

ΑΤΕΙ ΠΑΤΡΩΝ

ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ

ΤΜΗΜΑ ΕΠΙΧΕΙΡΗΜΑΤΙΚΟΥ ΣΧΕΔΙΑΣΜΟΥ

&

ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΘΕΜΑ :

**«ΟΝΤΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΓΙΑ ΤΗΝ ΑΝΑΠΤΥΞΗ
ΣΥΣΤΗΜΑΤΩΝ»**

ΕΙΣΗΓΗΤΗΣ :

Κος ΠΑΠΑΪΩΑΝΝΟΥ ΒΑΪΟΣ

ΕΠΙΜΕΛΕΙΑ :

ΚΩΤΣΙΑ ΠΑΡΑΣΚΕΥΗ

ΠΑΤΡΑ 2006

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ.....	
ΕΙΣΑΓΩΓΗ.....	1
ΚΕΦΑΛΑΙΟ 1^ο	
A. ΘΕΩΡΙΑ ΟΝΤΟΛΟΓΙΩΝ.....	3
A.1. ΠΕΡΙΓΡΑΦΗ ΑΝΑΠΤΥΞΗΣ ΕΝΟΣ ΠΛΗΡΟΦΟΡΙΑΚΟΥ ΣΥΣΤΗΜΑΤΟΣ.....	3
A.2. ΟΡΙΣΜΟΙ ΟΝΤΟΛΟΓΙΑΣ	4
A.3. ΟΝΤΟΛΟΓΙΑ (ΕΠΙΣΤΗΜΗ ΥΠΟΛΟΓΙΣΤΩΝ).....	6
A.3.1. ΕΥΦΥΗΣ ΠΡΟΣΒΑΣΗ ΣΤΗΝ ΨΗΦΙΑΚΗ ΤΕΧΝΟΛΟΓΙΑ.....	8
A.4. ΤΥΠΟΙ ΟΝΤΟΛΟΓΙΑΣ.....	13
A.5. ΔΙΑΔΙΚΑΣΙΕΣ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΟΛΟΚΛΗΡΩΣΗΣ.....	14
A.6. ΕΞΑΓΩΓΗ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΠΛΗΡΟΦΟΡΙΑΣ.....	17
A.7. ΣΥΓΚΡΙΣΗ ΟΝΤΟΛΟΓΙΩΝ.....	18
A.8. ΟΛΟΚΛΗΡΩΣΗ ΟΝΤΟΛΟΓΙΩΝ.....	19
A.9. ΟΝΤΟΛΟΓΙΑ ΓΙΑ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΣΥΜΒΑΤΟΤΗΤΑ.....	21
A.10 ΟΝΤΟΛΟΓΙΕΣ ΣΥΣΤΗΜΑΤΩΝ.....	22
A.11. . ΣΥΜΠΕΡΑΣΜΑΤΑ ΟΝΤΟΛΟΓΙΩΝ ΣΥΣΤΗΜΑΤΩΝ.....	23
A.12. ΜΕΘΟΔΟΛΟΓΙΕΣ ΑΝΑΠΤΥΞΗΣ ΣΥΣΤΗΜΑΤΩΝ.....	24
A.13. ΧΡΗΣΗ ΣΥΣΤΗΜΑΤΩΝ ΠΟΛΛΑΠΛΩΝ ΠΡΑΚΤΟΡΩΝ.....	27
A.14. ΟΝΤΟΛΟΓΙΚΗ ΚΑΙ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ.....	29
A.15. ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ ΙΣΤΟΣ.....	30
A.16. ΜΕΤΑΔΕΔΟΜΕΝΑ ΚΑΙ ΣΗΜΑΣΙΟΛΟΓΙΚΟΣ ΧΑΡΑΚΤΗΡΙΣΜΟΣ.....	31
ΚΕΦΑΛΑΙΟ 2^ο	
B. ΓΛΩΣΣΑ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ OWL.....	34
B.1. OWL: ΓΛΩΣΣΑ ΟΝΤΟΛΟΓΙΑΣ ΤΟΥ ΠΑΓΚΟΣΜΙΟΥ ΙΣΤΟΥ.....	34
B.2. Η OWL ΑΠΟΦΕΡΕΙ ΟΝΤΟΛΟΓΙΕΣ ΠΟΥ ΔΟΥΛΕΥΟΥΝ ΣΤΟΝ ΠΑΓΚΟΣΜΙΟ ΙΣΤΟ.....	34
B.3. ΤΑ ΎΓΓΡΑΦΑ OWL ΠΟΥ ΠΑΡΑΓΟΝΤΑΙ ΑΠΟ ΤΟ W3C.....	35

B.4. Η ΘΕΣΗ ΤΗΣ OWL ΣΤΗΝ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ΣΗΜΑΣΙΟΛΟΓΙΚΟΥ ΙΣΤΟΥ :XML, RDF	36
B.5. ΣΗΜΑΣΙΟΛΟΓΙΚΕΣ ΚΑΙ ΟΝΤΟΛΟΓΙΚΕΣ ΓΛΩΣΣΕΣ.....	37
B.6. ΜΕΘΟΔΟΛΟΓΙΕΣ ΜΕΤΑΦΡΑΣΗΣ ΚΑΙ ΟΛΟΚΛΗΡΩΣΗΣ ΣΧΗΜΑΤΩΝ.....	40
B.7. ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΚΟΙΝΟΠΡΑΞΙΑ ΠΑΓΚΟΣΜΙΟΥ ΙΣΤΟΥ [W3C].....	43
B.8. ΒΙΟΜΗΧΑΝΙΚΟΙ ΚΑΙ ΑΚΑΔΗΜΑΪΚΟΙ ΗΓΕΤΕΣ ΠΡΟΩΘΟΥΝ ΤΗΝ OWL ΜΠΡΟΣΤΑ.....	43
B.9. ΠΕΡΙΠΤΩΣΕΙΣ ΕΜΠΛΟΥΤΙΣΜΟΥ ΚΑΙ ΣΥΓΧΩΝΕΥΣΗΣ ΒΑΣΕΩΝ ΓΝΩΣΗΣ.	44
ΚΕΦΑΛΑΙΟ 3^ο	
Γ. ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ ΑΝΑΠΤΥΞΗΣ ΣΥΣΤΗΜΑΤΟΣ ΑΕΡΟΠΟΡΙΚΗΣ ΕΤΑΙΡΙΑΣ.....	54
Γ.1. ΕΦΑΡΜΟΓΗ ΓΙΑ ΤΗΝ ΓΛΩΣΣΑ OWL (ΑΝΑΛΥΣΗ ΠΑΡΑΔΕΙΓΜΑΤΟΣ ΜΙΑΣ ΑΕΡΟΠΟΡΙΚΗΣ ΕΤΑΙΡΙΑΣ).....	54
Γ.2. ΠΑΡΑΔΕΙΓΜΑ ΑΕΡΟΠΟΡΙΚΗΣ ΕΤΑΙΡΙΑΣ ΣΕ ER ΔΙΑΓΡΑΜΜΑ.....	55
Γ.3. ΑΝΑΠΤΥΞΗ ΟΝΤΟΛΟΓΙΑΣ ΓΙΑ ΤΗΝ ΑΕΡΟΠΟΡΙΚΗ ΕΤΑΙΡΙΑ	58
Γ.3.1.ΠΕΡΙΓΡΑΦΗ ΤΗΣ ΑΕΡΟΠΟΡΙΚΗΣ ΕΤΑΙΡΙΑΣ ΣΕ OWL	61
ΚΕΦΑΛΑΙΟ 4^ο	
ΣΥΜΠΕΡΑΣΜΑΤΑ.....	70
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	

ΠΡΟΛΟΓΟΣ

Στην επιστήμη των υπολογιστών μία οντολογία είναι ένα μοντέλο δεδομένων που αντιπροσωπεύει ένα πεδίο και χρησιμοποιείται για να βγάξει λογικά συμπεράσματα σχετικά με τα αντικείμενα σε αυτό το πεδίο και τις σχέσεις μεταξύ των. Οι οντολογίες χρησιμοποιούνται στην τεχνητή νοημοσύνη, στον σημασιολογικό ιστό, στο σχεδιασμό λογισμικού, και στην αρχιτεκτονική πληροφοριών ως μία μορφή παρουσίασης γνώσεων σχετικά με τον κόσμο ή τμήμα του.

