορες Σημασιολογικού Ιστού

Οι *πράκτορες* θεωρούνται εφαρμογές λογισμικού. Πιο συγκεκριμένα, αφορούν προγράμματα που επιτελούν ορισμένες λειτουργίες, χωρίς αναγκαία την επίβλεψη του χρήστη. Εκείνος λαμβάνει αποτελέσματα μετά το πέρας της εκτέλεσης αυτών των λειτουργιών. Οι πράκτορες μέχρι τώρα χρησιμοποιούνται με δύο τρόπους.

Στον πρώτο, οι πράκτορες περιηγούνται στο Διαδίκτυο, επεξεργάζονται και εξάγουν πληροφορίες από τις ιστοσελίδες που επισκέπτονται ενώ στον δεύτερο, λειτουργούν ως διαμεσολαβητές (mediators), μεταξύ ανόμοιων πηγών πληροφόρησης με αυστηρά καθορισμένη σύνταξη και σημασιολογία ανταλλαγής πληροφοριών.

Στον Σημασιολογικό Ιστό οι λειτουργίες που θα εκτελούν οι προσωπικοί πράκτορες διαδοχικά συνοψίζονται ως εξής: Οι πράκτορες πρώτα θα δέχονται κάποια αιτήματα εξυπηρέτησης από τους χρήστες, έπειτα θα αναζητούν πληροφορίες από τις πηγές του Διαδικτύου, θα επικοινωνούν με άλλους πράκτορες, θα συγκρίνουν τις πληροφορίες σχετικά με τα αιτήματα και τις προτιμήσεις των χρηστών, θα επιλέγουν μεταξύ των εναλλακτικών λύσεων και τέλος θα δίνουν τη λύση στον χρήστη ο οποίος θα λαμβάνει και την τελική απόφαση. (Παπαταξιάρχης, 2008)

Ο ρόλος των πρακτόρων στο Σημασιολογικό Ιστό δεν θα είναι ούτε η αντικατάσταση της ανθρώπινης βούλησης, ούτε και η λήψη τελικών αποφάσεων. Ο χρήστης θα είναι εκείνος που θα δίνει εντολές και θα παίρνει αποφάσεις. Οι πράκτορες θα συγκεντρώνουν και θα οργανώνουν πληροφορίες καθώς επίσης θα παρουσιάζουν τις όποιες επιλογές θα υπάρχουν σε κάθε περίπτωση έτσι ώστε ο χρήστης να διευκολυνθεί στην τελική απόφαση που θα λάβει. Η εξεύρεση τέτοιων πληροφοριών θα μπορούσε να γίνει καθημερινή πρακτική αρκεί οι δυνατότητες των διαδικτυακών υπηρεσιών και των πρακτόρων να εκφράζονται με τέτοιον τρόπο ώστε να είναι χρησιμοποιήσιμες και από άλλους πράκτορες. Για αυτόν τον λόγο χρησιμοποιείται από τους πράκτορες και η τεχνολογία των οντολογιών.

Επιπρόσθετα, οι πράκτορες του Σημασιολογικού Ιστού χρησιμοποιούν μεταδεδομένα, με απώτερο στόχο να εξάγουν τα δεδομένα που βρίσκουν στον Ιστό, αλλά και τη λογική για να βγάλουν αποτελέσματα και το κύριο πόρισμα. Ακόμη, είναι αναγκαία μια γλώσσα επικοινωνίας πρακτόρων, η OWL-S. Τέλος, οι πράκτορες θα θεωρούνται ολοένα και περισσότερο χρήσιμοι κατά την πλοήγηση στο Διαδίκτυο, καθώς το περιεχόμενο αυτού θα μεταβάλλεται και θα γίνεται πιο προσιτό και επεξεργάσιμο από τα προγράμματα αυτά.

2.10 Οι Μηχανές Αναζήτησης στο Σημασιολογικό Ιστό

Ο σημασιολογικός ιστός θα αποτελέσει καινοτομία στον τομέα των μηχανών αναζήτησης. Με αποδοτικότερη διαχείριση της πληροφορίας μέσω των θησαυρών και οντολογιών, δύναται να υπάρξουν καλύτερα και ποιοτικότερα αποτελέσματα χωρίς να φορτώνουν τόσο πολύ το δίκτυο. Συνάμα, ο χρήστης θα έχει τη δυνατότητα να διαμορφώσει πιο ελεύθερα ερωτήματα, όχι λέξεις – κλειδιά ή με τελεστές Boolean και η μηχανή αναζήτησης να του επιφέρει ικανοποιητικά αποτελέσματα, καθώς θα τα κατανοεί. Η χρήση της XML και των οντολογιών θα μπορέσει να το επιτύχει, καθώς με αυτές θα μπορεί η μηχανή να κατανοήσει και να αξιολογήσει λογικά το περιεχόμενο μιας σελίδας. Ταυτόχρονα, θα έχει την ικανότητα να αντιληφθεί τα ερωτήματα σημασιολογικά και όχι απλά σαν λέξεις – κλειδιά. Συνεπώς, αφορά τη μεγαλύτερη ακρίβεια στην ανάκτηση και πιο ποιοτικά αποτελέσματα. Αυτή τη στιγμή

οι μηχανές αναζήτησης έχουν δύο προσεγγίσεις στον ιστό. Εκείνη των μηχανών μεγάλης κλίμακας που στηρίζονται σε έξυπνους πράκτορες και αυτή των μηχανών μικρής κλίμακας που βασίζονται στην ανασκόπηση σελίδων.

Σχετικά με την πρώτη περίπτωση, διαπιστώνεται μεγάλη ανάκτηση, αλλά μικρή ακρίβεια. Αυτή 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.

Στην δεύτερη περίπτωση, τα αποτελέσματα φανερώσουν μεγάλη σχετικότητα, επειδή έχει γίνει ανασκόπηση της κάθε σελίδας. Έτσι, λοιπόν, υπάρχει μεγάλη ακρίβεια. Όμως, όπως συνεπάγεται, αυτή η δουλειά γίνεται σε περιορισμένο όγκο ιστοσελίδων, και έχουμε μικρό ποσοστό ανάκτησης.

Οι λέξεις – κλειδιά αποτελούν την βάση των δύο ειδών μηχανών αναζήτησης. Αυτό σημαίνει ότι και στις δύο περιπτώσεις παρουσιάζεται πρόβλημα, αναφορικά με το ερώτημα που διατυπώνεται, εξαιτίας των πολλαπλών ερμηνειών και σημασιών που μπορεί να έχει μια λέξη. Προκειμένου να ξεπεραστεί αυτό το πρόβλημα, έχουν αναπτυχθεί διάφορες τεχνικές. Ειδικότερα, ένα παράδειγμα είναι οι αλγόριθμοι που προσπαθούν να καθορίσουν τις συνώνυμες λέξεις, αλλά συναντάται πάντα το πρόβλημα ότι ποικίλες λέξεις με άσχετη ερμηνεία που πιθανόν προέρχονται από κοινή ρίζα ή κοντινές λέξεις σημασιολογικά, να έχουν εντελώς άσχετη ρίζα.

Η λύση σε αυτό το ζήτημα τίθεται με προφανή τρόπο. Είναι απαραίτητο να υπάρχουν semantics στις ιστοσελίδες, για να καθορίζεται το νόημά τους, αν δεν έχει επικρατήσει κάποια γλώσσα που να καθιερώνει κάποιες σταθερές HTML ετικέτες. Η XML παρουσιάζει πολλά πλεονεκτήματα, αφού κρατάει το περιεχόμενο, την δομή και την παρουσίαση χωριστά και είναι πολύ πιο επαρκής για την αναπαράσταση της γνώσης. Ωστόσο, την περιορίζει η συντακτική δομή της, καθώς έχει τη δυνατότητα να διαχειριστεί ορισμένες μόνο σημασιολογικές ιδιότητες. Για την πιο κατανοητή περιγραφή των ιστοσελίδων θα εισχωρήσουν οι οντολογίες σε αυτόν τον τομέα.

2.11 Σημαιολογικός Ιστός και Ψηφιακές Βιβλιοθήκες

Είναι γνωστό, ότι οι βιβλιοθήκες είναι ο φορέας της οργανωμένης γνώσης. Στο νέο περιβάλλον του σηματολογικού ιστού είναι απαραίτητο να προσφέρεται η εμπειρία που διαθέτουν για την οργάνωση της γνώσης από το συμβατικό περιβάλλον αλλά και η αξιοποίηση όλων των νέων εργαλείων, προκειμένου να συνεχίσουν να πραγματώνουν και να επιτελούν το κοινωνικό τους έργο διατηρώντας τα ίδια υψηλά επίπεδα ποιότητας. Έχουν την υποχρέωση προς την κοινωνία να φέρουν σε επαφή το συμβατικό περιβάλλον λειτουργίας τους με το ψηφιακό.

Η κύρια αλλαγή που συναντάται στις Ψηφιακές Βιβλιοθήκες είναι η αλλαγή υπόστασης του τεκμηρίου, δηλαδή έχει περάσει από την υλική στην ψηφιακή. Επιπρόσθετα, έχουν μεταβληθεί τα δεδομένα πρόσβασης στο υλικό τους, το περιεχόμενο των συλλογών τους και το οικονομικό τους περιβάλλον. Πλέον, οι συλλογές τους αποτελούνται από ψηφιακά πολυμεσικά αντικείμενα σε διάφορες formats, ηλεκτρονικά βιβλία και περιοδικά, videos, φωτογραφίες κ.ά. Αναμφίβολα, έχει αλλάξει και το κοινό. Πρόκειται για ένα παγκόσμιο κοινό που μπορεί να ζητά πρόσβαση στην πληροφορία από κάθε μέρος του πλανήτη, γεγονός που θέτει και το ζήτημα της πολυγλωσσικότητας. Σαφέστατα, το οικονομικό περιβάλλον είναι διαφορετικό, τώρα πια η βιβλιοθήκη εκτός από κοινωφελής οργανισμός πρέπει να είναι και βιώσιμος, παρέχοντας υπηρεσίες προστιθέμενης αξίας. Ο σηματολογικός ιστός παρουσιάζει νέες δυνατότητες οργάνωσης, διαχείρισης και διάθεσης των συλλογών τους. Ειδικότερα, με την χρήση των δυνατοτήτων που προσφέρουν τα σηματολογικά εργαλεία, μπορούν να παρέχουν νέες υπηρεσίες, γρηγορότερα και με μεγαλύτερη αξιοπιστία, με την ελάχιστη δυνατή ανθρώπινη παρέμβαση. Αυτά που προαναφέρθηκαν σχετικά με τον σηματολογικό ιστό δύναται να εφαρμοστούν άμεσα στις ψηφιακές βιβλιοθήκες.

2.12 Εργαλεία Ανάπτυξης Οντολογιών

Εκτός από τις γλώσσες ορισμού οντολογιών, υπάρχουν και εργαλεία με γραφικό περιβάλλον που κάνουν δυνατή τη δημιουργία οντολογιών. Ορισμένα από αυτά είναι:

Protégé, WebODE, Swoop κ.α. Το Protégé είναι ευρέως διαδεδομένο ως ένα εργαλείο δημιουργίας οντολογιών, το οποίο «γεννήθηκε» στο πανεπιστήμιο του Stanford.

Με τη βοήθεια ενός γραφικού περιβάλλοντος, δίνει τη δυνατότητα στη δημιουργία σύνθετων οντολογιών και την περιήγησή τους. Η ευρεία χρήση του Protégé σε συνδυασμό με την ανοιχτή του αρχιτεκτονική έχει ανοίξει το δρόμο προς τη δημιουργία ενός μεγάλου αριθμού από plug-ins που επεκτείνουν τις δυνατότητές του και το καθιστούν ένα ολοκληρωμένο πεδίο ανάπτυξης και διαχείρισης οντολογιών.

3. WEB ONTOLOGY LANGUAGE (OWL)

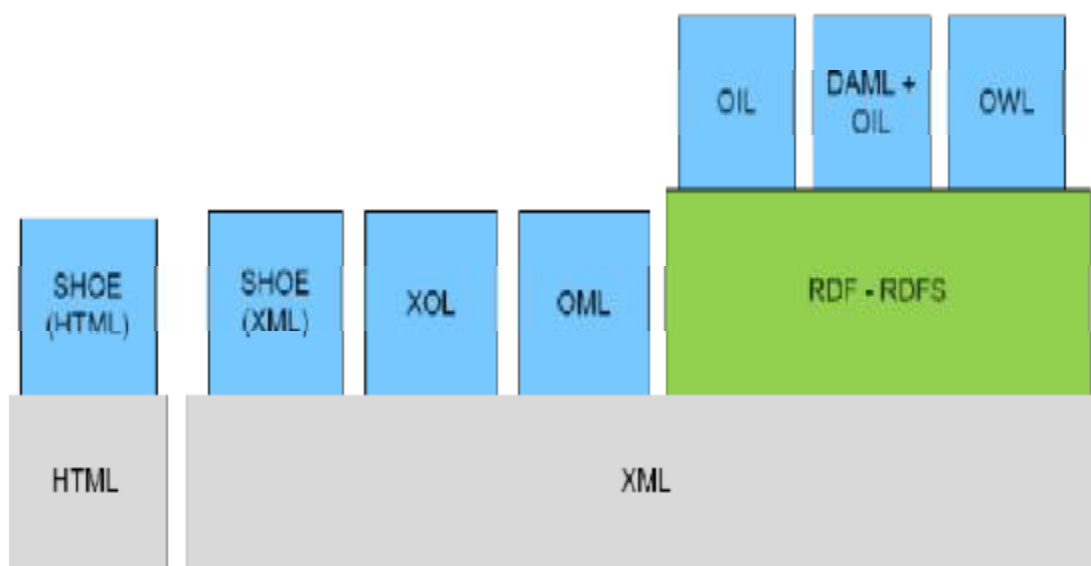
3.1 Γλώσσες Περιγραφής Οντολογίας

Στην επιστήμη της πληροφορικής, οι γλώσσες οντολογίας είναι τυπικές γλώσσες που χρησιμοποιούνται για την κωδικοποίηση οντολογιών. Επιτρέπουν την κωδικοποίηση της γνώσης για συγκεκριμένα πεδία γνώσης και περιλαμβάνουν συχνά κανόνες συλλογισμού που υποστηρίζουν την επεξεργασία εκείνης της γνώσης. Οι γλώσσες οντολογίας είναι συνήθως δηλωτικές γλώσσες, είναι σχεδόν πάντα γενικεύσεις των γλωσσών πλαισίων (frame language), και είναι συνήθως βασισμένες είτε στη λογική πρώτης τάξης είτε στη λογική περιγραφής. Σήμερα, υπάρχουν αρκετές γλώσσες αναπαράστασης οντολογιών και μπορούν να διαιρεθούν στους εξής κύριους τύπους :

Πίνακας 4 Γλώσσες αναπαράστασης οντολογιών

Παραδοσιακές γλώσσες	Οι γλώσσες αυτές δημιουργήθηκαν στις αρχές της δεκαετίας του 1990 κυρίως στα πλαίσια του επιστημονικού κλάδου της Τεχνητής Νοημοσύνης. Κάποιες από αυτές είναι οι εξής : (1) Flogic, (2) LOOM, (3) OCML, (4) Ontolingua, (5) KIF και (6) Cari
Web-based γλώσσες	Με την αύξηση της χρήσης του Διαδικτύου όλο και περισσότερες γλώσσες οντολογίας αναπτύχθηκαν για να εκμεταλλευτούν την αφθονία στοιχείων του Παγκόσμιου Ιστού. Αυτές οι βασισμένες στον Ιστό γλώσσες οντολογίας (web-based ontology languages) ή οι γλώσσες σήμανσης οντολογίας(ontology markup languages) είναι συντακτικά βασισμένες στις υπάρχουσες γλώσσες σήμανσης όπως η HTML και η XML. Ακολουθεί μια σειρά από αυτές : (1) Simple HTML Ontology Extensions (SHOE), (2) XML-Based Ontology