Επίσης μία οντολογία είναι προϊόν μιας απόπειρας να διατυπωθεί ένα εξονυχιστικό και αυστηρό εννοιολογικό σχήμα γύρω από ένα πεδίο. Μία οντολογία τυπικά είναι μία ιεραρχική δομή δεδομένων που περιέχει όλες τις σχετικές οντότητες και τις σχέσεις τους και τους κανόνες που υπάρχουν σε αυτό το πεδίο.

ΕΙΣΑΓΩΓΗ

Οι οντολογίες αποτελούν από τις σημαντικότερες ερευνητικές κατευθύνσεις για την επίτευξη σημασιολογικής διαλειτουργικότητας. Παρέχουν δυνατότητα συστηματικότερης κατανόησης ενός πληροφοριακού συστήματος. Επιπλέον αποτελούν την μόνη διέξοδο για την αποτελεσματική ολοκλήρωση συλλογών δεδομένων.

Στην συγκεκριμένη εργασία γίνεται αναφορά για την οντολογία ενός πληροφοριακού συστήματος. Στην αρχή δείχνουμε διαφόρους ορισμούς της οντολογίας από επιστήμονες που ασχολήθηκαν με το συγκεκριμένο αντικείμενο. Επίσης γίνεται εκτενέστερη αναφορά στον Σημασιολογικό Ιστό και την σημασία που έχει, καθώς τα τελευταία δυο χρόνια γίνεται πολύς λόγος. Ο Σημασιολογικός Ιστός αποτελεί πρωτοβουλία της Κοινοπραξίας του Παγκόσμιου Ιστού (W3C), η οποία δημιουργήθηκε για να οδηγήσει τον Ιστό στην πλήρη αξιοποίηση του αναπτύσσοντας κοινά πρωτόκολλα που προάγουν την εξέλιξη του και εξασφαλίζουν τη δια-λειτουργικότητα του.

Επίσης γίνεται αναφορά στην χρήση μιας γλώσσας για την οντολογική προσέγγιση ενός συστήματος, της OWL. Η OWL παρέχει μια γλώσσα για τον ορισμό δομημένων οντολογιών που βασίζονται στον Παγκόσμιο Ιστό, που αποφέρει πλουσιότερη ενσωμάτωση και δια-λειτουργικότητα δεδομένων ανάμεσα σε περιγραφικές κοινότητες. Επίσης χτίζει πάνω στο Μοντέλο και Σχήμα RDF, προσθέτοντας πλουσιότερο λεξιλόγιο για την περιγραφή ιδιοτήτων και τάξεων.

Τέλος σ' αυτήν την εργασία θα εξηγήσουμε πως συγκεκριμένα πρότυπα μπορούν να εκφραστούν μέσα από την γλώσσα προγραμματισμού OWL χρησιμοποιώντας το πρόγραμμα Protégé. Για αυτό στο τελευταίο μας κεφάλαιο δείχνουμε μέσω μιας εφαρμογής την χρήση της γλώσσας OWL στην ανάπτυξη ενός συστήματος. Επίσης θα γίνει πρώτα η ανάλυση με την χρήση του ER διαγράμματος και μετά την χρήση της γλώσσας OWL.

Αναλυτικά στο πρώτο κεφάλαιο περιγράφουμε ένα πληροφοριακό σύστημα καθώς και τα χαρακτηριστικά των οντολογιών. Επίσης αναφέρουμε την σημασία της οντολογίας στην επιστήμη των υπολογιστών, τα πλεονεκτήματα και τις δυνατότητες που προσφέρει. Στη συνέχεια

παρουσιάζουμε τις διαδικασίες για τη σημασιολογική ολοκλήρωση των οντολογιών καθώς κάνουμε σύγκριση των οντολογιών μεταξύ τους. Σ' αυτό το κεφάλαιο γίνεται αναφορά για τις οντολογίες συστημάτων και τις μεθοδολογίες ανάπτυξης τους. Παράλληλα περιγράφουμε την χρήση συστημάτων πολλαπλών πρακτόρων για την αυτόματη και έξυπνη διαχείριση πολύπλοκων διεργασιών στο διαδίκτυο καθώς και σε μικρότερα δίκτυα. Τέλος γίνεται αναλυτική περιγραφή του Σημασιολογικού Ιστού (Semantic Web), για την ραγδαία εξέλιξή του τα τελευταία χρόνια.

Στο δεύτερο κεφάλαιο γίνεται αναλυτική περιγραφή της γλώσσας προγραμματισμού OWL, και την σχέση που έχει με τον Σημασιολογικό Ιστό. Επίσης κάνουμε αναφορά και σε άλλες γλώσσες προγραμματισμού καθώς και τις σχέσεις ή διαφορές τους με την OWL. Τέλος σ' αυτό το κεφάλαιο δείχνουμε κάποιες περιπτώσεις εμπλουτισμού και συγχώνευσης βάσεων γνώσης.

Στο τρίτο κεφάλαιο κάνουμε μία εφαρμογή πάνω στα παραπάνω που έχουμε αναπτύξει με την ανάλυση ενός προβλήματος σε μια αεροπορική εταιρία. Πρώτα περιγράφουμε το πρόβλημα με ER διάγραμμα και μετά το απεικονίζουμε σε OWL γλώσσα.

Τέλος στο τέταρτο κεφάλαιο αναλύουμε όσα έχουμε περιγράψει στα προηγούμενα κεφάλαια καθώς και τις αντιλήψεις μας, που έχουμε αποκομίσει από αυτήν την εργασία.

ΚΕΦΑΛΑΙΟ 1

ΘΕΩΡΙΑ ΟΝΤΟΛΟΓΙΩΝ

1.1 Περιγραφή ανάπτυξης ενός Πληροφοριακού Συστήματος

Τα Πληροφοριακά Συστήματα (ΠΣ) αποτελούν ένα από τα σημαντικότερα στοιχεία της σημερινής επιχείρησης. Είναι πολύ δύσκολο στη σημερινή εποχή να φανταστούμε μια επιχείρηση ή έναν οργανισμό να λειτουργεί χωρίς την υποστήριξη πληροφοριακών συστημάτων, δηλαδή συστημάτων τα οποία δέχονται, αποθηκεύουν, επεξεργάζονται και αναλύουν δεδομένα με τη βοήθεια υπολογιστών.

Πληροφοριακό Σύστημα είναι ένα σύστημα το οποίο αξιοποιεί διάφορες πηγές πληροφόρησης και διαχειρίζεται τις πληροφορίες που συγκεντρώνει με κύρια συστατικά τους ανθρώπους, το υλικό, το λογισμικό και τις διαδικασίες. Ένα ΠΣ αποτελείται από στοιχεία που συνδέονται μεταξύ τους με σχέσεις αλληλεξάρτησης και συνλειτουργούν με στόχο την επίτευξη συγκεκριμένου σκοπού. Η απόδοσή του εξαρτάται από τους πόρους που έχει στην διάθεσή του, οι οποίοι μπορεί να είναι οι πληροφορίες, τα δεδομένα και τα διατεθειμένα στο σύστημα μέσα.

Το ΠΣ περιλαμβάνει έναν αριθμό υποσυστημάτων. Κάθε υποσύστημα μέσω μηχανισμών επεξεργασίας πληροφοριών που διαθέτει, μετατρέπει τις εισερχόμενες πληροφορίες και τις δίνει επεξεργασμένες σε άλλο υποσύστημα. Επίσης κάθε υποσύστημα καλύπτει ένα κύκλο μηχανογραφικών εφαρμογών.

Για την ανάπτυξη ενός ΠΣ χρειάζεται να γίνουν κάποιες ενέργειες. Οπότε οι φάσεις ανάπτυξης ενός ΠΣ είναι οι εξής:

- Διενέργεια προκαταρκτικής έρευνας : έχει σκοπό να εξακριβώσει την αναγκαιότητα δημιουργίας νέου ΠΣ.
- Σύσταση μελέτης εφικτότητας : εξετάζει την δυνατότητα δημιουργίας του ΠΣ και διερευνά εάν η σχετική επένδυση για την ανάπτυξη του αντισταθμίζεται από τα προβλεπόμενα οφέλη.
- Ανάλυση υπάρχοντος ΠΣ: λεπτομερή καταγραφή και ενημέρωση για τις λειτουργίες του οργανισμού ή της επιχείρησης. Στόχος του είναι να συλλέξει στοιχεία για την διακίνηση εντύπων που διακινούνται, καταγραφή των λειτουργιών και χωροταξική καταγραφή των θέσεων εργασίας.
- Σχεδιασμός του νέου ΠΣ : καταγράφονται οι λειτουργίες με το νέο ΠΣ και δίνονται λύσεις των προβλημάτων του παλαιού ΠΣ.

Έτσι, μ' αυτόν τον τρόπο δημιουργείται ένα νέο Πληροφοριακό Σύστημα, που σκοπό έχει να διευκολύνει στην διεξαγωγή της ομαλής λειτουργίας ενός οργανισμού ή μιας επιχείρησης.

1.2. Ορισμοί οντολογίας

Σημαντικό ρόλο στην δημιουργία ενός Πληροφοριακού Συστήματος παίζει η οντολογική ανάπτυξη του συστήματος. Έτσι, είναι μία συστηματική κατάταξη όλων των σημαντικών κατηγοριών αντικειμένων ή εννοιών που υπάρχουν σε κάποιο γνωστικό πεδίο που δείχνει τις σχέσεις μεταξύ των οντολογιών. Όταν είναι ολοκληρωμένη, μία οντολογία είναι μία κατηγοριοποίηση όλων των εννοιών σε κάποιο πεδίο γνώσης, και περιλαμβάνει τα αντικείμενα και όλες τις ιδιότητες, σχέσεις και τις λειτουργίες που χρειάζονται για να οριστούν τα

αντικείμενα και να καθοριστούν οι ενέργειές τους. Μία απλουστευμένη οντολογία μπορεί να περιέχει μόνο μία ιεραρχική ταξινόμηση (μία ταξινομία) που να δείχνει το είδος των σχέσεων υπαγωγής μεταξύ των εννοιών στο γνωστικό πεδίο. Μπορούμε να οπτικοποιήσουμε μία οντολογία ως ένα αφηρημένο γράφημα που έχει σημεία τομής και χαρακτηρισμένες καμπύλες που αντιπροσωπεύουν τα αντικείμενα και τις σχέσεις των. Θα δείξουμε πιο κάτω διάφορους ορισμούς οντολογίας.

Σημείωση: Οι έννοιες που περιλαμβάνονται σε μία οντολογία και η ιεραρχική κατάταξη είναι μέχρι ενός σημείου αυθαίρετες, ανάλογα με το σκοπό για τον οποίο δημιουργήθηκε η οντολογία. Αυτό πηγάζει από το γεγονός ότι τα αντικείμενα έχουν διάφορες ιδιότητες των αντικειμένων μπορούν να επιλεχθούν ως τα κριτήρια βάσει των οποίων τα αντικείμενα θα ταξινομηθούν. Επιπρόσθετα μπορεί να χρησιμοποιηθεί η συνάθροιση των εννοιών σε διαφορετικό βαθμό κατά προσέγγιση, και οι διαχωρισμοί των μπορεί να είναι σημαντικοί για κάποιο σκοπό αλλά μηδαμινής σημασίας για κάποιον άλλο. (Από Webster's dictionary)

Η δομή ενός συστήματος. Ένα μοντέλο συστήματος. Η λέξη οντολογία κυριολεκτικά σημαίνει τον κλάδο της μεταφυσικής που ασχολείται με τη φύση της πραγματικότητας ή της ύπαρξης. Επομένως είναι μία κυριλέ λέξη για το "τι υπάρχει" σε ένα σύστημα. (Tech Encyclopedia)

Οντολογία είναι η επιστήμη για το τι είναι, από ποια είδη αποτελείται και τις δομές των αντικειμένων, τις ιδιότητες, τα γεγονότα, τις διαδικασίες και τις σχέσεις σε κάθε περιοχή της πραγματικότητας. (Smith)

Για τα πληροφοριακά συστήματα, η οντολογία είναι παρουσίαση μερικών παλαιότερων κατακτήσεων της πραγματικότητας.

1. Αντανακλά τις ιδιότητες των αντικειμένων μέσα από μία περιοχή με τέτοιο τρόπο με τον οποίο τα αποκτά μια συστηματική σχέση μεταξύ πραγματικότητας και παρουσίασης του
2. Είναι κατανοητή σε έναν εμπειρογνώμονα περιοχών
3. Είναι προσδιορισμένο με τέτοιο τρόπο που επιτρέπει να υποστηρίξει αυτοματοποιημένα συστήματα πληροφοριών

Στην επιστήμη των υπολογιστών μία οντολογία είναι ένα μοντέλο δεδομένων που αντιπροσωπεύει ένα πεδίο και χρησιμοποιείται για να βγάξει λογικά συμπεράσματα σχετικά με τα αντικείμενα σε αυτό το πεδίο και τις σχέσεις μεταξύ των. Οι οντολογίες χρησιμοποιούνται στην τεχνητή νοημοσύνη, στον σημασιολογικό ιστό, στο σχεδιασμό λογισμικού, και στην αρχιτεκτονική πληροφοριών ως μία μορφή παρουσίασης γνώσεων σχετικά με τον κόσμο ή τμήμα του.

Οι οντολογίες γενικώς περιγράφουν:

- Μεμονωμένα "άτομα": τα βασικά ή "πρώτου (ground) επιπέδου" αντικείμενα
- Τάξεως : συλλογές ή τύποι αντικειμένων
- Αντιπροσωπευτικές Ιδιότητες : ιδιότητες, χαρακτηριστικά ή παράμετροι που έχουν ή μοιράζονται τα αντικείμενα
- Συσχετισμοί : οι τρόποι με τους οποίους τα αντικείμενα αλληλοσυσχετίζονται

1.3. Οντολογία (επιστήμη υπολογιστών)

Στην επιστήμη των υπολογιστών μία οντολογία είναι προϊόν μιας απόπειρας να διατυπωθεί ένα εξονυχιστικό και αυστηρό εννοιολογικό σχήμα

γύρω από ένα πεδίο. Μία οντολογία τυπικά είναι μία ιεραρχική δομή δεδομένων που περιέχει όλες τις σχετικές οντότητες και τις σχέσεις τους και τους κανόνες που υπάρχουν σε αυτό το πεδίο (π.χ. μία οντολογία πεδίο domain ontology).

Η χρήση του όρου οντολογία στην επιστήμη των υπολογιστών πηγάζει από την παλαιότερη χρήση του όρου οντολογία στην φιλοσοφία. Η φιλοσοφική έννοια του όρου οντολογία είναι η μελέτη της "ύπαρξης". Ο σκοπός μιας υπολογιστικής οντολογίας δεν είναι να καθορίσει "τι υπάρχει" ή "δεν υπάρχει", αλλά είναι να δημιουργήσει μία βάση δεδομένων η οποία να περιέχει έννοιες που αναφέρονται σε οντότητες που παρουσιάζουν ενδιαφέρον για τον οντολόγο, και οι οποίες θα είναι χρήσιμες στην εκτέλεση συγκεκριμένων μορφών υπολογισμών. Γι' αυτό το λόγο, η λογική που χρησιμοποιείται από τους φιλόσοφους οντολόγους μπορεί να μας βοηθήσει να αναγνωρίσουμε και να αποφύγουμε πιθανές λογικές ασάφειες.

Όπου εναλλακτικές αναπαραστάσεις εξυπηρετούν το ίδιο καλά τον σκοπό του υπολογιστικού οντολόγου, συνήθως οι χρονικοί περιορισμοί των υποδεικνύουν μία επιλογή που πρέπει να γίνει και ποιες θα αγνοηθούν. Για κάποιους σκοπούς, μπορεί να είναι καλύτερο να αγνοήσουμε πολλές λεπτομέρειες των αντικειμένων ενδιαφέροντος μας. Ως αποτέλεσμα υπολογιστικές οντολογίες που αναπτύχθηκαν ανεξάρτητα για διαφορετικούς σκοπούς συχνά διαφέρουν πολύ η μία από την άλλη. Οι διαφορετικές οντολογίες του ίδιου πεδίου αντίληψης μπορούν επίσης να οφείλονται σε διαφορετικές αντιλήψεις του πεδίου εξαιτίας του κοινωνικού υπόβαθρου, μορφωτικού επιπέδου κ.α.

Μία οντολογία η οποία δεν συνδέεται με ένα πεδίο που αναφέρεται σε ένα συγκεκριμένο πρόβλημα-αντικείμενο αλλά επιχειρεί να περιγράψει γενικές οντότητες είναι γνωστή ως **θεμέλια οντολογία** ή **ανώτερη οντολογία**. Τυπικά πρέπει να δημιουργηθούν ποιο ειδικευμένα σχήματα που αναφέρονται σε

συγκεκριμένα πεδία (κυριαρχίας) ώστε να μπορούν να χρησιμοποιηθούν τα δεδομένα (data) για τη λήψη αποφάσεων στον πραγματικό κόσμο.