	<p>exchanga Language (XOL) , (3) Ontology Markup Language (OML and KML), (4) Resource Description Framework Schema language (RDFS), (5) DARPA Agent Markup Language (DAML), (6) Ontology Interchange Language (OIL) και (7) Ontology Web Language (OWL).</p>
<p>Γλώσσες που αναπτύχθηκαν για να αναπαραστήσουν συγκεκριμένες οντολογίες και να χρησιμοποιηθούν σε συγκεκριμένες εφαρμογές</p>	<p>Παραδείγματα: CycL, GRAIL, NKRL Η διαφορά ανάμεσα στις παραδοσιακές και στις web-based γλώσσες είναι πως οι τελευταίες διαθέτουν καλά ορισμένη σύνταξη και σημασιολογία και ικανοποιητική συλλογιστική (reasoning) υποστήριξη. Επίσης, παρέχουν δύναμη και ευελιξία στην εκφραστικότητα και το συντακτικό τους είναι συμβατό με ήδη υπάρχοντα πρότυπα του Ιστού (XML, RDF, RDFS). Η διαφοροποίηση και ο διαχωρισμός των γλωσσών αναπαράστασης οντολογιών βασίζεται κυρίως (α) στη σύνταξη, (β) στην ορολογία, για παράδειγμα οι ιδιότητες σε άλλες γλώσσες αναφέρονται ως properties και σε άλλες ως slots (πχ Class-concept, Instance object, Slot-property), (γ) στην εκφραστικότητα, δηλαδή κάτι που μπορούμε να εκφράσουμε σε μία γλώσσα δεν μπορούμε σε μία άλλη και (δ) στη σημασιολογία καθώς η ίδια δήλωση μπορεί να σημαίνει διαφορετικά πράγματα σε διαφορετικές γλώσσες.</p>



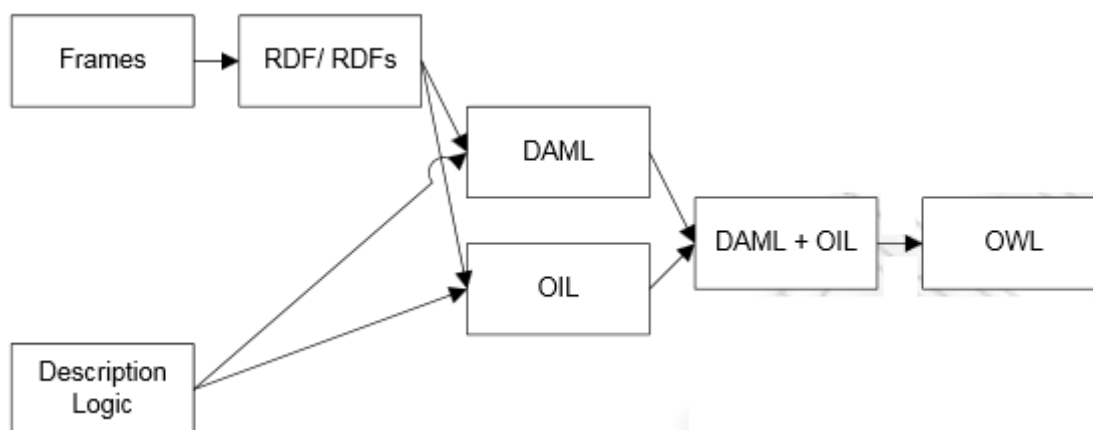
Εικόνα 6 Γλώσσες περιγραφής οντολογίας

3.2 Γλώσσα Οντολογίας Ιστού (OWL)

Η OWL είναι η πιο πρόσφατη εξέλιξη στις γλώσσες οντολογιών για το (Σημαντικό) Ιστό. Είναι σημαντικό να αναφερθεί ότι η OWL (που αναπτύσσεται από το W3C) δεν έχει περάσει πολύς καιρός από τότε που οριστικοποιήθηκε, καθώς μόλις τον Φεβρουάριο του 2004 πέρασε σε κατάσταση Σύστασης (W3C Recommendation).

Η OWL έχει δεχτεί επιρροές από τις περισσότερες γλώσσες που έχουν μέχρι τώρα αναφερθεί, αλλά ιδιαίτερα η DAML+OIL είναι αυτή που αποτέλεσε το εναρκτήριο σημείο για την ανάπτυξη της γλώσσας. Η σημαντική ομοιότητα με την DAML+OIL αποτελεί ουσιαστικό πλεονέκτημα, καθώς τα υπάρχοντα εργαλεία μπορούν, σχετικά εύκολα, να τροποποιηθούν, ώστε να υποστηρίζουν την OWL. (Πετράκης, 2011)

Η OWL, όπως και οι προκάτοχοί της, ακολουθεί το συντακτικό της RDF(S), εμπλουτίζοντας όμως το λεξιλόγιό της για την περιγραφή των κλάσεων και των ιδιοτήτων, υποστηρίζοντας, μεταξύ άλλων, σχέσεις μεταξύ των κλάσεων (π.χ. ότι δύο κλάσεις είναι ξένες), πληθικότητα (περιορισμοί αριθμού), ισότητα, περισσότερα χαρακτηριστικά των ιδιοτήτων (π.χ. συμμετρία) και κλάσεις απαρίθμησης (χρήση ονοματικών). Η μεγαλύτερη εκφραστικότητα που προσδίδεται στη γλώσσα με τον τρόπο αυτό, είναι αναγκαία για το Σημαντικό Ιστό, καθώς επιτρέπει τη χρήση συστημάτων συλλογισμού για την διεξαγωγή αποτελεσματικών συμπερασμών πάνω στις οντολογίες που αναπαρίσταται στον Ιστό με τη γλώσσα αυτή.



Εικόνα 7 Πρόγονοι της OWL (Πετράκης, 2011)

Μεγαλύτερη όμως εκφραστική δύναμη συνεπάγεται τη διεξαγωγή συμπερασμών αυξανόμενης πολυπλοκότητας έως και μη αποφασισιμότητας. Για το λόγο αυτό, η OWL ακολουθεί τη γνωστή προσέγγιση των διαδοχικών επιπέδων:

Η *OWL Lite* έχει σχεδιαστεί για την έκφραση ιεραρχιών ταξινόμησης και απλών περιορισμών ιδιοτήτων. Για παράδειγμα, ενώ η *OWL Lite* υποστηρίζει περιορισμούς πληθικότητας, οι μόνες τιμές που επιτρέπονται είναι 0 και 1. Είναι πιο εύκολο να σχεδιαστούν εργαλεία και να αντιστοιχιστούν θησαυροί όρων και ταξινομίες στην *OWL Lite* από ότι στα άλλα εκφραστικότερα επίπεδα. Η *OWL Lite* αντιστοιχεί στη Λογική Περιγραφής *SHIF(D)*. Η OWL [McGuinness & Harmelen, 2004] είναι η πιο πρόσφατη εξέλιξη στις γλώσσες οντολογιών για το (Σημαντικό) Ιστό. Είναι σημαντικό να αναφερθεί ότι η OWL (που αναπτύσσεται από το W3C) δεν έχει περάσει πολύς καιρός από τότε που οριστικοποιήθηκε, καθώς μόλις τον Φεβρουάριο του 2004 πέρασε σε κατάσταση Σύστασης (W3C Recommendation).

□ Η *OWL DL* υποστηρίζει τη μέγιστη δυνατή εκφραστικότητα, χωρίς απώλεια της αποφασισιμότητας. Η *OWL DL* ονομάζεται έτσι λόγω της αντιστοιχίας της με τις Λογικές Περιγραφής και συγκεκριμένα με τη Λογική Περιγραφής *SHOIN(D)*.

□ Η *OWL Full* προορίζεται για χρήστες που επιθυμούν μέγιστη εκφραστικότητα και την πλήρη εκφραστική ελευθερία του RDF χωρίς όμως εγγυήσεις επιλυσιμότητας. Για παράδειγμα η *OWL Full* επιτρέπει σε μια κλάση να ορίζεται ως συλλογή ατόμων

και συγχρόνως να θεωρείται άτομο ή ίδια (πράγμα που οδηγεί σε κυκλικούς ορισμούς εννοιών και επομένως σε μη αποφασισιμότητα). Θεωρείται απίθανο ότι κάποιο σύστημα συλλογισμού θα μπορέσει ποτέ να υποστηρίξει όλα τα χαρακτηριστικά της OWL Full.

Η OWL DL και η OWL Full χρησιμοποιούν το ίδιο λεξιλόγιο, αλλά η OWL DL υπάγεται σε κάποιους περιορισμούς, κυριότερος από τους οποίους είναι η απαίτηση για διαχωρισμό των τύπων: Μια κλάση δεν μπορεί να είναι συγχρόνως άτομο ή ιδιότητα και μια ιδιότητα δεν μπορεί να είναι συγχρόνως άτομο ή κλάση. Όσον αφορά τη σχέση με το RDF, η OWL Full μπορεί να θεωρηθεί επέκταση του RDF, ενώ οι OWL Lite και OWL DL μπορούν να θεωρηθούν επεκτάσεις μιας περιορισμένης όψης του RDF: Κάθε OWL (Lite, DL, Full) έγγραφο είναι ένα RDF έγγραφο και κάθε RDF έγγραφο είναι ένα OWL Full έγγραφο.

Η OWL έχει αυξημένη εκφραστικότητα και σε σχέση με την DAML+OIL. Η OWL DL μπορεί να θεωρηθεί ως η περισσότερο ισοδύναμη με την DAML+OIL, με τη διαφορά ότι έχουν αφαιρεθεί οι προσδιορισμένοι περιορισμοί αριθμού από τη γλώσσα, δεδομένου ότι δεν παρουσιάζοταν ουσιαστική ανάγκη για τη διατήρησή τους και επιπλέον ήταν δύσκολο να υλοποιηθούν. Άλλες αλλαγές περιλαμβάνουν τυπογραφικές αλλαγές και μετονομασίες, ενώ στην OWL έχουν ενσωματωθεί και οι πρόσφατες ενημερώσεις στον ορισμό της RDF(S).

3.3 Προγενέστερες Γλώσσες (XML, RDF)

Η XML είναι μία γλώσσα αναπαράστασης πληροφοριών στο Διαδίκτυο που επιτρέπει την αναπαράσταση δομημένων εγγράφων με χρήση λεξιλογίων που ορίζονται από τον χρήστη. Η γλώσσα αυτή δημιουργήθηκε για να καλύψει τις αδυναμίες της γλώσσας HTML, παρόλο που έχουν πολλές ομοιότητες. Ένα έγγραφο XML είναι δομημένο ιεραρχικά σε μορφή δέντρου, με στοιχεία που περικλείονται μέσα σε άλλα στοιχεία και με ένα στοιχείο ανώτερου επιπέδου γνωστό ως βασικό στοιχείο, το οποίο περιέχει όλα τα υπόλοιπα. Κάθε στοιχείο αποτελείται από μία ετικέτα αρχής, το περιεχόμενο και μία ετικέτα τέλους.

Το XML Schema είναι η γλώσσα που καθορίζει τη δομή των XML εγγράφων. Ορίζει τα επιτρεπόμενα στοιχεία, τις ιδιότητές τους, καθώς και τον τρόπο με τον οποίο

συνδυάζονται μεταξύ τους μέσα στο XML κείμενο. Με απλά λόγια, το XML Schema αποτελεί το «συντακτικό» του XML κειμένου.

HTML vs. XML

Παράδειγμα XML

HTML: παρουσίαση	<pre><Order> <to> Tom </to> <oderDate> 22 Nov 2003</orderDate> <orderItem> <product> milk </product> <quantity> 500 ml </quantity> </orderItem> <orderItem> <product> biscuits </product> <quantity> 1 box </quantity> </orderItem> <orderItem> <product> coffee </product> <quantity> 250 grams </quantity> </orderItem> </Order></pre>
XML: δομή	
HTML: χρήση από ανθρώπους	
XML: χρήση από μηχανές	
HTML: καθορισμένα tags	
XML: επιπρόσθετα tags	

Παράδειγμα HTML

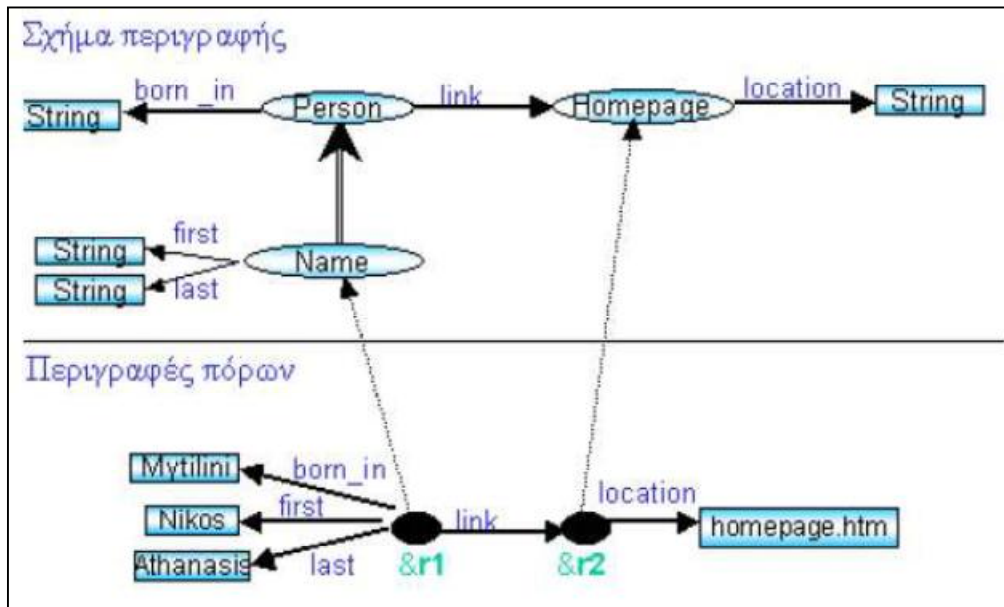
```
<h1> My Order </h1>
<h2> Tom </h2>
<h3> 22 Nov 2003 </h3>
<table> <tr><td> 500 ml <td> milk
        <tr><td> 1 box <td> biscuits
        <tr><td> 250 grams <td> coffee
</table>
```

Εικόνα 8 HTML vs. XML

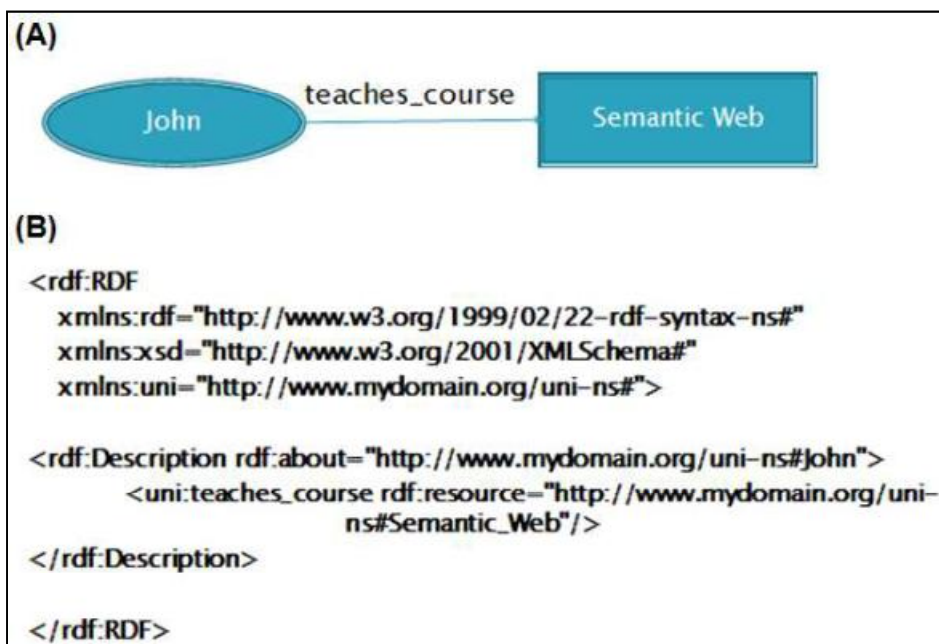
Το πλαίσιο περιγραφής διαδικτυακών πόρων RDF (Resource Description Framework) είναι ένα μοντέλο δεδομένων για την αναπαράσταση της πληροφορίας που σχετίζεται με τους πόρους στον Παγκόσμιο Ιστό, δηλαδή για τα μεταδεδομένα. Τα μεταδεδομένα αφορούν τις συμπληρωματικές πληροφορίες όπως είναι ο τίτλος, ο συντάκτης και η ημερομηνία τελευταίας τροποποίησης μιας ιστοσελίδας, η άδεια χρήσης ενός εγγράφου στο Διαδίκτυο ή ακόμα και το χρονοδιάγραμμα διαθεσιμότητας ενός κοινόχρηστου διαδικτυακού πόρου. Το RDF προορίζεται για περιπτώσεις που τις πληροφορίες τις επεξεργάζονται εφαρμογές και όχι για παρουσίαση πληροφοριών σε χρήστες. Μέσω αυτού του μοντέλου παρέχεται ένα κοινό πλαίσιο εργασίας ώστε οι πληροφορίες μεταξύ των εφαρμογών να ανταλλάσσονται χωρίς να χάνεται το νόημα. Επιπλέον, οι πληροφορίες πλέον θα είναι διαθέσιμες και σε εφαρμογές διαφορετικές από αυτές για τις οποίες είχαν αρχικά σχεδιαστεί. Η σύνταξη των διαδικτυακών πόρων υλοποιείται με τη βοήθεια αναπαραστάσεων που έχουν τη μορφή πόρος - ιδιότητα - τιμή. Οι διαδικτυακοί πόροι αναπαριστώνται με τη χρήση αναφορών URI και τους αποδίδονται απλές ιδιότητες και οι αντίστοιχες τιμές τους, με τη βοήθεια της ορολογίας που παρέχεται από τη γλώσσα. Ο πόρος για τον οποίο γίνεται λόγος λέγεται υποκείμενο (subject), ο πόρος που προσδιορίζει μια ιδιότητα ένα χαρακτηριστικό του υποκειμένου ονομάζεται κατηγορημα (predicate), ενώ το στοιχείο που δίνει τιμή σε αυτήν την ιδιότητα καλείται αντικείμενο (object). Ο συνδυασμός των τριών καλείται δήλωση ή ισχυρισμός ή πρόταση (statement). Πέρα από τις RDF περιγραφές των πληροφοριακών πόρων, το μοντέλο δεδομένων RDF αποτελείται και από μια γλώσσα ορισμού σχημάτων (συνόλων κλάσεων και ιδιοτήτων), την RDF Schema Specification Language (RDFS). Με τη βοήθεια του RDFS, είναι δυνατός ο προσδιορισμός μηχανισμών καθορισμού κλάσεων πόρων, καθώς και ο περιορισμός των πιθανών συνδυασμών κλάσεων μεταξύ τους χρησιμοποιώντας κατάλληλες συσχετίσεις. Ένα RDF σχήμα αποτελείται από τις δηλώσεις κλάσεων, γνωρισμάτων και των σχέσεων μεταξύ των κλάσεων. Όμοιοι πόροι είναι ομαδοποιημένοι κάτω από την ίδια κλάση. Με βάση τα παραπάνω, μπορούμε να διακρίνουμε τρία διαφορετικά επίπεδα αφαίρεσης στο μοντέλο δεδομένων RDFS: Στο κατώτερο επίπεδο υπάρχουν οι ίδιοι οι πόροι (έγγραφα, δικτυακοί τόποι, πρόσωπα ή οτιδήποτε άλλο). Το επόμενο επίπεδο αφαίρεσης είναι το επίπεδο δεδομένων, όπου γίνεται η περιγραφή των

πληροφοριακών πόρων με τη χρήση λεξιλογίων, τα οποία περιγράφονται στο τελευταίο επίπεδο, το επίπεδο σχήματος. Το επίπεδο σχήματος είναι το επίπεδο αφαίρεσης όπου αναπτύσσονται RDF σχήματα για να διευκολύνουν τη σημασιολογική περιγραφή των πόρων. Σε αυτό το επίπεδο, οι κλάσεις αναπαριστούν αφηρημένες οντότητες και αναφέρονται συλλογικά σε σύνολα παρόμοιων αντικειμένων. Για την γραφική απεικόνιση των RDF/S περιγραφών και σχημάτων χρησιμοποιείται ένα μοντέλο κατευθυνόμενων γράφων με ετικέτες τόσο στις ακμές όσο και στους κόμβους που μπορεί εύκολα να συνδυάσει πολλά διαφορετικά λεξιλόγια και να επεκταθεί προσθέτοντας απλώς περισσότερες ακμές. Οι κόμβοι αναπαριστούν αντικείμενα (πόρους ή κλάσεις) και οι ακμές συμβολίζουν σχέσεις μεταξύ των κόμβων (ιδιότητες). Ως κόμβους αναπαριστούμε επίσης και τύπους Literal, δηλαδή αλφαριθμητικά και άλλους βασικούς τύπους δεδομένων, όπως αυτοί ορίζονται στο πρότυπο XML Schema. Οι κόμβοι συμβολίζονται ως ελλείψεις και οι ατομικές τιμές ως παραλληλόγραμμα. Οι ακμές μπορεί να είναι τριών ειδών: απόδοσης γνωρισμάτων (attributes), δημιουργίας στιγμιότυπων (instances) και υπαλληλίας. Οι ακμές απόδοσης γνωρισμάτων αναπαριστούν γνωρίσματα κόμβων και σχέσεις μεταξύ τους, ενώ οι ακμές υπαλληλίας χρησιμοποιούνται για να δηλώσουν σε επίπεδο σχήματος ότι ένας κόμβος (κλάση) ή ιδιότητα είναι υποκατηγορία ενός σημασιολογικά ευρύτερου κόμβου ή ιδιότητας αντίστοιχα. Τέλος, οι ακμές δημιουργίας στιγμιότυπων αποτελούν τη σύνδεση ανάμεσα στα πρότυπα RDF και RDFS, επιτρέποντας τη δημιουργία στιγμιότυπων μίας κλάσης και την απόδοση τύπων σε πληροφοριακούς πόρους που περιγράφονται. Οι γράφοι του RDF σχήματος περιγραφής και των RDF περιγραφών πόρων για την πληροφορία και αποδεικνύεται στην ακόλουθη εικόνα :

Για τον καλύτερο διαμοιρασμό των RDF περιγραφών πόρων και των σχημάτων περιγραφών στο διαδίκτυο, το μοντέλο RDF «δανείζεται» το συντακτικό της XML. Έτσι προκύπτει η RDF/XML σύνταξη των RDF σχημάτων περιγραφής και RDF περιγραφής πόρων που απεικονίζονται στους αντίστοιχους γράφους αναπαράστασης.



Εικόνα 9 Γράφοι αναπαράστασης πληροφορίας σύμφωνα με το μοντέλο RDF



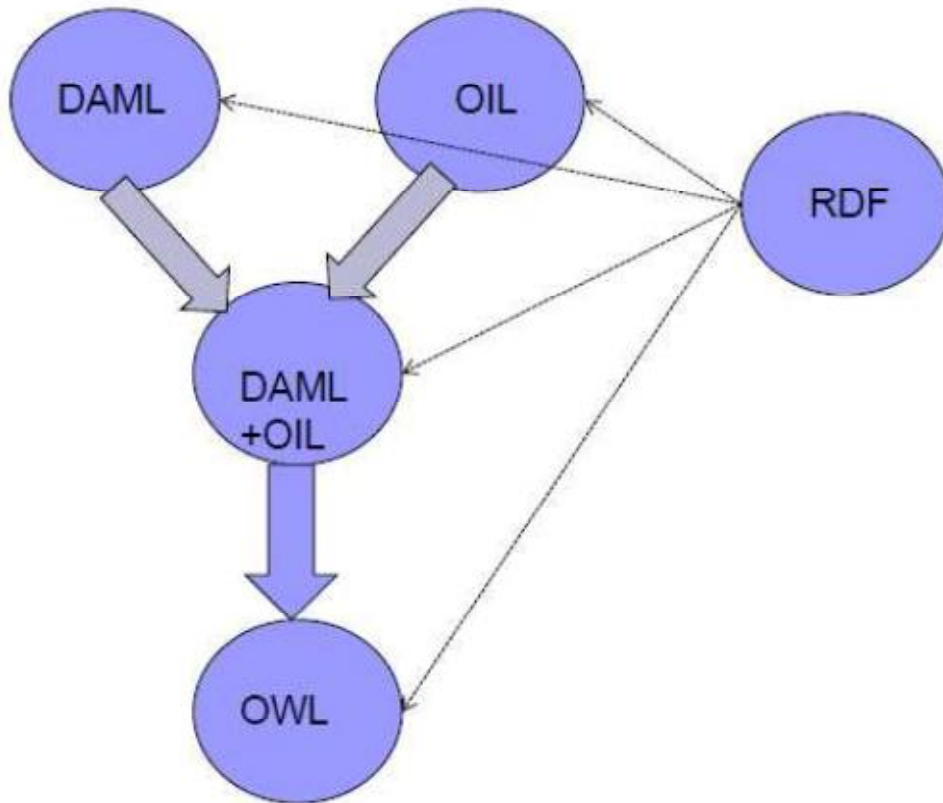
Εικόνα 10 (Α) Γράφος αναπαράστασης πληροφορίας - (Β) RDF/XML σύνταξη

3.4 Χαρακτηριστικά της Γλώσσας Οντολογίας OWL

Η OWL δημιουργήθηκε για να ικανοποιήσει την ανάγκη για μια γλώσσα οντολογίας Ιστού κι αποτελεί μια από τις προτάσεις του W3C σχετικών με το Σημασιολογικό Ιστό. 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 Διαδικτύου και του Σημασιολογικού Ιστού. Η τυποποίηση των οντολογιών σε γλώσσα OWL θα κάνει τα δεδομένα στο Web περισσότερο επεξεργάσιμα από μηχανές και επαναχρησιμοποιήσιμα στις εφαρμογές. Έτσι, η επεκτασιμότητα, η μετατρεψιμότητα και η διαλειτουργικότητα κατέχουν υψηλή προτεραιότητα στο σχεδιασμό της γλώσσας. Συγκριτικά με τις υπάρχουσες τεχνολογίες, η OWL υπερβαίνει τις γνωστές και ευρέως διαδεδομένες 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. Η OWL προσθέτει περισσότερο λεξιλόγιο για την περιγραφή των ιδιοτήτων και των κλάσεων όπως για παράδειγμα σχέσεις μεταξύ των κλάσεων (disjointness), αριθμός στοιχείων συνόλου (cardinality), λογικοί συνδυασμοί κλάσεων (ένωση, τομή, συμπλήρωμα), ισότητα, πλουσιότερη δακτυλογράφηση ιδιοτήτων, χαρακτηριστικά ιδιοτήτων (για παράδειγμα συμμετρία, μοναδικότητα, μεταβατικότητα) και απαριθμημένες κλάσεις. Ένα πλεονέκτημα των OWLοντολογιών είναι η διαθεσιμότητα των εργαλείων που μπορούν να εξάγουν συμπεράσματα για αυτές. Τα εργαλεία παράγουν γενική υποστήριξη που δεν είναι συγκεκριμένη για την ιδιαίτερη υπαγόμενη περιοχή, κάτι το οποίο θα συνέβαινε εάν επρόκειτο να δημιουργηθεί ένα σύστημα για την εξαγωγή συλλογισμών για ένα συγκεκριμένο βιομηχανικό πρότυπο XML σχήματος. Η γλώσσα OWL αναπτύσσεται ως επέκταση του λεξιλογίου RDF, προέρχεται από τη γλώσσα οντολογίας Ιστού DAML+OIL και είναι γραμμένη σε μορφή XML ώστε να μπορεί εύκολα να είναι ανεξάρτητη του λειτουργικού συστήματος και των γλωσσών εφαρμογής ενός υπολογιστή. Επιπλέον, ο Σημασιολογικός Ιστός χρησιμοποιώντας βασισμένες στην XML οντολογίες κι όχι παραδοσιακές δομές δέντρων θα καταστήσει ευκολότερη για τις εφαρμογές Υπηρεσιών Ιστού την επεξεργασία στοιχείων από πολλαπλές πηγές. Μια OWL οντολογία περιλαμβάνει τις περιγραφές των κλάσεων, των ιδιοτήτων και των στιγμιότυπων τους. Λαμβάνοντας υπόψη μια τέτοια οντολογία, η επίσημη OWL σημασιολογία διευκρινίζει πώς να παράγει τα λογικά συμπεράσματά 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 καταναμημένα έγγραφα που έχουν συνδυαστεί χρησιμοποιώντας καθορισμένους OWL μηχανισμούς.



Εικόνα 11 Η δημιουργία της γλώσσας OWL

Δεδομένου ότι ο Σημασιολογικός Ιστός κατανέμεται, η OWL πρέπει να επιτρέψει να συγκεντρώνονται πληροφορίες από καταναμημένες πηγές. Αυτό γίνεται εν μέρει επιτρέποντας στις οντολογίες να είναι συσχετισμένες, συμπεριλαμβανομένης ρητά της εισαγωγής πληροφοριών από άλλες οντολογίες. Οι πρακτικές εφαρμογές της OWL, περιλαμβάνουν τα Web portals, όπου μπορεί να χρησιμοποιηθεί για να δημιουργήσει τους κανόνες κατηγοριοποίησης, προκειμένου να βελτιωθεί η αναζήτηση, οι Συλλογές Πολυμέσων, όπου μπορεί να χρησιμοποιηθεί για να επιτρέψει βασισμένες σε περιεχόμενο αναζητήσεις των μέσων και οι Υπηρεσίες Ιστού (Web Services), όπου μπορεί να χρησιμοποιηθεί για την ανακάλυψη και τη σύνθεση Υπηρεσιών Ιστού (composition) καθώς επίσης και για τη διαχείριση δικαιωμάτων και τον έλεγχο πρόσβασης.

3.5 Υπο-Γλώσσες OWL

Η OWL μπορεί να ταξινομηθεί σε τρεις επιμέρους υπο-γλώσσες αυξανόμενης εκφραστικότητας. Η κάθε μια στοχεύει στην ικανοποίηση διαφορετικών πλευρών του πλήρους συνόλου των απαιτήσεων. Στον ακόλουθο πίνακα παρουσιάζονται οι υπο-γλώσσες OWL.