1.3.1 Ευφυής πρόσβαση στην ψηφιακή τεχνολογία

Τα τελευταία χρόνια παρατηρείται μία αλματώδη αύξηση στον ρυθμό παραγωγής πληροφοριών, αποτέλεσμα των εξελίξεων σε διάφορα επιστημονικά πεδία αλλά και των απαιτήσεων των σύγχρονων εταιριών για αυξημένη παραγωγικότητα και αποτελεσματικότητα. Αποτέλεσμα αυτής της πληροφοριακής έκρηξης είναι το γεγονός ότι η εξαγωγή γνώσης από την πληροφορία είναι εξίσου βαρυσήμαντη όσο και η ίδια η πληροφορία. Ιδιαίτερα στον χώρο των επιχειρήσεων, η εκμετάλλευση της γνώσης αποτελεί κρίσιμο παράγοντα για την αύξηση της ανταγωνιστικότητας και για την ικανοποιητική υποστήριξη διαδικασιών λήψης αποφάσεων. Ως γνώση θα μπορούσε να χαρακτηριστεί κάθε επιπλέον πληροφορία (μετα-δεδομένα) που περιγράφει την πληροφορία και τις σχέσεις που υπάρχουν μεταξύ των δεδομένων. Η ερευνητική κοινότητα έχει επενδύσει σημαντικά στον τομέα της διαχείρισης γνώσης μέσω της εκμετάλλευσης των μετα-δεδομένων και την ανάπτυξη δομών, των οντολογιών, που περιγράφουν αποδοτικά αντικείμενα, ιδέες, διαδικασίες καθώς και των σχέσεων που τις διέπουν.

Η οντολογική προσέγγιση για την υποστήριξη πληροφοριακών συστημάτων έχει ήδη παρουσιάσει ενθαρρυντικά αποτελέσματα μέσω της κωδικοποίησης και εκμετάλλευσης της σημασιολογίας της πληροφορίας. Τα σημαντικότερα προβλήματα που παρουσιάζονται σε αυτήν την προσέγγιση αφορούν την δυσκολία της διαχείρισης των μετα-δεδομένων και των οντολογιών. Ένα άλλο σημαντικό πρόβλημα είναι η απουσία ενιαίας

σημασιολογικής κωδικοποίησης αφού κάθε πεδίο (ή κάθε οργανισμός) αντιλαμβάνεται με διαφορετικό τρόπο την πληροφορία.

Οι περισσότερες προσεγγίσεις στην κατασκευή πληροφοριακών συστημάτων δεν έχουν επικεντρώσει την προσοχή τους σε παράγοντες ποιότητας στο θέμα της ζήτησης /ανάκτησης πληροφορίας, κυρίως γιατί τα κριτήρια απόδοσης αποτελούσαν το κύριο μέλημα των σχεδιαστών. Τα τελευταία χρόνια όμως η κατάσταση έχει αλλάξει, με αποτέλεσμα παράγοντες όπως η σχετικότητα, η ακρίβεια, η πληρότητα και η συνοχή των αποτελεσμάτων αναζήτησης να θεωρούνται εξίσου σπουδαίοι με τους χρόνους απόκρισης. Μία πρώτη προσέγγιση για την λύση του προβλήματος της ποιοτικής αναζήτησης / ανάκτησης πληροφορίας αποτελεί η ανάκτηση πληροφορίας με βάση το περιεχόμενο (content based retrieval). Η προσέγγιση αυτή στηρίζεται στη ανάλυση του περιεχομένου της πληροφορίας μέσω στατικών ή και υπολογιστικών μεθόδων για την επίτευξη όσο το δυνατόν καλύτερων αποτελεσμάτων. Παρόλη την πρόοδο που έχει επιτευχθεί στον τομέα της αναζήτησης /ανάκτησης πληροφορίας με βάση το περιεχόμενο, τα αποτελέσματα έχουν δείξει ότι υπάρχουν σημαντικοί περιορισμοί σε αυτήν την προσέγγιση. Αυτό οφείλεται στο γεγονός ότι η συνολική πληροφορία δεν περιορίζεται μόνο στο περιεχόμενο ή στους χρήστες αλλά αποτελείται και από διαδικασίες και σχέσεις που διέπουν την ίδια την πληροφορία. Τα σύγχρονα πληροφοριακά συστήματα, και ιδιαίτερα τα εταιρικά, βασίζονται στο μοντέλο ροής πληροφοριών "pull" (pull information model) και κατά συνέπεια οι ανάγκες των χρηστών είναι αυτές οι οποίες θα οδηγήσουν τις μελλοντικές τεχνολογικές εξελίξεις σε αυτόν τον τομέα.

Η ανάγκη για την ικανοποίηση των αναγκών των χρηστών έχει οδηγήσει την ερευνητική κοινότητα στην υιοθέτηση τεχνικών που εκμεταλλεύονται κάθε διαθέσιμη πληροφορία. Μία προσέγγιση η οποία έχει ελκύσει την προσοχή της ερευνητικής κοινότητας είναι ο τομέας της διαχείρισης γνώσης μέσω της

εκμετάλλευσης των μετα-δεδομένων και την ανάπτυξη δομών, των οντολογιών, που περιγράφουν αποδοτικά αντικείμενα, ιδέες, διαδικασίες καθώς και των σχέσεων που τις διέπουν.

Η οντολογία είναι ένας απαραίτητος παράγοντας για την μοντελοποίηση και την διαχείριση της γνώσης. Στην περίπτωση των πληροφοριακών συστημάτων, ως οντολογία χαρακτηρίζεται η ενιαία και γενικά αποδεκτή σημασιολογική κωδικοποίηση της πληροφορίας ενός θεματικού χώρου. Η διαδικασία της μετατροπής της πληροφορίας σε σημασιολογική είναι δυνατή μέσω ενός πλαισίου περιγραφής, του **RDF (Resource Description Framework)**. Αυτό το πλαίσιο επιτρέπει την σύνδεση του περιεχομένου με ειδικά λεξιλόγια τα οποία ερμηνεύουν την σημασία του. Αυτά τα λεξιλόγια αποτελούν τις οντολογίες. Ο συνδυασμός του περιεχομένου με τις οντολογίες επιτρέπουν την καλύτερη οργάνωση, ανάκτηση, αναζήτηση και ολοκλήρωση του περιεχομένου και συνεπώς την βελτιωμένη απόδοση των αντίστοιχων υπηρεσιών ή και την δυνατότητα κατασκευής νέων.

Ο συνδυασμός των οντολογιών και του αντίστοιχου περιεχομένου αναφέρονται ως σημασιολογικό περιεχόμενο. Στην περίπτωση της συγκεκριμένης έρευνας, το πεδίο εφαρμογής αφορά περιβάλλοντα Intranet /Extranet στηριγμένα σε τεχνολογίες Web, οπότε και ο αντίστοιχος στόχος είναι η προώθηση του semantic Intranet /Extranet. Η ιδέα αυτή είναι παράλληλη με αυτή του Semantic Web, αν και τα ιδιαίτερα χαρακτηριστικά της την διαφοροποιούν σε πολύ μεγάλο βαθμό.

Η δυναμική που έχει αναπτυχθεί επιβεβαιώνει την σημασία των οντολογιών για τα σύγχρονα πληροφοριακά συστήματα. Ειδικά στο πεδίο της ανάκτησης πληροφορίας, η χρήση όλης της διαθέσιμης πληροφορίας (περιεχόμενο και γνώση) έχει ήδη δώσει πολύ ενθαρρυντικά αποτελέσματα αλλά έχει παρουσιάσει και νέα προβλήματα. Ειδικά στο πεδίο της ανάκτησης

της πληροφορίας σε εφαρμογές **e-Business** έχουν εντοπιστεί τα παρακάτω χαρακτηριστικά:

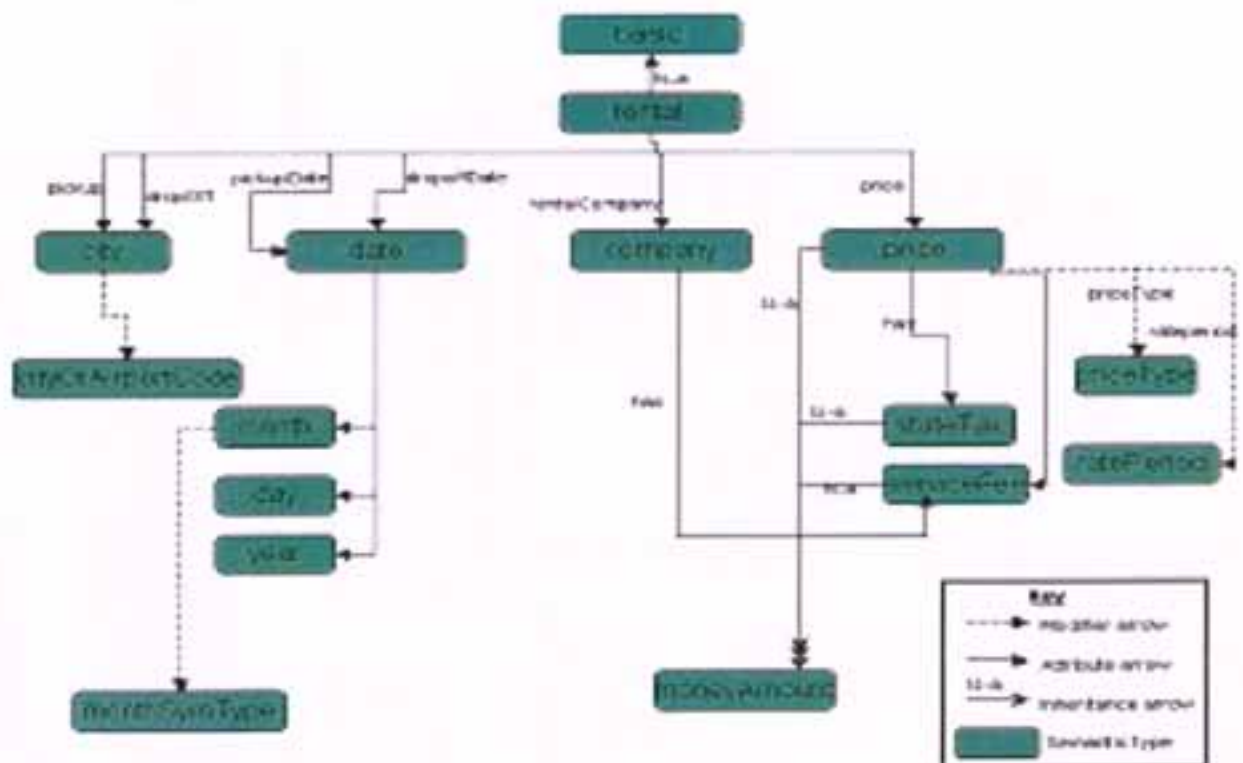
Πλεονεκτήματα /Δυνατότητες

- Αυξημένη ακρίβεια αναζήτησης μέσω της άρσης προβλημάτων όπως διαφορούμενες έννοιες, ανεπαρκής πληροφορία από τον χρήστη κ.λ.π
- Δυνατότητα σύνδεσης σχετικών πληροφοριών
- Μείωση του θορύβου της πληροφορίας
- Διευκόλυνση διαδικασιών επίλυσης προβλημάτων ή /και λήψης προβλημάτων
- Συμμετοχή και επαναχρησιμοποίηση γνώσης
- Αυτοματοποίηση
- Προσαρμοζόμενες υπηρεσίες
- Ανάπτυξη αποδοτικότερων εργαλείων εξαγωγής / διαχείρισης γνώσης
- Εκμετάλλευση πολυμεσικών δεδομένων

Ανοικτά προβλήματα

- Δυσκολία στην παραγωγή μετα-δεδομένων και οντολογιών από τους χρήστες
- Ο ανακριβής ορισμός των οντολογιών οδηγεί σε προβλήματα ολοκλήρωσης
- Ο μη ανοικτός και ευέλικτος ορισμός των οντολογιών οδηγεί σε προβλήματα ευχρηστίας

- Η μη σωστή κατανόηση του επιχειρηματικού μοντέλου και η λανθασμένη ολοκλήρωση των υπηρεσιών που βασίζονται σε οντολογίες οδηγούν σε αδυναμία εκμετάλλευσης των δυνατοτήτων τους
- Ανομοιογένεια των υπαρχόντων σημασιολογικών μοντέλων (ontology interoperability)
- Είναι μη εφικτή (προς το παρόν) η εφαρμογή ενός συνολικού οντολογικού μοντέλου για όλες τις εφαρμογές ηλεκτρονικού εμπορίου (δηλαδή e-Business, e-Commerce, e-CRM κτλ).



1.4. Τύποι οντολογίας

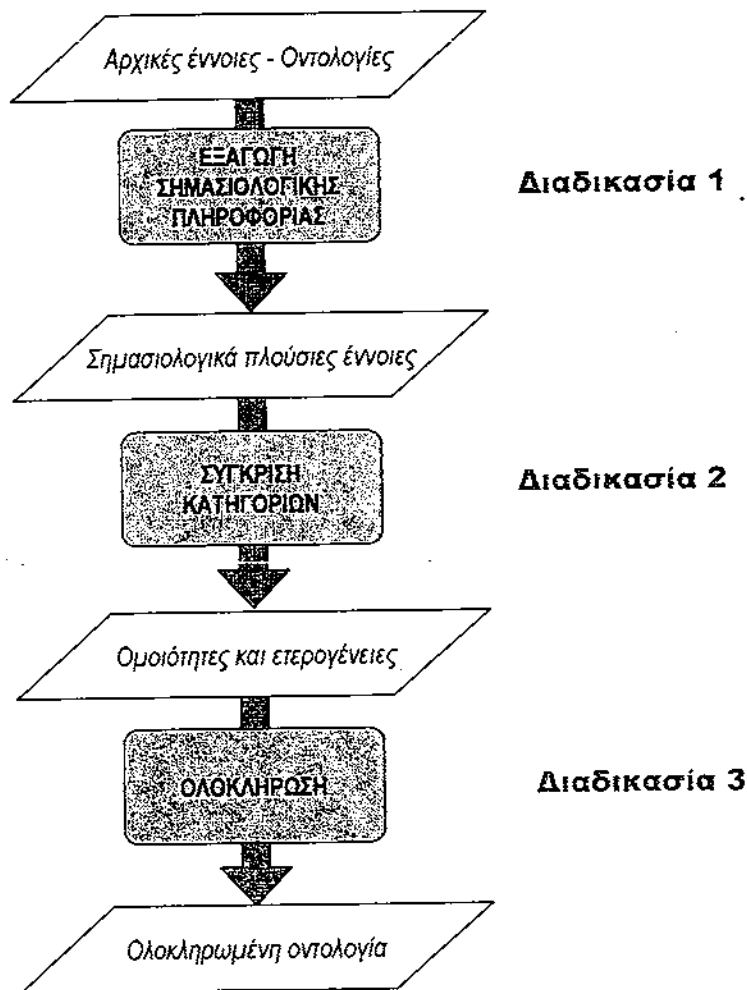
Ένα πεδίο οντολογίας είναι μία οντολογία που συνδέεται με ένα συγκεκριμένο πεδίο. Μία θεμέλια οντολογία είναι μία μορφή οντολογίας που προσπαθεί να είναι λιγότερο συγκεκριμένη και μ' αυτό τον τρόπο να είναι πιο γενικός εφαρμόσιμη. Περιέχει ένα βασικό γλωσσάριο (core glossary), του οποίου οι όροι χρησιμοποιούνται για την περιγραφή των πάντων σε ένα ευρύ πεδίο. Μία θεμέλια οντολογία στην επιστήμη των υπολογιστών θα χρησιμοποιούνταν ως βασική οντολογία (core ontology) και για τα προγράμματα και τους χρήστες των υπολογιστών, επηρεάζοντας την αντίληψη που θα έχουν για τα δεδομένα και τα γεγονότα.

Ένα παράδειγμα στην επιστήμη των υπολογιστών είναι ότι όλα τα προγράμματα υπολογιστών από προκαθορισμένη έχουν μία θεμέλια οντολογία που αποτελείται από ένα σετ οδηγιών της κεντρικής μονάδας, μία στάνταρ βιβλιοθήκη σε κάποια γλώσσα προγραμματισμού, αρχεία που βρίσκονται σε προσβάσιμα συστήματα αρχειοθέτησης, ή κάποιο άλλο κατάλογο που περιέχει «αυτά που υπάρχουν». Επειδή αυτά μπορούν να είναι ανεπαρκείς αναπαραστάσεις για κάποια πεδία αντικειμένων, πρέπει να δημιουργηθεί πιο ειδικευμένο σχήμα για να είναι τα δεδομένα χρήσιμα στη λήψη αποφάσεων του πραγματικού κόσμου. Έτσι για να έχουμε την ανάγκη ύπαρξης κάποιων στάνταρς που παίρνουν τις «βασικές» οντολογίες (π.χ. η Dublin Core σε SGML) και τις κάνουν θεμέλιες.