Πίνακας 5 Υπο-γλώσσες OWL

OWL-Lite	Είναι η πιο απλή συντακτικά και η λιγότερο εκφραστική υπο-γλώσσα. Χρησιμοποιείται για οντολογίες που αποτελούνται από μια απλή ιεραρχία κλάσεων και απλούς περιορισμούς. Υποστηρίζει cardinality constraints, αλλά επιτρέπει μόνο τις τιμές 0 ή 1. Καθιστά απλούστερη την κατασκευή ενός εργαλείου που θα στηρίζεται στην OWL Lite σε σύγκριση με τις άλλες πιο εκφραστικές υπο-γλώσσες. Παρέχει τη δυνατότητα μιας γρήγορης μετάβασης από θησαυρούς λέξεων και άλλες ταξινομήσεις.
OWL-DL	Είναι πιο εκφραστική από την OWL-Lite και βασίζεται στα Description Logics. Μια οντολογία εκφρασμένη σε OWL-DL μπορεί να χρησιμοποιηθεί από κάποιο μηχανισμό εξαγωγής συμπερασμάτων (reasoner) καθώς και να ελεγχθεί αυτόματα αν περιέχει ασυνέπειες. Με την OWL-DL διατηρείται η υπολογιστική πληρότητα (computational completeness), δηλαδή όλοι οι υπολογισμοί είναι πραγματοποιήσιμοι και η λήψη απόφασης (decidability), δηλαδή όλοι οι υπολογισμοί ολοκληρώνονται σε πεπερασμένο χρόνο. Περιλαμβάνει όλες τις γλωσσικές κατασκευές της OWL, αλλά με συγκεκριμένους περιορισμούς.
OWL-Full	Είναι η πιο εκφραστική υπο-γλώσσα. Χρησιμοποιείται σε περιπτώσεις που απαιτείται μεγάλη εκφραστικότητα, ενώ δεν είναι δυνατόν να γίνει αυτόματος συμπερασμός σε μια OWL-Full οντολογία. Παρέχει τη συντακτική ελευθερία του RDF χωρίς όμως εγγυήσεις για τους υπολογισμούς. Για παράδειγμα, στην OWL-Full μια κλάση μπορεί να

	αντιμετωπίζεται παράλληλα σαν μια συλλογή από στοιχεία αλλά και ως ένα στοιχείο η ίδια. Η OWL Full επιτρέπει σε μια οντολογία να επεκτείνει τη σημασία ενός προ-ορισμένου (RDF ή OWL) λεξικού.
--	--

Υπάρχουν συγκεκριμένες έννοιες για την ανωφερή συμβατότητα μεταξύ των τριών υπο-γλωσσών. Ειδικότερα:

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

Η OWL Full, ούσα πιο κοντά στη θεωρία μοντέλου του RDF και μη έχοντας το χαρακτηριστικό της αποφασισιμότητας, δεν αντιστοιχεί ούτε συντακτικά ούτε σημασιακά σε κάποια Λογική Περιγραφής. Ωστόσο, ακόμα και η OWL DL, που ισοδυναμεί εκφραστικά με την Λογική Περιγραφής *SHOIN(D)*.

έχει ορισμένες διαφορές από τις Λογικές Περιγραφής, κυρίως ως προς το συντακτικό, τις εκφράσεις σχολιασμού (annotations) και τις δυνατότητες εισαγωγής (import). Οι διαφορές αυτές παρέχουν τη γέφυρα μεταξύ των τυπικών Λογικών Περιγραφής και του Σημαντικού Ιστού: (KAON, 2008).

Η OWL χρησιμοποιεί αναφορές URI ως ονόματα και κατασκευάζει τις αναφορές αυτές με τον τρόπο που χρησιμοποιείται στο RDF. Είναι επομένως σύνηθες στην OWL να χρησιμοποιούνται προσδιορισμένα ονόματα σαν συντομογραφίες για τις αναφορές URI, όπως, για παράδειγμα, το προσδιορισμένο όνομα owl:Thing για την αναφορά URI <http://www.w3.org/2002/07/owl#Thing>.

Η OWL συγκεντρώνει πληροφορίες σε οντολογίες, που είναι γενικά αποθηκευμένες ως έγγραφα Ιστού, γραμμένα σε RDF/XML. Οι οντολογίες μπορούν να εισάγουν

άλλες οντολογίες, προσθέτοντας τις πληροφορίες από την εισαγόμενη οντολογία στην τρέχουσα.

Η OWL επιτρέπει τη χρήση ιδιοτήτων σχολιασμού RDF για την επισύναψη πληροφοριών σε κλάσεις, ιδιότητες και οντολογίες, όπως π.χ. owl:DeprecatedClass. Οι σχολιασμοί αυτοί είναι τριπλέτες RDF και απαιτείται επομένως να φέρουν πλήρες σημασιακό βάρος στο RDF. Στην OWL DL όμως, οι σχολιασμοί αυτοί φέρουν μια ξεχωριστή, περιορισμένη σημασία.

Η OWL χρησιμοποιεί τις υπηρεσίες τύπων δεδομένων του RDF και του XML Schema για την παροχή τύπων και τιμών δεδομένων (μια πολύ περιορισμένη μορφή απτών πεδίων).

Η OWL DL και η OWL Lite έχουν αφηρημένη σύνταξη τύπου πλαισίων, ενώ η RDF/XML είναι το επίσημο συντακτικό ανταλλαγής και για τα τρία είδη της OWL.

3.6 Σύγκριση της OWL με Προγενέστερες Γλώσσες (XML, RDF και RDFS)

Η XML παρέχει ένα συντακτικό για τη δημιουργία δομημένων εγγράφων, το οποίο όμως δεν εισάγει σημασιολογικούς περιορισμούς στη σημασία αυτών των εγγράφων. Το XML Schema είναι μια γλώσσα για τον προσδιορισμό/περιορισμό της δομής των XML εγγράφων και επιπλέον επεκτείνει την XML με τη χρήση των τύπων δεδομένων. Συνεπώς, ένα σύνολο από XML δηλώσεις δε μας οδηγεί στην εξαγωγή συμπερασμάτων για οποιοδήποτε άλλο XML σύνολο δηλώσεων, πράγμα το οποίο συμβαίνει στην OWL με OWL δηλώσεις αντίστοιχα. Μία οντολογία αποτελεί αναπαράσταση γνώσης, παρά μορφή μηνύματος. Τα XML μηνύματα επικεντρώνονται σε συγκεκριμένα δεδομένα και δεν επιτρέπουν την εξαγωγή πληροφοριών πέρα από το πλαίσιο στο οποίο χρησιμοποιούνται. Άλλες γλώσσες ενσωμάτωσαν τη σημασιολογία των λέξεων που περιγράφουν, αλλά η εκφραστικότητα που προσφέρουν είναι πολύ περιορισμένη. Συγκεκριμένα, το RDF είναι ένα μοντέλο δεδομένων για πόρους και σχέσεις μεταξύ των πόρων που παρέχει μια απλή σημασιολογία και μπορεί να αναπαρασταθεί με χρήση XML σύνταξης, RDF/XML. Το RDF Schema είναι ένα λεξικό για την περιγραφή ιδιοτήτων και

κλάσεων RDF πόρων, που εισάγει κάποια σημασιολογία σχετικά με την ιεραρχία τέτοιων ιδιοτήτων και κλάσεων. Το RDF σχήμα και η OWL είναι συμβατές, πράγμα το οποίο σημαίνει ότι μπορούν να υπάρχουν RDFS στοιχεία μέσα σε ένα σύνολο OWL δηλώσεων. Ωστόσο, η OWL παρέχει διάφορες έννοιες που είναι μοναδικές σε αυτή, όπως π.χ. περιορισμοί ιδιοτήτων, Boolean συνδυασμοί εκφράσεων των κλάσεων κ.ά. Η OWL σχεδιάστηκε με τον ίδιο σκοπό που σχεδιάστηκε και η RDFS. Συγκεκριμένα, να παρέχει ένα XML λεξιλόγιο για τον ορισμό των κλάσεων και των μεταξύ τους σχέσεων και των ιδιοτήτων. Τα RDF σχήματα μας επιτρέπουν να εκφράσουμε πολύ στοιχειώδεις σχέσεις και παρουσιάζουν μειωμένη ικανότητα συλλογισμού, σε αντίθεση με την OWL που μπορεί να εκφράσει πιο σύνθετες σχέσεις και προσφέρει τη δυνατότητα συλλογισμού. Η OWL εξακολουθεί να χρησιμοποιεί την RDF και το RDF Schema σε μεγάλο βαθμό. Όλα τα είδη της OWL χρησιμοποιούν την RDF για τη σύνταξή τους και τα στιγμιότυπα ορίζονται όπως στην RDF, χρησιμοποιώντας τις RDF περιγραφές και εισάγοντας την πληροφορία. Η OWL προσθέτει περισσότερο λεξιλόγιο για την περιγραφή ιδιοτήτων και κλάσεων. Συγκεκριμένα με την OWL μπορούν να εκφραστούν περιορισμοί ως προς το πλήθος των στοιχείων ενός συνόλου (cardinality constraints), περιορισμοί στο εύρος ή το πλήθος τιμών μιας ιδιότητας σε σχέση με την κλάση στην οποία ανήκει ο πόρος στον οποίο αναφέρεται η ιδιότητα, μεταβατικές ιδιότητες, νέες κλάσεις με συνδυασμό (ενώσεις ή τομές) άλλων κλάσεων ή πλήρως διαχωρισμένες κλάσεις (δηλαδή κανένα στιγμιότυπο της μιας κλάσης δεν μπορεί να ανήκει και στην άλλη). Επιπλέον, παρέχει τη δυνατότητα καθορισμού πως μια συγκεκριμένη ιδιότητα αποτελεί μοναδικό αναγνωριστικό (κλειδί) για αντικείμενα μιας συγκεκριμένης κλάσης, πως δύο διαφορετικές κλάσεις ή στιγμιότυπα στην πραγματικότητα αναπαριστούν την ίδια κλάση ή το ίδιο αντικείμενο.

3.7 Έγγραφο OWL

Τα βασικά στοιχεία που ορίζουν μία οντολογία σε ένα OWL έγγραφο είναι οι κλάσεις και οι ιδιότητές τους, οι σχέσεις μεταξύ των κλάσεων και περιορισμοί στις σχέσεις και στις ιδιότητες των κλάσεων. Έτσι, σε ένα OWL έγγραφο καθορίζονται οι κλάσεις,

η ιεραρχία των κλάσεων, ισοδύναμες κλάσεις και ιδιότητες, συσχετίσεις των κλάσεων και μεταδεδομένα των ιδιοτήτων. Τα OWL έγγραφα μοιάζουν αρκετά με τα XML και τα RDF έγγραφα, με την έννοια ότι χρησιμοποιούν και αυτά στοιχεία, ετικέτες και ονόματα χώρου (namespaces). Ένα OWL έγγραφο ξεκινά με μια κεφαλίδα (header) και στη συνέχεια ορίζει τις κλάσεις και τις ιδιότητες.

3.6.1 Κεφαλίδα (Header)

Ένα OWL έγγραφο δεν είναι υποχρεωτικό να περιέχει κεφαλίδα, ωστόσο η συμπερίληψή της αποτελεί καλή πρακτική. Στην κεφαλίδα (header) ορίζεται το εξωτερικό Ontology block `<owl:Ontology>`, το οποίο μπορεί να περιλαμβάνει ορισμούς για την οντολογία, σχόλια (comments), ετικέτες (labels), πληροφορίες για την έκδοση και τη συμβατότητά της με προηγούμενες εκδόσεις, καθώς επίσης και import δηλώσεις. Οι import δηλώσεις προσδιορίζουν έναν άλλο RDF πόρο που παρέχει ορισμούς για την οντολογία. Όλοι οι OWL ορισμοί βρίσκονται εμφωλευμένοι σε ένα *rdf:RDF* στοιχείο.

3.6.2 Κλάσεις και στιγμιότυπα

Οι κλάσεις αναπαριστούν σύνολο πόρων που έχουν κοινά χαρακτηριστικά. Ένας πόρος που είναι μέλος μιας κλάσης καλείται Individual και είναι στιγμιότυπο της κλάσης. Τα Individuals μπορούν να ανήκουν σε παραπάνω από μία κλάσεις. Κλάσεις με ίδια Individuals δε σήμαινε ότι είναι ίδιες. Η συμμετοχή σε μία κλάση μπορεί να γίνει ορίζοντάς την ρητά αλλά και έμμεσα αναφέροντας απλά τις απαραίτητες συνθήκες συμμετοχής σε μία κλάση. Στην OWL η πρωταρχική κλάση είναι η *owl:Thing*, το URI της οποίας είναι <http://www.w3.org/2002/07/owl#Thing>. Αυτός ο πόρος αναπαριστά την κλάση όλων των Individuals. Ο πόρος *owl:Class* περιέχει όλες τις κλάσεις. Κάθε κλάση πρέπει να είναι μέλος της *owl:Class* και κάθε πόρος του οποίου η τιμή της *rdf:type* είναι *owl:Class* είναι κλάση. Στην επόμενη εικόνα παρουσιάζεται η δομή των κλάσεων όπως ορίζεται στην OWL.

Μία κλάση περιγράφεται με σχέσεις υποκλάσης, τα μέλη της, περιορισμούς ιδιοτήτων και λογικές πράξεις με άλλες κλάσεις (π.χ. ένωση). Συγκεκριμένα, η σχέση υποκλάσης ορίζεται με την ιδιότητα *rdfs:subclassOf* και χρησιμοποιείται για να ορίσει μία ιεραρχική σχέση μεταξύ κλάσεων. Τα μέλη μιας κλάσης μπορούν να περιγραφούν με πράξεις συνόλων, με απαρίθμηση και με σχέσεις ισοδυναμίας ή διαφορετικότητας. Αναλυτικότερα, οι πράξεις συνόλων που μπορούν να εφαρμοστούν στις κλάσεις είναι *owl:intersectionOf*, *owl:unionOf* και *owl:complementOf*, οι οποίες αντιστοιχούν στην τομή, την ένωση και το συμπλήρωμα όπως ακριβώς ορίζονται στη θεωρία συνόλων. Στην απαρίθμηση τα μέλη μιας κλάσης ορίζονται ρητά, δηλαδή αναγράφοντάς τα σε λίστα ένα-ένα. Τούτο συνεπάγεται ότι αν ένα στιγμιότυπο δεν αναφέρεται στη λίστα απαρίθμησης, τότε δε μπορεί να γίνει μέλος της κλάσης. Τα ισοδύναμα στιγμιότυπα δηλώνονται με την ιδιότητα *owl:sameAs* και οι ισοδύναμες κλάσεις με την ιδιότητα *owl:equivalentClass*. Τέλος, οι ξένες μεταξύ τους κλάσεις (disjoint classes) ορίζονται με την ιδιότητα *owl:disjointWith*, όπου κανένα στιγμιότυπο οποιασδήποτε κλάσης δε μπορεί να είναι μέλος της άλλης, και τα διαφορετικά στιγμιότυπα χρησιμοποιώντας την ιδιότητα *owl:differentFrom*.

3.6.3 Ιδιότητες

Οι ιδιότητες στην OWL είναι παρόμοιες με αυτές του RDF σχήματος. Συγκεκριμένα, για τον ορισμό ιδιοτήτων σε OWL έγγραφα χρησιμοποιούνται τα *rdfs:domain* και *rdfs:range* και προστίθενται επιπλέον περιορισμοί.

Στην OWL υπάρχουν δύο είδη ιδιοτήτων.

- Η *ObjectProperty* η οποία συσχετίζει αντικείμενα με αντικείμενα, δηλαδή το πεδίο ορισμού και το πεδίο τιμών αποτελείται από κλάσεις.
- Η *DatatypeProperty* η οποία συσχετίζει αντικείμενα με τύπους δεδομένων, δηλαδή το πεδίο τιμών είναι ένας τύπος δεδομένων. Η OWL, εκτός από τους τύπους δεδομένων που ορίζονται στο XML Schema, παρέχει τη δυνατότητα δημιουργίας κλάσεων τύπων δεδομένων που μπορούν να χρησιμοποιηθούν για να περιορίσουν το

πεδίο τιμών των ιδιοτήτων.

Στην OWL, οι ιδιότητες μπορούν να είναι μεταβατικές, συμμετρικές, συναρτησιακές ή αντίστροφες. Μία σχέση-ιδιότητα s είναι μεταβατική (transitive) αν η s συνδέει το στιγμιότυπο a με το στιγμιότυπο b και το στιγμιότυπο b με το στιγμιότυπο c , τότε και το στιγμιότυπο a συνδέεται με το στιγμιότυπο c μέσω της σχέσης s . Για μία μεταβατική ιδιότητα P ισχύει η εξής λογική σχέση: $P(x,y)$ και $P(y,z)$ τότε $P(x,z)$ για οποιαδήποτε x,y,z . Μεταβατική είναι η ιδιότητα της υποκλάσης. Αν η κλάση A είναι υποκλάση της B και η κλάση B είναι υποκλάση της Γ , τότε η A είναι υποκλάση της Γ .

4. ΔΙΑΧΕΙΡΙΣΗ ΟΝΤΟΛΟΓΙΩΝ

4.1 Εργαλεία Ανάπτυξης & Διαχείρισης Οντολογιών

Υπάρχουν αρκετά εργαλεία τα οποία χρησιμεύουν στην ανάπτυξη και διαχείριση οντολογιών. Έχουν αναπτυχθεί καθαρές γλώσσες ορισμού οντολογιών σε γραμμή εντολών, αλλά υπάρχουν και πιο σύγχρονα εργαλεία (κυρίως Java based) τα οποία διαθέτουν παραθυρικό περιβάλλον και συνεπώς υπερτερούν των πρώτων λόγω της ευκολίας και της εποπτείας με εικόνα της διαχείρισης. Μερικά εργαλεία γραμμής εντολών είναι τα εξής : Methontology, On-To-Knowledge methodology, AFM (Activity-First Method) κλπ. Τα σημαντικότερα εργαλεία ανάπτυξης οντολογιών με γραφικό περιβάλλον είναι τα εξής :

- OntoEdit,
- Ontolingua Server,
- WebOnto,
- WebODE,
- OIEd Protege

Αδιαμφισβήτητα το Protege κατέχει την πρώτη θέση ως προς την αποτελεσματικότητα, την εξελισσιμότητα αλλά και την φιλικότητα του γραφικού του περιβάλλοντος προς το χρήστη.

Το Protege αποτελεί μια πλατφόρμα «ανοιχτού κώδικα» (open source) βασισμένη σε JAVA που παρέχει ένα σύνολο από εργαλεία για την μοντελοποίηση πεδίων γνώσης και την ανάπτυξη και επεξεργασία οντολογιών. Πρωτοεμφανίστηκε το 1988 και στην αρχή αποτελούσε απλώς ένα μέσο για τη δημιουργία εργαλείων ανάκτησης γνώσης για έμπειρα συστήματα. Σήμερα, χάρη στις προσπάθειες του τμήματος ιατρικής πληροφορικής του Πανεπιστημίου του Stanford (Stanford Medical Informatics – SMI) το Protege έχει εξελιχθεί σε ένα σύγχρονο εργαλείο μοντελοποίησης γνώσης, το οποίο εξελίσσεται με ταχύτατους ρυθμούς. Το βασικό τμήμα του Protege παρέχει ένα πλούσιο σύνολο από δομές μοντελοποίησης γνώσης και λειτουργίες που υποστηρίζουν τη δημιουργία, απεικόνιση και επεξεργασία οντολογιών σε μια πληθώρα τύπων αναπαράστασης. Επιπλέον, το Protege μπορεί να επεκταθεί κάνοντας χρήση της αρχιτεκτονικής των plugins και μιας προγραμματιστικής διεπαφής (Application Programming Interface – API) βασισμένη σε JAVA, την οποία υποστηρίζει για την ανάπτυξη εργαλείων και εφαρμογών επεξεργασίας οντολογιών. Μια οντολογία περιγράφει το σύνολο των ρητών προσδιοριστικών εννοιών και συσχετίσεων που είναι σημαντικές σε ένα συγκεκριμένο πεδίο γνώσης, παρέχοντας ένα λεξικό για το πεδίο αυτό. Επίσης, προσδιορίζει τη σημασία των όρων που χρησιμοποιούνται στο συγκεκριμένο πεδίο και βρίσκονται στο λεξικό. Τα τελευταία χρόνια, οι οντολογίες συναντώνται σε αρκετές επιστημονικές και επαγγελματικές κοινότητες ως μέσο διαμερισμού, επαναχρησιμοποίησης και επεξεργασίας γνώσης σε συγκεκριμένα πεδία. Η πλατφόρμα του Protege υποστηρίζει δυο βασικούς τρόπους μοντελοποίηση οντολογιών:

- Η εφαρμογή «**Protege – Frames Editor**» που δίνει τη δυνατότητα στους χρήστες να κατασκευάσουν και να δημοσιοποιήσουν οντολογίες που

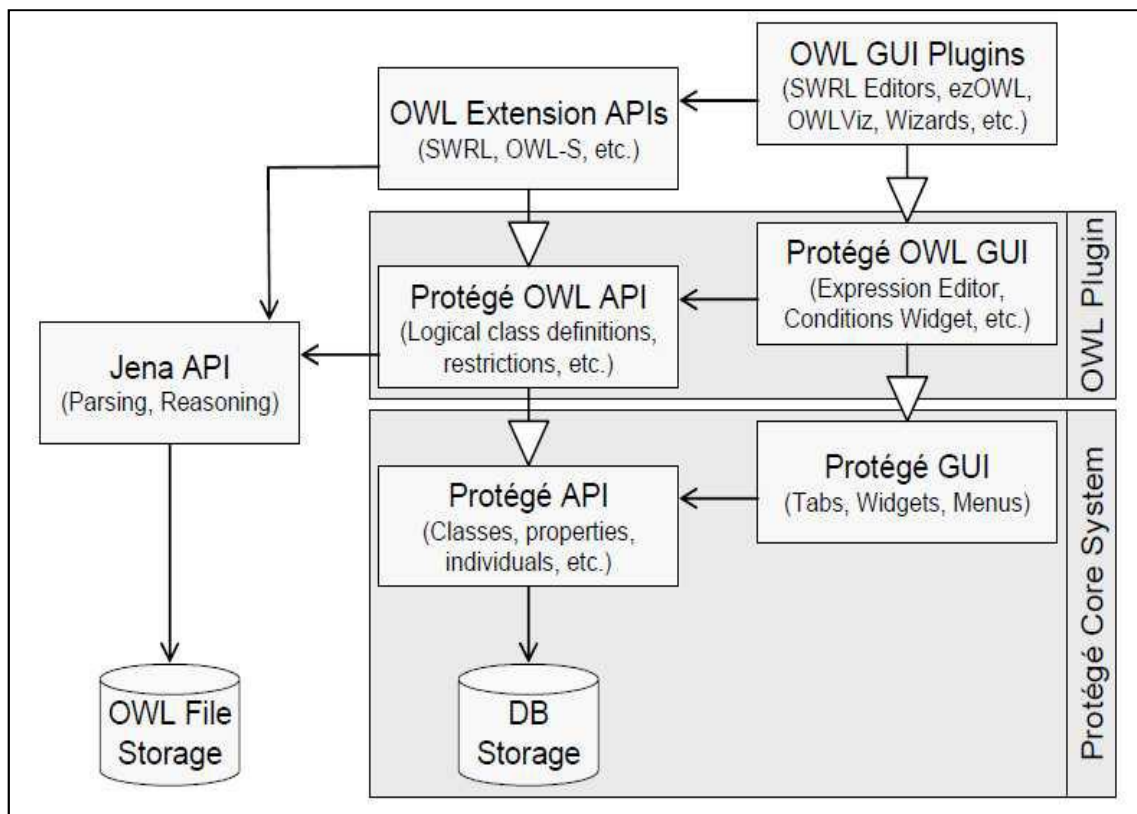
ακολουθούν το πλαίσιο του πρωτοκόλλου Open Knowledge Base Connectivity (OKBC).

- Η εφαρμογή «**Protege – OWL Editor**» που δίνει τη δυνατότητα στους χρήστες να κατασκευάσουν οντολογίες για τον Σημασιολογικό Ιστό και συγκεκριμένα για τη γλώσσα OWL.

4.2 Η Αρχιτεκτονική του Protege

Όπως στα περισσότερα εργαλεία μοντελοποίησης, η αρχιτεκτονική του Protege είναι καθαρά χωρισμένη στο τμήμα του «μοντέλου» και στο τμήμα της «απεικόνισης». Το «μοντέλο» είναι ο μηχανισμός εσωτερικής αναπαράστασης για οντολογίες και βάση γνώσης. Τα εργαλεία της «απεικόνισης» παρέχουν ένα περιβάλλον για απεικόνιση και επεξεργασία του μοντέλου της γνώσης. Το «μοντέλο» του Protege βασίζεται σε ένα απλό, αλλά ευέλικτο μετα-μοντέλο, το οποίο συγκρίνεται με αυτό των αντικειμενοστραφών και βασισμένων σε πλαίσιο (frame based) συστημάτων. Γενικά, μπορεί να αναπαριστά οντολογίες που αποτελούνται από κλάσεις, ιδιότητες, χαρακτηριστικά των ιδιοτήτων και στιγμιότυπα. Το Protege παρέχει μια ανοικτή προγραμματιστική διεπαφή βασισμένη σε JAVA για να εκτελεί ερωτήματα και να διαχειρίζεται τα μοντέλα. Σημαντικό είναι επίσης το γεγονός ότι το μεταμοντέλο του Protege αποτελεί το ίδιο μια οντολογία, με κλάσεις που αναπαριστούν κλάσεις, ιδιότητες κλπ., κάτι που του δίνει τη δυνατότητα εύκολα να μπορεί να επεκταθεί και να προσαρμοστεί σε άλλες αναπαραστάσεις. Χρησιμοποιώντας τα εργαλεία της «απεικόνισης» του Protege, οι σχεδιαστές οντολογιών δημιουργούν κλάσεις, αναθέτουν ιδιότητες στις κλάσεις και περιορίζουν τα χαρακτηριστικά των ιδιοτήτων ορισμένων κλάσεων. «Φορτώνοντας» μια οντολογία, το Protege είναι σε θέση να παράγει επιφάνειες αλληλεπίδρασης με το χρήστη, μέσω των οποίων θα μπορούν οι χρήστες να δημιουργούν νέα στιγμιότυπα. Για κάθε κλάση της οντολογίας και για κάθε ιδιότητα της κλάσης δημιουργείται μια φόρμα με επεξεργάσιμα συστατικά (components – widgets). Οι φόρμες που δημιουργούνται μπορούν επιπλέον να διαμορφωθούν με το εργαλείο «Protege Forms Editor», όπου οι χρήστες μπορούν να επιλέξουν εναλλακτικά widgets. Εκτός των widgets που παρέχει η βιβλιοθήκη του Protege, το Protege διαθέτει μια ευέλικτη αρχιτεκτονική που επιτρέπει στους προγραμματιστές να αναπτύσσουν τα δικά τους widgets, τα οποία στην συνέχεια

μπορούν να συνδεθούν με τον πυρήνα του συστήματος της εφαρμογής. Ανάλογα μπορούν να αναπτυχθούν πλαίσια επιφανειών αλληλεπίδρασης με το χρήστη σε πλήρες μέγεθος (tabs), τα οποία θα περιέχουν διάφορα άλλα συστατικά. Εκτός από τη βασική συλλογή των tabs για τη συγγραφή κλάσεων, ιδιοτήτων, φορμών και στιγμιότυπων, υπάρχει και μια βιβλιοθήκη με επιπλέον tabs για την εκτέλεση ερωτημάτων, για την πρόσβαση σε αποθήκες δεδομένων, για τη γραφική απεικόνιση οντολογιών και για την διαχείριση των εκδόσεων των οντολογιών.



Εικόνα 12 Η αρχιτεκτονική των συστημάτων του Protege

4.2.1 Protege – Frames Editor

Στα χαρακτηριστικά του Protege – Frames Editor ανήκουν τα εξής :

- Ένα περιβάλλον εργασίας που μπορεί να προσαρμοστεί ώστε να δώσει στους χρήστες τη δυνατότητα να μοντελοποιήσουν τη γνώση και να εισάγουν δεδομένα με φόρμες που να αντιπροσωπεύουν το εκάστοτε πεδίο.
- Μια αρχιτεκτονική plugin που επιτρέπει την ενσωμάτωση στοιχείων, όπως γραφικά, πολυμέσα, χρήση τυποποιήσεων (π.χ. RDF, XML, HTML), αλλά και πρόσθετων υποστηρικτικών εργαλείων για επεξεργασία ή απεικόνιση οντολογιών, εξαγωγής συμπερασμάτων, κ.ά.
- Μια προγραμματιστική διεπαφή (API) βασισμένη σε JAVA που επιτρέπει στα plugins και σε άλλες εφαρμογές να έχουν πρόσβαση στη χρήση και την απεικόνιση οντολογιών που έχουν δημιουργηθεί μέσω του Protege – Frames Editor.

4.2.2 Protege – OWL Editor

Η εφαρμογή αυτή επιτρέπει στους χρήστες :

- Να φορτώνουν και να σώζουν OWL και RDF οντολογίες.
- Να συντάσσουν και να απεικονίζουν κλάσεις, ιδιότητες και SWRL κανόνες.
- Να ορίζουν λογικά χαρακτηριστικά κλάσεων ως OWL παραστάσεις.
- Να εκτελούν λογικούς συλλογισμούς, όπως τα Description Logic Classifiers.
- Να συντάσσουν την OWL με στόχο την οικοδόμηση του Σημασιολογικού Ιστού.

4.3 Τεχνολογίες Εξόρυξης Οντολογικής Γνώσης

Σε διαφορετικά μοντέλα αναπαράστασης γνώσης ή σε διαφορετικά σχήματα είναι γραμμένη η γνώση που προέρχεται από τα εργαλεία εξόρυξης γνώσης καθώς και η

γνώση των ειδικών που περιέχεται σε μια οντολογία. Σε κάθε περίπτωση, η ορολογία που χρησιμοποιείται διαφέρει, όπως και η σημασιολογία. Για την ολοκλήρωση της γνώσης που εξάγεται με τους διάφορους μηχανισμούς εξόρυξης, με τη γνώση που περιέχεται στις οντολογίες, απαιτείται, αρχικά, η αναπαράσταση της γνώσης σε ένα κοινό σχήμα και κατά συνέπεια μηχανισμοί μετάφρασης από τα διάφορα σχήματα προς το κοινό και στη συνέχεια η συγχώνευση των διαφόρων πηγών. Το δεύτερο βήμα περιλαμβάνει προβλήματα σύγκρουσης απόψεων και επίλυσης διαφορών και δεν αποτελεί στόχο της συγκεκριμένης έρευνας.

Πρωτίστως, θα επιχειρήσουμε να καταγράψουμε, να ερευνήσουμε και να αναλύσουμε τις παρούσες προσπάθειες μετάφρασης σχημάτων γνώσης καθώς και των τεχνικών για προσθήκη γνώσης σε βάσεις γνώσης. Διακρίνουμε στην προσπάθεια αυτή δύο βασικές κατευθύνσεις:

- Τον ορισμό γλωσσών επερωτήσεων σε σχήματα γνώσης
- Την ανάπτυξη συστημάτων για τη διαχείριση και επερώτηση σχημάτων γνώσης

4.3.1 Γλώσσες Επερωτήσεων

Η πρώτη κατηγορίας εμπεριέχει *γλώσσες επερωτήσεων* προς RDF σχήματα, όπως η **RQL** και η **SPARQL**. Η οντολογική γνώση, με τις γνώσεις αυτές που έχει αναπαρασταθεί σε RDF μπορεί εύκολα να ερωτηθεί, να αποκοπεί, να μετασχηματιστεί ή να συγκεντρωθεί. Συχνά, προέρχεται από κάποια Υπηρεσία Ιστού και είτε ενημερώνεται από κάποιον ειδικό, είτε είναι το αποτέλεσμα μιας διαδικασίας εξόρυξης γνώσης.