Έτσι ο T.R Gruber και ο R. Studer περιέγραψαν μία οντολογία ως μία ρητή και τυπική ειδίκευση μίας εννοιολογίας.

1.5. Διαδικασίες σημασιολογικής ολοκλήρωσης

Η σημασιολογική ολοκλήρωση των οντολογιών, εκτός από τη διαδικασία ολοκλήρωσης, προϋποθέτει δύο επιπλέον διαδικασίες: τη διαδικασία εξαγωγής σημασιολογικής πληροφορίας για τη δημιουργία ή τον εμπλουτισμό των αρχικών οντολογιών και τη διαδικασία σύγκρισης των οντολογιών – προσδιορισμού των μεταξύ τους ομοιοτήτων και ετερογενειών (Σχήμα 1). Παρόλο που οι διαδικασίες αυτές είναι διακριτές και αντιμετωπίζονται ανεξάρτητα στη βιβλιογραφία, η συνεργασία τους είναι απαραίτητη σε μια μεθοδολογία σημασιολογικής ολοκλήρωσης. Έτσι, τα στοιχεία που εξάγονται από κάθε μια από τις επιμέρους διαδικασίες εισάγονται στην επόμενη διαδικασία.



Σχήμα 1. Διαδικασίες για τη σημασιολογική ολοκλήρωση οντολογιών

Στη συνέχεια παρουσιάζονται έξι χαρακτηριστικά που διαδραματίζουν βασικό ρόλο στην ανάλυση και τυποποίηση των τριών διαδικασιών:

1. υποθέσεις - στόχοι: προκειμένου να είναι κατανοητή κάθε διαδικασία πρέπει να είναι γνωστές οι υποθέσεις που γίνονται σχετικά με τη σημασιολογία των στοιχείων, καθώς και οι στόχοι της.

2. σημασιολογικό επίπεδο: κάθε προσέγγιση έχει διαφορετική αντιμετώπιση της σημασιολογίας των οντολογιών. Άλλες χρησιμοποιούν λιγότερα σημασιολογικά στοιχεία, όπως όρους και χαρακτηριστικά, ενώ άλλες προχωρούν σε μεγαλύτερο σημασιολογικό βάθος αξιοποιώντας στοιχεία όπως μέρη, λειτουργίες, σημασιολογικές ιδιότητες και σχέσεις. Όσο πλουσιότερη σημασιολογία περιλαμβάνει κάθε προσέγγιση, τόσο βαθύτερη «επικοινωνία» επιτυγχάνεται μεταξύ των συσχετιζόμενων οντολογιών.
3. εισαγόμενα-εξαγόμενα στοιχεία: τα στοιχεία που εισάγονται και εξάγονται από κάθε διαδικασία καθορίζουν κατά πόσο ικανοποιούνται οι απαιτήσεις της μεθοδολογίας που εφαρμόζεται και κατά πόσο εκπληρώνονται οι στόχοι της κάθε διαδικασίας.
4. μέθοδος: η μέθοδος που εφαρμόζεται σε κάθε διαδικασία είναι επίσης καθοριστικός παράγοντας για την επίτευξη των στόχων της. Σε κάθε διαδικασία είναι δυνατόν να εφαρμόζονται διαφορετικές μέθοδοι. Οι μέθοδοι εξαγωγής σημασιολογικής πληροφορίας αξιοποιούν με διαφορετικό τρόπο τα εισαγόμενα στοιχεία. Κάποιες μέθοδοι σύγκρισης διαχειρίζονται ανεξάρτητα έννοιες, ιδιότητες, οντολογίες, ενώ άλλες τις επεξεργάζονται ενιαία. Όσον αφορά τη συσχέτιση, είναι δυνατόν να γίνεται, είτε οριζόντια μεταξύ των αρχικών οντολογιών, είτε μέσω μιας οντολογίας «υψηλού» επιπέδου, μιας οντολογίας-στόχου (target ontology) ή μιας κοινής οντολογίας (shared ontology) .
5. βαθμός αλλαγής: ο βαθμός αλλαγής – αλλοίωσης των αρχικών οντολογιών που προκαλείται από κάθε διαδικασία καθορίζει κατά πόσον θα είναι δυνατή η επαναχρησιμοποίηση των αρχικών οντολογιών.
6. βαθμός αλληλεπίδρασης: ο βαθμός αλληλεπίδρασης ή εμπλοκής του χρήστη χωρίζει την κάθε διαδικασία σε τρεις τύπους: αυτόματη, ημιαυτόματη και χειροκίνητη. Αυτό το βασικό χαρακτηριστικό δεν

αναφέρεται τόσο στο βαθμό αυτοματοποίησης της διαδικασίας που αποτελεί ζήτημα υλοποίησης, αλλά στο κατά πόσο η διαδικασία γίνεται αντικειμενικά ή η εφαρμογή της απαιτεί αποφάσεις που λαμβάνονται με υποκειμενικό τρόπο.

1.6. Εξαγωγή Σημασιολογικής Πληροφορίας

Η πρώτη διαδικασία εφαρμόζεται προκειμένου να εξαχθεί σημασιολογική πληροφορία από τις διαθέσιμες πηγές. Τα εισαγόμενα στοιχεία στη διαδικασία διαφέρουν και μπορεί να περιλαμβάνουν ελεύθερο κείμενο, θησαυρούς, ειδικό κείμενο (π.χ., ορισμούς), ιεραρχικές ταξινομήσεις, όρους, λεξικά, σχήματα βάσεων δεδομένων, κλπ. Παρόλη την ποικιλία των διαθέσιμων πηγών, η διαδικασία εξαγωγής σημασιολογικής πληροφορίας δεν είναι τετριμμένη. Η πιο συνήθης περίπτωση συνίσταται στην εξαγωγή της σημασιολογικής πληροφορίας από τις αρχικές οντολογίες: συχνά οι αρχικές οντολογίες είναι ονοματολογικές (terminological) και πιο σπάνια τυπικές (formal).

Μια σημαντική απόφαση που λαμβάνεται σε αυτήν τη διαδικασία αφορά τις υποθέσεις σχετικά με το τι αποτελεί σημασιολογική πληροφορία και τι όχι. Έτσι, σύμφωνα με αυτήν την υπόθεση, χρησιμοποιούνται πιο απλές, εμπειρικές ή πιο σύνθετες μέθοδοι για την εξαγωγή διαφορετικών στοιχείων. Οι εμπειρικές προσεγγίσεις τυποποιούν έννοιες και τις μεταξύ τους σχέσεις εξειδίκευσης που περιγράφουν συγκεκριμένα γνωστικά πεδία με σκοπό το σχεδιασμό των αντίστοιχων βάσεων δεδομένων. Οι πιο προηγμένες και συστηματικές προσεγγίσεις εξάγουν πιο πολύπλοκα στοιχεία που αφορούν σημασιολογικές

ιδιότητες και σχέσεις με βάση μεθόδους Επεξεργασίας Φυσικής Γλώσσας (Natural Language Processing) (Jensen και Binot, 1987; Vanderwende, 1995).

1.7. Σύγκριση Οντολογιών

Η δεύτερη διαδικασία αφορά τη σύγκριση των εννοιών των αρχικών οντολογιών με βάση τη σημασιολογική πληροφορία που εξήχθη στην προηγούμενη φάση. Διάφορες μέθοδοι εφαρμόζονται για τη σύγκριση εννοιών με σκοπό την αποκάλυψη και τη μέτρηση ομοιοτήτων και ετερογενειών (Hony, 2002; Klein, 2001; Kanouras κ.α., 2003).