RQL

Οι RQL, SeRQL και eRQL αποτελούν την οικογένεια γλωσσών *RQL*. Το κοινό χαρακτηριστικό μεταξύ αυτών των γλωσσών είναι ότι υποστηρίζουν το συνδυασμό ερωτημάτων στο σχήμα και τα δεδομένα του μοντέλου RDF. Ωστόσο, το μοντέλο των δεδομένων διαφέρει από αυτό της RDFS. Στις κύριες διαφορές ανήκει το ότι δεν

επιτρέπονται κύκλοι στην ιεραρχία και ότι για κάθε ιδιότητα χρειάζεται να ορίζεται το domain και το range της. Η RQL αποτελεί μια από τις γλώσσες ερωτημάτων σε RDF αρχεία. Έχει πραγματοποιηθεί στο RDF Suite του ICS-FORTH και ένα τμήμα της είναι διαθέσιμο στο σύστημα Sesame. Αποτελεί πολύ εκφραστική γλώσσα ερωτημάτων, αλλά τα σημασιολογικά στοιχεία που εμπεριέχει δε συμφωνούν απόλυτα με αυτά της RDF.

SPARQL

Η οικογένεια *SPARQL* αποτελείται από τις SPARQL, SquishQL, RDQL και TriQL. Τα κοινά χαρακτηριστικά μεταξύ αυτών των γλωσσών είναι ότι αντιλαμβάνονται τα δεδομένα RDF ως απλές τριάδες χωρίς κάποιο σχήμα ή άλλη οντολογική πληροφορία εκτός κι αν αυτή έχει ρητά οριστεί στο RDF έγγραφο. Οι SquishQL και RDQL προσφέρουν απλά μια πρόσβαση σε RDF γράφους χωρίς περαιτέρω δυνατότητες. Βασικό στόχο, όπως εξάλλου και στην RQL, αποτελεί η ευκολία στη χρήση και η ομοιότητα με την SQL. Η SquishQL – όπως και η RDQL – δε μπορεί να μεταχειριστεί σύνθετα ερωτήματα. Η RDQL έχει σύνταξη παρόμοια με αυτή της SQL. Είναι μια εξέλιξη της SquishQL επηρεασμένη από την rdfDB. Μια υλοποίηση της RDQL, η RDQLPlus παρέχει μια επέκταση στην RDQL χρησιμοποιώντας την inference engine της Jena. Η γλώσσα αυτή που προκύπτει ονομάζεται RIDIQL. Η RDQL υποστηρίζεται από την Jena, μέσω του RAP, το οποίο είναι ένα RDF API για τη γλώσσα PHP, και έχει υποβληθεί στο W3C για προτυποποίηση. Η TriQL επεκτείνει την RDQL υποστηρίζοντας επιπλέον την υποβολή ερωτημάτων σε Γράφους.

4.3.2 Συστήματα μετατροπής σχημάτων γνώσης

Τα **συστήματα που μετατρέπουν τα σχήματα γνώσης** από γνωστά εργαλεία (π.χ. το εργαλείο εξόρυξης γνώσης WEKA) σε εύκολα αναγνώσιμες μορφές (π.χ. PMML) ανήκουν στην ίδια κατηγορία με τις γλώσσες επερωτήσεων. Τα συστήματα αυτά λειτουργούν ως wrappers και διευκολύνουν την ολοκλήρωση Σχημάτων.

4.3.3 Προγραμματιστικές Διεπαφές

Εν κατακλείδι, εμπεριέχονται και οι *προγραμματιστικές διεπαφές (APIs)*, όπως τα **Jena**, **Protege-OWL**, **JeSS** και **KAON** που εφαρμόζονται από μεγαλύτερα συστήματα διαχείρισης γνώσης για την υποβολή ερωτήσεων και την εξαγωγή κανόνων (inference engines).

Protege-OWL API

Το *Protege-OWL API* είναι μια open source βιβλιοθήκη της JAVA για την γλώσσα OWL και το μοντέλο RDF. Το API παρέχει τις κατάλληλες κλάσεις και μεθόδους για να φορτωθούν και να αποθηκευτούν OWL οντολογίες, να διερευνηθούν και να ελεγχθούν OWL μοντέλα δεδομένων, να υλοποιηθούν διαδικασίες συλλογιστικής(reasoning) βασισμένες στις μηχανές συμπερασμού και τις περιγραφικές λογικές. Επιπλέον, το Protege-OWL API είναι βελτιστοποιημένο για την εφαρμογή γραφικών διεπαφών χρήστη. Το Protege-OWL API έχει σχεδιαστεί για να χρησιμοποιείται για δυο διαφορετικούς λόγους : (Herman, 2007)

- Για την ανάπτυξη των στοιχείων που εκτελούνται στο εσωτερικό της διεπαφής του χρήστη του Protege-OWL editor.
- Για την ανάπτυξη stand-alone εφαρμογών (π.χ., εφαρμογές Swing, Servlets, ή Eclipse plugins).

Jena API

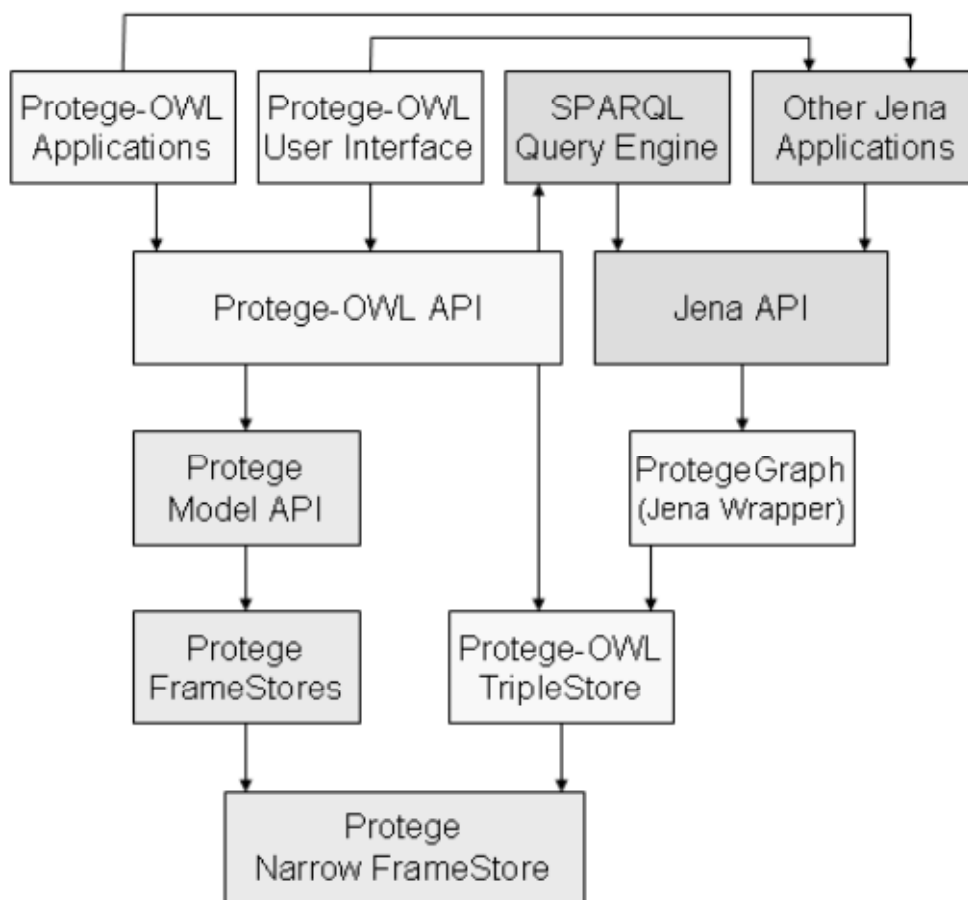
Το *Jena API* αποτελεί μια προγραμματιστική διεπαφή σε JAVA, είναι ανοιχτού κώδικα και παρέχει υποστήριξη για τη διαχείριση οντολογιών και προγραμματισμό εφαρμογών Σημαιολογικού Ιστού. Αναπτύσσεται κατά κύριο λόγο από τα εργαστήρια της Hewlett Packard. Η πρώτη έκδοση της Jena ήταν διαθέσιμη το 2000. Η αρχιτεκτονική της επανασχεδιάστηκε και η Jena2 έγινε διαθέσιμη τον Αύγουστο του 2003, αποτελώντας πλέον την πρακτικότερη λύση για προγραμματιστική

διαχείριση οντολογιών. Η Jena παρέχει ουσιαστικά μια συλλογή οντολογικών εργαλείων. Εκτός από το API για διαχείριση οντολογιών, στη Jena ενσωματώνονται επίσης τα εξής:

- Υλοποίηση του πρωτοκόλλου SPARQL (ARQ). Η γλώσσα ερωτημάτων της Jena ήταν η RDQL μέχρι την έκδοση 2.2. Από την 2.3 και μετά γλώσσα που υιοθετείται είναι η SPARQL ενώ η RDQL εξακολουθεί να υποστηρίζεται για λόγους προς τα πίσω συμβατότητας.
- Διαχείριση τριάδων αποθηκευμένων σε Βάση Δεδομένων (SDB).
- Εξυπηρετητής εγγράφων RDF μέσω HTTP (Joseki). Ο Joseki μπορεί να διαχειρίζεται μέσω της διεπαφής Jena τριάδες οι οποίες μπορεί να υπάρχουν σε έγγραφα RDF σε οποιαδήποτε μορφή διάταξης (RDF/XML, N3, Turtle), είτε σε μοντέλο αποθηκευμένο σε Βάση Δεδομένων σύμφωνα με τις προδιαγραφές της Jena.

Η διεπαφή Jena μπορεί να λειτουργήσει επίσης ως εξυπηρετητής υπηρεσιών συλλογιστικής καθώς ενσωματώνει μηχανή εξαγωγής συμπερασμάτων η οποία βασίζεται σε κανόνες. Μπορεί να εξάγει συμπεράσματα από τα μοντέλα που της δίνονται και υποστηρίζει σχεδόν κάθε έκφραση RDFS. Όσον αφορά όμως την OWL DL, η Jena είναι ακόμα ημιτελής. Παρέχει τρεις υλοποιήσεις για OWL reasoners, standard, mini και macro αλλά κανένας από αυτούς δεν καλύπτει την εκφραστικότητα της OWL DL. Υποστηρίζεται ωστόσο η **διαλειτουργικότητα DIG** για τη σύνδεση των μοντέλων με εξωτερικό reasoner. Η **διεπαφή DIG** παρέχει ενιαία πρόσβαση στους DL Reasoners ορίζοντας ένα απλό πρωτόκολλο (βάσει HTTP PUT / GET) μαζί με ένα σχήμα XML που περιγράφει μια σημασιολογική γλώσσα και συνοδευτικές ενέργειες. Η διεπαφή δεν προορίζεται ως μια βαρέων βαρών προδιαγραφή της υπηρεσίας συλλογισμού. Αντίθετα, παρέχει ένα ελάχιστο σύνολο λειτουργιών (π.χ. έλεγχος ικανοποιησιμότητας, υπαγωγή και λογική ταξινόμηση) που έχουν αποδειχθεί ότι είναι χρήσιμες σε πολλές εφαρμογές. Πολλοί reasoners συμπεριλαμβανομένων των CEL, FACT++, Pellet και RacerPro παρέχουν υποστήριξη για την διεπαφή DIG. Ένας αριθμός από τις εφαρμογές ανάπτυξης και διαχείρισης οντολογιών (υπεριλαμβανομένου του Protege) μπορούν να χρησιμοποιήσουν τη διασύνδεση DIG

για να επικοινωνούν με εξωτερικούς reasoners. Η προγραμματιστική διεπαφή της Jena κάνει επίσης χρήση της διασύνδεσης DIG για την υλοποίηση διαδικασιών συλλογιστικής πάνω σε OWL οντολογίες. Το Protege είναι στενά συνδεδεμένο με το Jena API. Αυτή η σχέση επιτρέπει τη χρήση συγκεκριμένων μεθόδων του Jena σε χρόνο εκτέλεσης. Το βασικό σημείο είναι πως και τα δύο συστήματα λειτουργούν σε ένα χαμηλό επίπεδο αναπαράστασης τριάδων. Ο μηχανισμός αποθήκευσης πλαισίων (frame store mechanism) του Protege έχει ενσωματωθεί από την αποθήκευση τριάδων του Protege-OWL (Protege triple store). Οι αντίστοιχοι μηχανισμοί στο Jena είναι τα Graph και Model. Η αποθήκευση τριάδων του Protege ενσωματώνεται σε έναν γράφο του Jena, έτσι ώστε κάθε λειτουργία ανάγνωσης του Jena στην ουσία να επιδρά στις τριάδες του Protege. Έτσι, το Jena μπορεί να χρησιμοποιηθεί για παράδειγμα για την επερώτηση οντολογιών μέσα από τη μηχανή SPARQL που διαθέτει, ή για άλλες λειτουργίες σε οντολογίες, οι οποίες έχουν υλοποιηθεί με χρήση του Protege. Για την επεξεργασία των τριάδων, χρησιμοποιείται το συμβατικό Protege-OWL API.



Εικόνα 13 Αρχιτεκτονική ενσωμάτωσης εργαλείων Jena και Protege

Στη βάση των εργαλείων της Jena είναι η δομή των γράφων (Graph Layer) η οποία αποτελείται από τριάδες (subject, predicate, object). Η αποθήκευση των γράφων μπορεί να γίνει είτε στην κύρια μνήμη είτε σε Βάση Δεδομένων. Το επίπεδο αφαίρεσης στην αποθήκευση επιτρέπει τη συνεργασία με οποιαδήποτε μορφή αποθήκευσης. Πάνω από το επίπεδο των γράφων βρίσκεται το επίπεδο των ενισχυμένων γράφων (Enh-Graph Layer) το οποίο αποτελεί ουσιαστικά τη λειτουργικότητα με τα παραπάνω επίπεδα. Το επίπεδο αυτό σχεδιάστηκε ώστε να επιτρέπει δημιουργία όψεων (views) πάνω στο γράφο και παράλληλα είναι σημαντικό ώστε οι δηλώσεις του RDFS να μπορούν να αντιστοιχηθούν σε συναρτήσεις Java. Το επίπεδο του μοντέλου (Model Layer) είναι αυτό με το οποίο ο προγραμματιστής μπορεί να έχει πρόσβαση μέσω ενός πλούσιου συνόλου από μεθόδους και συναρτήσεις, τόσο στους κόμβους του γράφου, όσο και στην ίδια τη δομή του γράφου. Οι γλώσσες περιγραφής οντολογιών που υποστηρίζονται από τη Jena είναι οι RDF(S), DAML+OIL και OWL. Σημαντική είναι και η παρουσία του Joseki SPARQL Server. Ο Joseki είναι ανοιχτού κώδικα εξυπηρετητής, δημιουργήθηκε και αναπτύσσεται από την Hewlett Packard, από τους συντελεστές που αναπτύσσουν την Jena. Η ιδέα πίσω από τον Joseki είναι η δημοσίευση στο Διαδίκτυο δεδομένων που είναι αποθηκευμένα στο σύστημα αποθήκευσης γράφων της Jena. Με τη χρήση του Joseki, μια Βάση Γνώσεως δημιουργημένη σε Jena αλλά και οποιαδήποτε αποθήκη σημασιολογικών δεδομένων, αποκτά υποστήριξη ερωτημάτων σε SPARQL μέσω διαδικτυακής διεπαφής (GET και POST ερωτημάτων). (Στοιλος, 2008)

4.3.4 Συστήματα Διαχείρισης Βάσεων Γνώσης

Στη δεύτερη κατηγορία περιέχονται τα συστήματα που επιτρέπουν τη διαχείριση βάσεων γνώσης με χρήση κάποιου από τους προηγούμενους μηχανισμούς ερωτήσεων ή με την υλοποίηση δικών τους μηχανισμών. Τα συστήματα αυτά μπορούν να διαχειριστούν ένα ή περισσότερα αρχεία RDF ή OWL και να απαντήσουν σε ερωτήσεις. Κατά την απάντηση των ερωτήσεων λαμβάνουν υπόψη τους ιδιότητες όπως η μεταβατική (Transitive Property), η αντίθεση (inverseOf), η υπόταξη (subsumption), η ισοδύναμη τάξη κλπ. και εμφανίζουν διαφορετικούς βαθμούς

πληρότητας και ορθότητας στις απαντήσεις. Χαρακτηριστικά συστήματα σε αυτή την κατηγορία είναι τα εξής :

- Sesame, που αποθηκεύει και διαχειρίζεται RDF και RDF Schema αρχεία (στη μνήμη ή σε βάση δεδομένων) και υποστηρίζει ερωτήσεις σε SeRQL, RQL και RDQL.
- OWLJessKB, που διαχειρίζεται έγγραφα OWL στη μνήμη χρησιμοποιώντας το Jess API.
- DLDB-OWL19, που αποθηκεύει δεδομένα OWL σε σχεσιακή Βάση Δεδομένων. Πριν τη δημιουργία της βάσης χρησιμοποιεί τον FaCT reasoner και εξάγει αυτόματα κανόνες ιεραρχίας τάξεων (subsumption) από τα OWL δεδομένα. Οι κανόνες αυτοί εμπλουτίζουν τη Βάση Δεδομένων με απόψεις που διευκολύνουν την απάντηση ερωτήσεων που αφορούν στιγμιότυπα. (extensional queries).
- TRIPLE, που διαχειρίζεται ερωτήσεις σε RDF αρχεία υποστηρίζοντας μεταβατικές ιδιότητες (subClassOf, subPropertyOf κλπ.). Το TRIPLE επίσης δεν διαθέτει API κάτι που το καθιστά δύσκολα επεκτάσιμο.
- OntoTrack, που υποστηρίζει RDF, DAML+OIL και OWL αρχεία και περιορισμένους τύπους ερωτήσεων και ενσωματώνει τον reasoner Racer. Η πλατφόρμα Racer λειτουργεί ως server που μπορεί να εξυπηρετεί αιτήματα reasoning μέσω HTTP ή TCP, δεν προσφέρει όμως API για τη δημιουργία επιπλέον ερωτήσεων από αυτά που υποστηρίζει (custom queries).

Στόχος των συστημάτων αυτών είναι η αποθήκευση στη μνήμη ή σε μια Βάση

Δεδομένων ενός ή περισσότερων αρχείων RDF, DAML+OIL ή OWL και η υποβολή ερωτήσεων σε αυτά. Οι ερωτήσεις συνήθως στοχεύουν στην αναζήτηση στιγμιότυπων που ικανοποιούν καθορισμένα κριτήρια αλλά και στην εξαγωγή γενικότερων συμπερασμάτων.

4.4 Συλλογιστική Οντολογιών (reasoners)

Παρ' όλες τις σαφείς αναλογίες μεταξύ των βάσεων δεδομένων και των OWL οντολογιών, υπάρχουν και πολυάριθμες σημαντικές διαφορές. Η OWL έχει μια αποκαλούμενη «σημασιολογία ανοιχτού κόσμου», στην οποία τα δεδομένα που απουσιάζουν, αντιμετωπίζονται ως άγνωστες και όχι ως λανθασμένες. Παράλληλα, τα αξιώματα της OWL συμπεριφέρονται σαν κανόνες συμπερασμού, και όχι ως περιορισμοί βάσεων δεδομένων. Σε αντίθεση με τις βάσεις δεδομένων, η OWL δεν κάνει καμία «Υπόθεση Μοναδικού Ονόματος» (UNA). Συγχρόνως, τα εργαλεία οντολογίας συνήθως δεν απορρίπτουν τις αλλαγές που οδηγούν στο να γίνει η οντολογία πλήρως ή εν μέρει ασυνεπής (inconsistent), απλά παρέχουν μια κατάλληλη προειδοποίηση. Οι οντολογίες μπορεί να είναι πολύ μεγάλες και σύνθετες³ με αποτέλεσμα δημιουργία και η διατήρηση τέτοιων οντολογιών να είναι πολύ δαπανηρή και χρονοβόρα, και η παροχή εργαλείων και υπηρεσιών για την υποστήριξη του μηχανισμού μιας οντολογίας, είναι μεγάλης σπουδαιότητας, και για το κόστος αλλά και για την ποιότητα της προκύπτουσας οντολογίας. Επιπλέον, όπως έχουμε δει και παραπάνω, η απάντηση ενός ερωτήματος στην OWL δεν είναι απλά θέμα ελεγχόντων δεδομένων, αλλά μπορεί να απαιτεί την εκτέλεση σύνθετου συλλογισμού (reasoning). Επομένως, ο συλλογισμός οντολογιών παίζει έναν πρωταρχικό ρόλο και στην ανάπτυξη υψηλής ποιότητας οντολογιών, και στη χρήση των οντολογιών από τις εφαρμογές. Παρά την πολυπλοκότητα του συλλογισμού στις OWL οντολογίες, τα ιδιαίτερος βελτιστοποιημένα συστήματα συλλογισμού λογικών περιγραφής, όπως το FaCT++, το RACER, και το Pellet, έχουν αποδειχθεί πολύ αποτελεσματικά στην πράξη. (Αγγελόπουλος, 2009)

Σύγχρονα εργαλεία ανάπτυξης οντολογιών, όπως το SWOOP, το Protege 4, και το TopBraid Composer, εφαρμόζουν συστήματα συλλογισμού λογικών περιγραφής για να παρέχουν πληροφορίες στο χρήστη για τις λογικές επιπτώσεις του σχεδιασμού τους. Συνήθως, αυτό περιλαμβάνει (τουλάχιστον) προειδοποιήσεις για ασυνέπειες και συνώνυμα. Ασυνεπής ή αλλιώς ανικανοποίητη (unsatisfiable), είναι μια κλάση της

³ για παράδειγμα, η πολύ γνωστή οντολογία κλινικών όρων SNOMED περιλαμβάνει περισσότερα από 400.000 ονόματα κλάσεων.

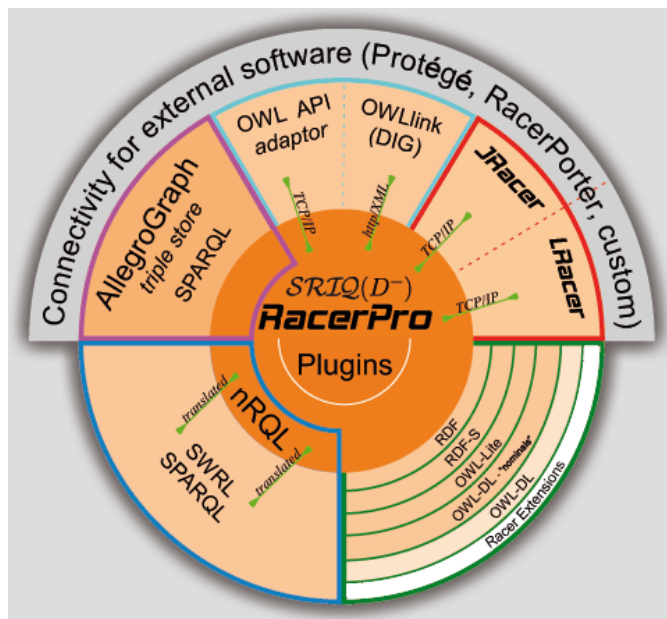
οποίας η περιγραφή εμπεριέχει τέτοιους περιορισμούς, ώστε να είναι αδύνατη η ύπαρξη κάποιου στιγμιότυπου της κλάσης αυτής. (Κορμπάκης, 2009)

Δημοφιλείς και απαραίτητοι έχουν γίνει οι reasoners, κατά την υλοποίηση του Σημασιολογικού Ιστού. Οι reasoners παρέχουν πολύ αποδοτικά κάποιες υπηρεσίες συμπερασμού. Ορισμένες από τις σημαντικότερες υπηρεσίες συμπερασμού είναι ο έλεγχος συνέπειας της βάσης γνώσης (consistency checking), η κατηγοριοποίηση των κλάσεων (classification) και ο υπολογισμός των κλάσεων στις οποίες ανήκει κάθε στιγμιότυπο (instance checking). Οι πιο δημοσιευμένοι reasoners είναι οι ακόλουθοι :

- Racer
- FaCT++
- Pellet
- KAON2

4.4.1 Racer

Ο *RacerPro* αποτελεί την εμπορική έκδοση του λογισμικού Racer. Ο **Racer (Renamed Abox and Concept Expression Reasoner)** θεωρείται ο πρώτος reasoner που κυκλοφόρησε για την γλώσσα OWL. Στη σύγχρονη εποχή, ο πιο διαδεδομένος reasoner είναι και ένας από τους γρηγορότερους. Η μηχανή Racer εφαρμόζει μια βελτιωμένη έκδοση του tableau calculus για πολύ εκφραστικές Description Logics, όπως είναι η OWL-DL. Όσες γλώσσες μπορεί να υποστηρίξει προέρχονται από το χώρο των Description Logics και εμπεριέχουν, μεταξύ άλλων, ιεραρχίες εννοιών και σχέσεων, τελεστές αριθμητικών περιορισμών, συμμετρικές και μεταβατικές σχέσεις.



Ο συμπερασμός νέας γνώσης προκύπτει λόγω της σαφώς ορισμένης σημασιολογίας της εκάστοτε DL γλώσσας.

Μερικές βασικές λειτουργίες που υποστηρίζει η μηχανή Racer είναι οι ακόλουθες:

- Έλεγχος συνέπειας μιας οντολογίας (Consistency checking)
- Εντοπισμός έμμεσων υποκλάσεων (Classification of taxonomy)
- Τοποθέτηση στιγμιότυπων σε άλλες κλάσεις (Individual Inference)
- Αλγεβρικός συμπερασμός (Algebraic Reasoning)

Εν κατακλείδι, είναι απαραίτητο να επισημανθεί ότι η μηχανή Racer δεν υποθέτει ότι τα στιγμιότυπα με διαφορετικά ονόματα είναι μεταξύ τους διαφορετικά. Αυτό αποτελεί κύριο συστατικό ενός αποτελεσματικού reasoner για οντολογίες, αφού οι οντολογίες λειτουργούν με την υπόθεση του κλειστού κόσμου. Το χαρακτηριστικό αυτό της μηχανής Racer σημαίνει ότι είναι εφικτό να επιτρέψει την ταύτιση δύο στιγμιότυπων με διαφορετικό όνομα, αν φυσικά αυτό προκύψει από τη διαδικασία συμπερασμού. (Παπαταξιάρχης, 2008)

4.4.2 FaCT++

Μετεξέλιξη του FaCT θεωρείται ο *FaCT++*, ικανός να υποστηρίζει συμπερασμό για τη γλώσσα OWL-DL. Ανάλογες είναι οι υπηρεσίες που προσφέρει με του RacerPro. Ειδικότερα, στηρίζεται σε tableau αλγορίθμους και υποστηρίζει ένα πλήθος από πρόσθετα χαρακτηριστικά (π.χ. υποστήριξη nominals). Βέβαια, είναι χρήσιμο να αναφερθεί ότι το μεγαλύτερο μειονέκτημα του Fact++ είναι η αδυναμία του για πλήρη συμπερασμό πάνω από ABox.

4.4.3 Pellet

Η γλώσσα JAVA υλοποιεί τη μηχανή συμπερασμού *Pellet*. Επίσης, αποτελεί λογισμικό ανοικτού κώδικα (open-source) ικανό να διαχειριστεί οντολογίες υψηλής εκφραστικότητας. Ειδικότερα, στηρίζεται σε βελτιστοποιημένους tableau αλγόριθμους, υπεύθυνους για το χειρισμό βάσεων γνώσης εκφρασμένες σε περιγραφικές λογικές. Συγχρόνως, υποστηρίζει ένα σύνολο από πρόσθετα χαρακτηριστικά, όπως υποστήριξη UNA, συλλογιστική βασισμένη σε CWA, εκτέλεση SPARQL ερωτημάτων κλπ. Επίσης, η μηχανή Pellet περιέχει ένα μηχανισμό επεξήγησης που έχει σκοπό τη διευκόλυνση της διαδικασίας σχεδιασμού και ανάπτυξης οντολογιών, μέσω του αποτελεσματικού και γρήγορου εντοπισμού λαθών και ασυνεπειών της βάσης γνώσης. Συνεπώς, αντίθετα με τις περισσότερες μηχανές συμπερασμού για περιγραφικές λογικές, η μηχανή Pellet όχι μόνο εντοπίζει σημεία ασυνέπειας αλλά την αιτία που οδήγησε τη βάση γνώσης σε μη-ικανοποιησιμότητα (unsatisfiability). Εκτενέστερα, ο Pellet προσφέρει στο χρήστη πρόσθετη γνώση, όπως αξιώματα (axioms) και περιορισμούς (restrictions), με στόχο να διευκολύνει την επίλυση του προκληθέντος προβλήματος. Εν κατακλείδι, η συγκεκριμένη μηχανή συμπερασμού προσφέρει τη συνατότητα στη χρήση τύπων δεδομένων (datatypes) προδιαγεγραμμένων από το συντακτικό της γλώσσας XML, αλλά και ορισμένων από το χρήστη (user-defined). (Gelfond & Horrocks, 2003)

4.4.4 KAON2

Το εργαλείο KAON2 αποτελεί το διάδοχο του KAON και αντίθετα με τις περισσότερες μηχανές συμπερασμού, δεν βασίζεται σε κάποιον tableau αλγόριθμο. Στην κυριολεξία, ο KAON2 αποτελεί ένα υβριδικό εργαλείο συμπερασμού ικανό να χειριστεί οντολογίες και προγράμματα Datalog με διάζευξη (Disjunctive Datalog). Πραγματοποιεί αλγόριθμους που ανάγουν την Περιγραφική Λογική SHIQ(D) σε Datalog με διάζευξη, εκμεταλλευόμενοι γνωστές τεχνικές από το πεδίο των επαγωγικών βάσεων δεδομένων (π.χ. magic sets). Επιπρόσθετα, η μηχανή δύναται να χειριστεί κανόνες SWRL, οντολογίες εκφρασμένες σε F-Logic και SPARQL επερωτήσεις. Συγχρόνως, προσφέρει μια διεπαφή για την υλοποίηση οντολογιών με

κανόνες. Ωστόσο, ο KAON2 δεν υποστηρίζει όλα τα χαρακτηριστικά που παρέχονται από τη γλώσσαOWL, όπως είναι οι αριθμητικοί περιορισμοί (cardinality restrictions). (Παπαταξιάρχης, 2008) (KAON, 2008)

4.5 Διαχείριση Οντολογικής Γνώσης με Υπηρεσίες Ιστού

Η Υπηρεσία Ιστού εστιάζει και περιγράφει ένα σύνολο λειτουργιών που είναι προσβάσιμες μέσω του Διαδικτύου, με έναν προτυποποιημένο τρόπο ανταλλαγής XML μηνυμάτων, καταγράφοντας όλες τις απαραίτητες λεπτομέρειες που είναι απαραίτητες για να εφαρμοστεί η εκάστοτε Υπηρεσία, όπως είναι τα πρωτόκολλα μηνυμάτων και μεταφοράς δεδομένων. Επιπρόσθετα, η διασύνδεση αποκρύπτει τις λεπτομέρειες της υλοποίησης και προσφέρει τη δυνατότητα στην Υπηρεσία να χρησιμοποιηθεί ανεξαρτήτως της πλατφόρμας υλικού και λογισμικού στην οποία έχει υλοποιηθεί. Αυτό επιτρέπει τη «χαλαρή» διασύνδεση των διαφόρων Υπηρεσιών Ιστού, είτε είναι στατική, είτε δυναμική, δημιουργώντας έτσι νέες Υπηρεσίες Ιστού από άλλες ήδη υπάρχουσες, εφαρμόζοντας κατά τον καλύτερο δυνατό τρόπο τις αρχές της τμηματοποιημένης ανάπτυξης του Παγκόσμιου Ιστού.