Εκτός από τον προσδιορισμό και τη μέτρηση ομοιοτήτων και ετερογενειών, η διαδικασία αυτή περιλαμβάνει και την επίλυση των ετερογενειών προκειμένου να είναι δυνατή η ολοκλήρωση των οντολογιών στην επόμενη φάση. Η σημασιολογική ετερογένεια προκαλείται εξαιτίας διαφορετικών «συλλήψεων» των ίδιων πραγματικών οντοτήτων – φαινομένων. Η ονομαστική ετερογένεια (ομώνυμα και συνώνυμα) αποτελεί την απλούστερη μορφή σημασιολογικής ετερογένειας. Συγκριτικά με τις άλλες μορφές ετερογένειας, η σημασιολογική ετερογένεια είναι η πιο δύσκολη να επιλυθεί. Σύμφωνα με την τυπολογία του Doerg (2001), υποθέτουμε ότι οι σημασιολογικές ετερογένειες μεταξύ διαφορετικών οντολογιών οφείλονται στις παρακάτω αιτίες:

1. **διαφορετική χρήση όρων** εξαιτίας διαφορετικών γλωσσών ή λόγω επικαλυπτόμενων γνωστικών πεδίων
2. **διαφορετική κάλυψη**, εξαιτίας διαφορετικών βαθμών ανάπτυξης, διαφορετικών αντικειμένων και αναγκών. Συγκεκριμένα, ορισμένες

οντολογίες μπορεί να έχουν αναπτύξει κάποια τμήματά τους με μεγαλύτερη λεπτομέρεια από άλλες.

3. **διαφορετικές σημασιολογικές σχέσεις**, συχνά λόγω επιβολής μονής κληρονομικότητας, αλλά επίσης εξαιτίας διαφορετικών οπτικών στην ταξινόμηση.
4. **διαφορετική σημασιολογία**, εξαιτίας διαφορετικών αντιλήψεων. Οι διαφορές στη σημασιολογία προκύπτουν λόγω διαφορετικών γλωσσών, αλλά και λόγω διαφορετικών οπτικών στην ταξινόμηση.

Η δεύτερη αιτία σημασιολογικής ετερογένειας δεν προκαλεί συγκρουόμενες, αλλά αλληλοσυμπληρούμενες όψεις του ίδιου γνωστικού πεδίου και επομένως δεν εμποδίζει τη διαδικασία ολοκλήρωσης. Η τρίτη αιτία σημασιολογικής ετερογένειας μπορεί να αντιμετωπιστεί μέσω της πολλαπλής κληρονομικότητας στην ολοκληρωμένη οντολογία. Η τέταρτη αιτία σημασιολογικής ετερογένειας είναι η δυσκολότερη να εντοπιστεί και επιλυθεί. Παρακάτω γίνεται μια προσπάθεια τυποποίησης της σημασιολογικής ετερογένειας, προκειμένου να διευκολυνθεί η διαδικασία εντοπισμού και επίλυσής της.

1.8. Ολοκλήρωση Οντολογιών

Στη βιβλιογραφία χρησιμοποιούνται διάφοροι όροι, π.χ., ευθυγράμμιση (*alignment*), μερική συμβατότητα (*partial compatibility*), ενοποίηση (*unification*) (Sowa, 2000) για την ολοκλήρωση οντολογιών. Παρόλο που οι περισσότερες από αυτές τις μορφές ολοκλήρωσης ονομάζονται «σημασιολογικές», προσβλέπουν στην επίλυση τεχνικών προβλημάτων προκαλώντας αλλοίωση των αρχικών οντολογιών προκειμένου να τις κάνουν υπολογιστικά ισοδύναμες.

Στην περίπτωση που η διαδικασία ολοκλήρωσης προκαλεί αλλοιώσεις στις αρχικές οντολογίες, τότε γίνεται ανέφικτη η προς τα πίσω συμβατότητα και διαλειτουργικότητα. Ένα άλλο σχετικό ζήτημα είναι κατά πόσον η διαδικασία ολοκλήρωσης καταλήγει σε μία οντολογία ή διατηρεί όλες τις αρχικές οντολογίες. Τέλος, ένα τρίτο ζήτημα είναι αν η διαδικασία στηρίζεται μόνο στις αρχικές οντολογίες για την ολοκλήρωση ή αν χρησιμοποιεί μια οντολογία -στόχο που λειτουργεί ως κοινή οντολογία. Οι οντολογίες – στόχοι μπορεί να είναι γενικές οντολογίες ή οντολογίες πεδίου, ανάλογα με τον τύπο των αρχικών οντολογιών και το στόχο της ολοκλήρωσης. Έτσι σε σχέση με τα παραπάνω προκύπτουν οι εξής ορισμοί (Σχήμα 2).

Η **ευθυγράμμιση** (alignment) είναι μια διαδικασία απλής αντιστοίχισης μεταξύ εννοιών διαφορετικών οντολογιών. Η διαδικασία αυτή συνήθως χρησιμοποιεί εργαλεία μετάφρασης/μετατροπής. Οι αρχικές οντολογίες δεν αλλοιώνονται. Μια οντολογία στόχος μπορεί και να χρησιμοποιηθεί στη διαδικασία ευθυγράμμισης.

Η **μερική συμβατότητα** (partial compatibility) είναι μια διαδικασία ενοποίησης εκείνων των τμημάτων των αρχικών οντολογιών που παρουσιάζουν μεγαλύτερη ομοιότητα. Τα μη κοινά τμήματα των οντολογιών διατηρούνται ως έχουν στην τελική οντολογία. Μια οντολογία στόχος μπορεί και να χρησιμοποιηθεί στη διαδικασία μερικής συμβατότητας.

Η **ενοποίηση** (unification) είναι μια διαδικασία ενοποίησης όλων των εννοιών των αρχικών οντολογιών. Η ενοποίηση προκαλεί αλλαγές – αλλοιώσεις στις αρχικές οντολογίες προκειμένου να γίνουν ισοδύναμες και πλήρως συμβατές με τις υπόλοιπες. Επομένως το αποτέλεσμα της ενοποίησης είναι μία

μοναδική οντολογία, γεγονός που καθιστά τις αρχικές οντολογίες μη χρησιμοποιήσιμες. Μια οντολογία στόχος μπορεί και να χρησιμοποιηθεί στη διαδικασία ενοποίησης.

Η **πραγματική ολοκλήρωση** (true integration) είναι μια διαδικασία δημιουργίας μιας μοναδικής οντολογίας η οποία αποτελείται από τις έννοιες των αρχικών οντολογιών και κάποιες επιπρόσθετες έννοιες που διευκολύνουν τη συσχέτιση. Οι αρχικές οντολογίες δεν αλλοιώνονται, αλλά διατηρούν την ανεξαρτησία και τη χρηστικότητα τους. Μια οντολογία στόχος μπορεί και να χρησιμοποιηθεί στη διαδικασία πραγματικής ολοκλήρωσης.

1.9. Οντολογία για σημασιολογική συμβατότητα

Μια οντολογία που χρησιμοποιείται για σημασιολογική συμβατότητα είναι η ICOM/CIDOC, και τα χαρακτηριστικά της είναι τα εξής:

- **ICOM/CIDOC Conceptual Reference Model, ISO/CD 21127**
 1. Το CIDOC/CRM αποτελεί μία οντολογία για μοντελοποίηση πολιτισμικής πληροφορίας και περιγράφει σε τυπική γλώσσα τις έννοιες και σχέσεις που εμφανίζονται στην πολιτισμική τεκμηρίωση.
 2. Προτείνει μία κοινή βάση ερμηνείας για τις διάφορες μορφές τεκμηρίωσης.
 3. Δεν υπαγορεύει τα εκάστοτε στοιχεία της τεκμηρίωσης.

- Χρήση της οντολογίας
 1. Πλαίσιο για τη σχεδίαση των πληροφοριακών δομών ενός συστήματος τεκμηρίωσης.
 2. Μέσο επικοινωνίας, σημασιολογικό επίπεδο, μεταξύ ετερογενών συστημάτων.
- Το CIDOC CRM αποτελεί ακρογωνιαίο λίθο στην οικοδόμηση προσπελάσιμου, διασυνδεδεμένου ψηφιακού πολιτισμικού αποθέματος.

1.10. Οντολογίες Συστημάτων

Από την άλλη πλευρά, η έρευνα στην περιοχή των οντολογιών βρίσκεται στα πρώτα στάδια της, αν και έχουν εμφανιστεί αρκετές γλώσσες που επιτρέπουν την αναπαράσταση του. Γλώσσες όπως η SHOE, η DAML, η OIL και η υβριδική DAML+OIL[32] αποτελούν επεκτάσεις της RDF δανειζόμενες χαρακτηριστικά από αντίστοιχες γλώσσες αναπαράστασης του πεδίου της τεχνητής νοημοσύνης. Όλες αυτές οι γλώσσες βρίσκονται σε προκαταρκτικό στάδιο.