Οι δυνατότητες μιας Υπηρεσίας Ιστού είναι οι εξής :

- Περιγραφή μέσω μιας γλώσσας περιγραφής.
- Δημοσίευση σε ένα ευρετήριο Υπηρεσιών Ιστού.
- Ανακάλυψη μέσω ενός προτυποποιημένου μηχανισμού.
- Κλήση μέσω μιας δηλωμένης προγραμματιστικής διεπαφής.
- Συνδυασμός με άλλες Υπηρεσίες Ιστού.

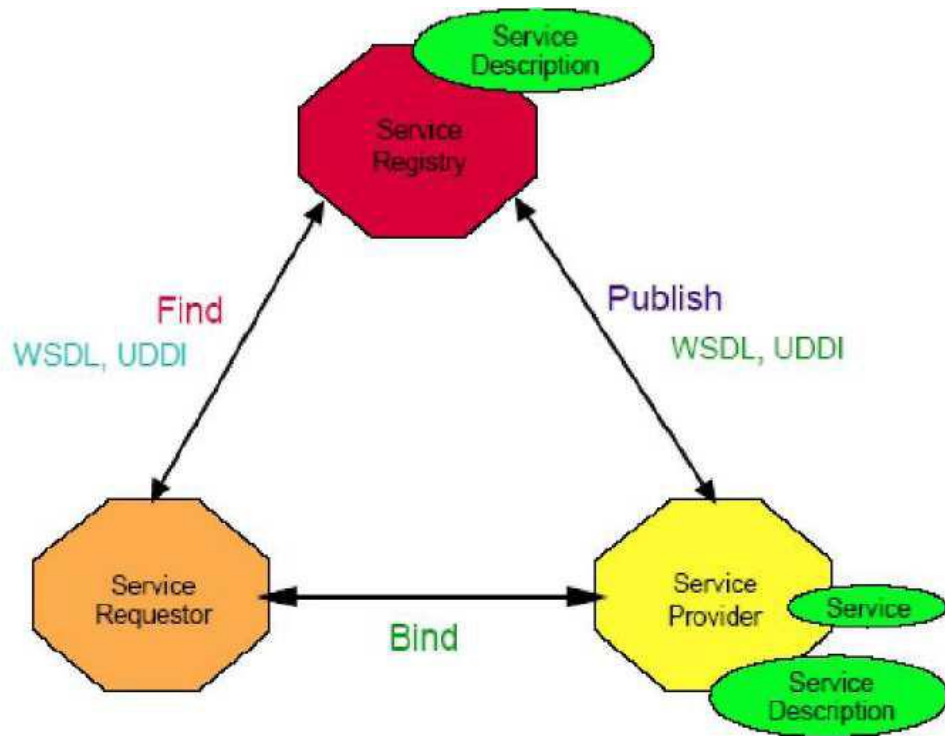
Ωστόσο, μια Υπηρεσία Ιστού μπορεί να υπάρχει σε οποιοδήποτε σημείο ενός δικτύου και δεν είναι απαραίτητο να χρησιμοποιείται μόνο μέσω του Διαδικτύου. Η ανάπτυξη

του Διαδικτύου έδωσε τη δυνατότητα στις Υπηρεσίες, οι οποίες προηγουμένως ήταν προσβάσιμες μόνο σε τοπικά δίκτυα, να έχουν πρόσβαση και μέσω Διαδικτύου. Το γεγονός αυτό βελτίωσε τις δυνατότητες και την αξία τους. Επομένως, οι Υπηρεσίες Ιστού είναι μια προσπάθεια να αντιμετωπιστεί το πρόβλημα της επικοινωνίας των καταναμημένων συστημάτων, διαμέσου του εντοπισμού και της πρόσβασης σε απομακρυσμένα συστήματα. (Στεφανίδης Β., 2009)

Θεωρείται αναγκαίο να επισημανθεί ότι οι Υπηρεσίες Ιστού δεν έχουν εξελιχθεί με βάση κάποια προκαθορισμένη αρχιτεκτονική, αλλά επιχειρώντας να αντιμετωπίσουν το ευρύτερο πρόβλημα της ολοκλήρωσης εφαρμογών. Απεναντίας, τα παραδοσιακά καταναμημένα συστήματα έχουν εξελιχθεί με βάση συγκεκριμένες τεχνικές αρχιτεκτονικές. Συνδυαστικά με τις Υπηρεσίες Ιστού είναι εφικτό να πραγματοποιηθούν εργασίες, οι οποίες είναι εκτός του αντικειμένου των παραδοσιακών καταναμημένων συστημάτων, όπως :

- Υποστήριξη μετάδοσης μηνυμάτων με βάση τα έγγραφα.
- Υποστήριξη μετάδοσης μηνυμάτων μέσω Απομακρυσμένης Κλήσης Διαδικασίας (Remote Procedure Call – RPC).
- Μεταφορά κωδικοποιημένων δεδομένων από τις εφαρμογές.
- Λειτουργία πάνω από πρωτόκολλα Ιστού όπως το HTTP και το SMTP.

Η αρχιτεκτονική των Υπηρεσιών Ιστού επικεντρώνεται στην αλληλεπίδραση τριών ρόλων : του Παροχέα (Provider), του Ευρετηρίου (Registry) και του Αιτούντος (Requestor) της υπηρεσίας. Με τις λειτουργίες της Δημοσίευσης (Publish), της Εύρεσης (Find) και της Σύνδεσης (Bind) υλοποιείται η αλληλεπίδραση.



Εικόνα 14 Ρόλοι και Λειτουργίες Υπηρεσιών Ιστού

5. ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΛΕΤΗΣ

Στην παρούσα πτυχιακή μελέτη παρουσιάστηκαν οι διάφορες προσεγγίσεις σχετικά με τον σημασιολογικό ιστό ενώ δόθηκε μεγαλύτερο βάρος στη προσέγγιση των οντολογιών και συγκεκριμένα στην γλώσσα αναπαράστασης OWL.

Αυτό που επιχειρείται να υλοποιηθεί με το Semantic Web είναι η δημιουργία μια στέρεας βάσης συσχετισμών μεταξύ των δεδομένων που αντιλαμβάνονται οι μηχανές. Μόλις αυτός ο στόχος επιτευχθεί και καταφέρουμε υμε να φτιάξουμε έναν τεράστιο όγκο δεδομένων, συνδυασμένων με έναν τρόπο ακριβή και αξιόπιστο – και θέλουμε τα δεδομένα αυτά να είναι οργανωμένα με έναν τέτοιο τρόπο -, θα έχουμε μια πολύ στέρεα βάση. Από εκεί και πέρα, το ενδιαφέρον και η μεγάλη πρόκληση είναι να διαμορφωθούν οι ευρετικές μέθοδοι που θα επιτρέψουν να κάνουμε δημιουργικές αναζητήσεις μέσα στα δεδομένα αυτά. Ωστόσο, δεν είναι αυτό που προσπαθούμε να κάνουμε αυτή τη στιγμή. Πρώτο μέλημά, λοιπόν, είναι η δημιουργία των σχέσεων μεταξύ των δεδομένων.

Στην συνέχεια τον ρόλο του διαχειριστή – αξιολογητή των πληροφοριών αναλαμβάνουν οι πράκτορες. Στον Σημασιολογικό Ιστό οι λειτουργίες που θα εκτελούν οι προσωπικοί πράκτορες διαδοχικά συνοψίζονται ως εξής: Οι πράκτορες πρώτα θα δέχονται κάποια αιτήματα εξυπηρέτησης από τους χρήστες, έπειτα θα αναζητούν πληροφορίες από τις πηγές του Διαδικτύου, θα επικοινωνούν με άλλους πράκτορες, θα συγκρίνουν τις πληροφορίες σχετικά με τα αιτήματα και τις προτιμήσεις των χρηστών, θα επιλέγουν μεταξύ των εναλλακτικών λύσεων και τέλος θα δίνουν τη λύση στον χρήστη ο οποίος θα λαμβάνει και την τελική απόφαση.

Ο ρόλος των πρακτόρων στο Σημασιολογικό Ιστό δεν θα είναι ούτε η αντικατάσταση της ανθρώπινης βούλησης, ούτε και η λήψη τελικών αποφάσεων. Ο χρήστης θα είναι εκείνος που θα δίνει εντολές και θα παίρνει αποφάσεις.

Για την υλοποίηση του σχεδιασμού απαιτείται μια γλώσσα οντολογίας και συγκεκριμένα η OWL, η πιο πρόσφατη εξέλιξη στις γλώσσες οντολογιών για το (Σημαντικό) Ιστό. Η OWL έχει δεχτεί επιρροές από τις περισσότερες γλώσσες που έχουν μέχρι τώρα αναφερθεί, αλλά ιδιαίτερα η DAML+OIL είναι αυτή που αποτέλεσε το εναρκτήριο σημείο για την ανάπτυξη της γλώσσας. Η OWL, όπως και οι προκάτοχοί της, ακολουθεί το συντακτικό της RDF(S), εμπλουτίζοντας όμως το λεξιλόγιό της για την περιγραφή των κλάσεων και των ιδιοτήτων, υποστηρίζοντας, μεταξύ άλλων, σχέσεις μεταξύ των κλάσεων (π.χ. ότι δύο κλάσεις είναι ξένες), πληθικότητα (περιορισμοί αριθμού), ισότητα, περισσότερα χαρακτηριστικά των ιδιοτήτων (π.χ. συμμετρία) και κλάσεις απαρίθμησης (χρήση ονοματικών). Η μεγαλύτερη εκφραστικότητα που προσδίδεται στη γλώσσα με τον τρόπο αυτό, είναι αναγκαία για το Σημαντικό Ιστό, καθώς επιτρέπει τη χρήση συστημάτων συλλογισμού για την διεξαγωγή αποτελεσματικών συμπερασμών πάνω στις οντολογίες που αναπαρίσταται στον Ιστό με τη γλώσσα αυτή.

Στα αποτελέσματα της έρευνας τοποθετείται η διαπίστωση πως ο Σημαντικός Ιστός μπορεί να πραγματοποιήσει ένα αξιόλογο μέρος όσων υπόσχεται, κυρίως μέσω των δυνατοτήτων αναπαράστασης, διαχείρισης και ανακάλυψης γνώσης. Στις δυνατότητες αυτές, σημαντικό ρόλο παίζουν οι οντολογίες και τα σύγχρονα εργαλεία Συλλογισμού (Reasoners).

6. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. **Berners-Lee B** The Semantic Web [Βιβλίο]. - [s.l.] : Scientific Retrieved on 2008, 2008.
2. **Gelfond & Horrocks G** Description Logic Programs [Βιβλίο]. - [s.l.] : World Wide Web Conference, 2003.
3. **Herman H** State of the Semantic Web [Βιβλίο]. - 2007.
4. **Mobasher B. Cooley R., Srivastava J.** Automatic Personalization Based on Web Usage MINING [Βιβλίο]. - [s.l.] : Communications of the ACM, 2000.
5. **Montainer M. Beatriz Lopez, Josef De la Rosa** , A taxonomy of Recommender agents on the internet [Βιβλίο]. - [s.l.] : Artificial Intelligence Review, 2003.
6. **Promoting Urban Sustainable Development in Local Authorities P** Leonardo da Vinci - Βιώσιμες Αστικές Κατασκευές [Βιβλίο]. - [s.l.] : Ajuntament de Sabadell, 2009.
7. **Sven R. Beck M. Freitag B.** , Generating Recommendation Dialogues from Product Models [Βιβλίο]. - Passau : University of Passau - Germany, 2003.
8. **Αγγελόπουλος Α** Σχεδιασμός και Ανάπτυξη Διεπαφής Πελάτη - Εξυπηρετητή για Υποστήριξη Συλλογισμού σε Κατανεμημένες Εφαρμογές του Σημαντικού Ιστού [Βιβλίο]. - Πάτρα : Πανεπιστήμιο Πατρών, 2009.
9. **KAON K** Orthology Manaegement fot the Semantic web [Βιβλίο]. - [s.l.] : <http://kaon2.semanticweb.org.>, 2008.
10. **Κάσσε Κ** Αξιοποίηση Τεχνολογιών Ανοικτού Κώδικα για την ανάπτυξη εφαρμογών [Βιβλίο]. - Πάτρα : Πανεπιστήμιο Πατρών, 2011.

- 11. Κορμπάκης Κ** Πρακτική Εφαρμογή των Οντολογιών ως Εργαλεία Αναπαράστασης και Διαχείρισης Γνώσης στην Ηλεκτρονική Υγεία [Βιβλίο]. - Μακεδονία : Πανεπιστήμιο Δυτικής Μακεδονίας , 2009.
- 12. Κουτσομητρόπουλος Κ** Αξιοποίηση Οντολογιών και Χρήση Τεχνολογιών του Σημαντικού Ιστού για Αναπαράσταση, Διαχείριση και Ανακάλυψη Γνώσης μέσω Διαδικτύου [Βιβλίο]. - Πάτρα : Πανεπιστήμιο Πατρών, 2004.
- 13. Παπαταξιάρχης Π** Τεχνολογίες Αναπαράστασης Γνώσης και Συμπερασμού : Η περίπτωση ενός Έξυπνου Βοηθού Προπονητή Ποδοσφαίρου [Βιβλίο]. - Αθήνα : ΕΚΠΑ, 2008.
- 14. Πετράκης Π** Σημασιολογικά Κατευθυνόμενη Εξόρυξη Γνώσης με Χρήση Οντολογιών [Βιβλίο]. - Πειραιάς : Πανεπιστήμιο Πειραιά, 2011.
- 15. Πρέντζα Π** Σημασιολογικός Ιστός - XML [Βιβλίο]. - Πανεπιστήμιο Πειραιώς : [s.n.], 2010.
- 16. Στασινόπουλος Σ** Διαδίκτυο & Εφαρμογές [Βιβλίο]. - Αθήνα : Εθνικό Μετσόβιο Πολυτεχνείο, 2008.
- 17. Στεφανίδης Β. Σ** Ανάπτυξη Μοντέλου και Πρωτοτύπου Συστήματος Context Aware E learning [Βιβλίο]. - Πάτρα : Ε.Α.Π., 2009.
- 18. Στοϊλος Σ** Γλώσσες Αναπαράστασης Γνώσης στο Σημασιολογικό Ιστό [Βιβλίο]. - Αθήνα : Εθνικό Μετσόβιο Πολυτεχνείο, 2008.