Ένας αριθμός οντολογιών έχει επίσης εμφανιστεί για την διευκόλυνση της ανταλλαγής δεδομένων μεταξύ εμπορικών εφαρμογών τύπου e-Commerce μέσω της αποδοχής κοινής σημασιολογίας για την περιγραφή της πληροφορίας. Αυτές περιλαμβάνουν την CBL (Common Business Library), την Cxml (commerce XML), την OCF (Open Financial Exchange) και άλλες. Ο σημασιολογικός χαρακτηρισμός της πολυμεσικής πληροφορίας καθώς και η διαχείριση των μετα-δεδομένων που προκύπτουν από αυτή αποτελεί ένα ανοικτό ερευνητικό πεδίο καθώς υπάρχουν πολλά ανοικτά θέματα τα οποία πρέπει να αντιμετωπιστούν. Μεταξύ αυτών η έλλειψη των κατάλληλων

εργαλείων για την εξαγωγή των κατάλληλων οντολογιών για την αναπαράστασή τους.

1.11. Συμπεράσματα Οντολογιών συστημάτων

Ήδη κυκλοφορούν κάποια συστήματα υπό την μορφή πρωτοτύπων, που λειτουργούν κατανεμημένοι χώροι αποθήκευσης (repository) για τον Παγκόσμιο Ιστό. Τα προϊόντα αυτά, υποστηρίζουν την ομοσπονδιακή (federated) ανάπτυξη και αποθήκευση των μοντέλων αναφοράς και των οντολογιών (ontologies), καθώς και την αντιστοίχιση και προβολή ανάμεσα σε διαφορετικά συστήματα μοντέλα/σχήματα, ενώ υποστηρίζουν τα πρότυπα του W3C (όπως XML και RDF). Ο σημασιολογικός Ιστός (Semantic Web) αποτελεί μια καινοτομία “εν τη γενέσει” της, η οποία υπόσχεται την ενσωμάτωση όψεων (views) βάσεων δεδομένων, “συναλλαγών” μεταξύ βάσεων δεδομένων (database transactions), λογικών αναπαραστάσεων, συνδέσμων Ιστού (Web Links) και αντικειμενοστρεφών αναπαραστάσεων, σε μια σημασιολογική βάση. Τα παραπάνω χαρακτηριστικά οδηγούν σε ένα νέο τρόπο στην επεξεργασία ερωτήσεων (query processing). Η τεχνολογία που παρουσιάζεται θα είναι υποχρεωτική, κατά τους δημιουργούς της, στις περισσότερες κατανεμημένες εφαρμογές-πελάτη, ενώ δεν υπάρχει περιορισμός ως προς τα πεδία (αγορές) εφαρμογής της.

Το όραμα της δημιουργίας του Semantic Web στηρίζεται στην επέκταση των υπάρχοντων πλαισίων περιγραφής μετα-δεδομένων και ειδικότερα στην ύπαρξη σημασιολογικού περιεχομένου που είναι δυνατόν να υπόκειται σε αυτόματη επεξεργασία από τον υπολογιστή χωρίς την επέμβασή του ανθρώπινου παράγοντα. Η ερευνητική προσπάθεια έγκειται στην δημιουργία

γενικών πλαισίων όπως το UN/SPSC και γλωσσών όπως η OWL τα οποία θα υποστηρίζουν οντολογικά όσο το δυνατόν περισσότερα πεδία εφαρμογών γίνεται.

1.12. Μεθοδολογίες ανάπτυξης συστημάτων

Η έμφαση σε μια ερμηνευτική επικοινωνία μεταξύ του σχεδιαστή και του χρήστη αποτελεί ένα βασικό επίκεντρο μιας εργασίας πάνω σε μεθοδολογίες ανάπτυξης συστημάτων, που έχει αναπτυχθεί τα τελευταία χρόνια στη Σκανδιναβία. Η συμβατική πρακτική για την ανάπτυξη συστημάτων υπολογιστών αποκαλείται «μοντέλο καταρράκτη». Αποτελείται από μια καταρρακτώδη ακολουθία βημάτων όπου οι σχεδιαστές ρωτούν πρώτα τους αγοραστές του συστήματος τι θέλουν (ένα «έγγραφο απαιτήσεων»), μετά δημιουργούν ένα επίσημο καθορισμό της λειτουργίας του συστήματος και στη συνέχεια μεταφέρουν την προδιαγραφή αυτή σε κάποιον που θα φτιάξει το πρόγραμμα. Το πρόγραμμα αφού ολοκληρωθεί, δοκιμάζεται μέχρι κάποιο βαθμό ενάντια στις προδιαγραφές, μετά αποστέλλεται πίσω, εγκαθίσταται και παραδίδεται προς χρήση. Αυτή η μεθοδολογία επικεντρώνεται στην παροχή μιας ορθής επίσημης περιγραφής των λειτουργιών του προγράμματος, συχνά με έναν τρόπο που απαιτείται από διάφορα πρωτόκολλα συμβολαίων και συμφωνιών λογισμικού, ιδιαίτερα αν οι αγοραστές είναι στρατιωτικές υπηρεσίες. Απέναντι σε αυτό το υπόβαθρο, οι υποστηρικτές μιας εναλλακτικής προσέγγισης, την οποία αποκαλούν «συμμετοχικό σχεδιασμό», θεωρούν ότι ο σχεδιασμός είναι μια ερμηνευτική δραστηριότητα, στην οποία η φύση του ολοκληρωμένου συστήματος προκύπτει από μια διαδικασία ανάπτυξης κοινών ερμηνειών μεταξύ των εμπλεκόμενων πλευρών, συμπεριλαμβανόμενων των

σχεδιαστών, των διευθυντών και των εργατών που θα έλθουν σε άμεση επαφή με τις προκείμενες συσκευές.

Ο συμμετοχικός σχεδιασμός ήταν το επίκεντρο διαφόρων σημαντικών εργασιών στη Δανία, τη Νορβηγία και τη Σουηδία, συμπεριλαμβανόμενης και μιας καλούμενης UTOPIA που πραγματοποιήθηκε στο Σουηδικό Κέντρο Εργάσιμης Ζωής σε συνεργασία με τα συνδικάτα που αντιπροσωπεύουν τους εργαζόμενους στη στοιχειοθέτηση και τους γραφίστες της Σουηδικής εκδοτικής βιομηχανίας εφημερίδων. Συμμετείχε ένας αριθμός ερευνητών οι οποίοι συνεργάστηκαν με τις εργατικές ενώσεις πάνω στην ανάπτυξη νέων εργαλείων παραγωγής ποιοτικών γραφικών.

Κατ' αρχάς, εξετάζοντας την αλληλεπίδραση μεταξύ σχεδιαστή και χρήστη, διαπιστώνεται ότι αντιμετωπίζει το ίδιο πρόβλημα που διέπει την πλαισιακή έρευνα, που περιγράφηκε στην προηγούμενη παράγραφο. Ο συμμετοχικός σχεδιασμός απαιτεί απόλυτη διαφάνεια, όπου η σχεδιαστική διαδικασία δεν ξεκινά με τον καθορισμό των λειτουργιών που θα εκτελούνται από τους υπολογιστές, αλλά από τη συνολική λειτουργική κατάσταση, ρωτώντας τι έχει νόημα να υπολογιστεί. Για την καλή σχεδίαση είναι απαραίτητοι να εργαστούμε πάνω στη διασύνδεση μεταξύ εργαλείων και εργασίας σαν ένα ζεύγος πρακτικής και τεχνογνωσίας. Όπως επισήμαναν οι ερευνητές του UTOPIA, το υπόβαθρο μέσα στο οποίο ένας ειδήμων υπολογιστών μεταφράζει την εργασία παραγωγής μιας εφημερίδας, είναι διαφορετικός από αυτόν ενός γραφίστα, που έχει αναπτύξει έναν ορίζοντα μέσα από χρόνια παρατήρησης και πρακτικής.

Η ανάλυση του ισοδυναμεί με την εφαρμογή μεθόδων όπου οι σχεδιαστικές παρεμβάσεις αναπτύσσονται εκτός της εμπειρίας. Υποστηρίζει ότι οι σχεδιαστές βιώνουν την πρακτική του χρήστη για μεγάλο διάστημα και

χρησιμοποιούν αυτό που αποκαλεί μεθόδους «σχεδίασης μέσω πρακτικής», όπως την κατασκευή μερικών πρωτοτύπων, την εκτέλεση δοκιμών πεδίων με πειραματικά συστήματα και την ανάπτυξη μακετών από τους χρήστες. Σε αντίθεση με το να προσπαθούμε να καθορίσουμε αρχικά τι θα γίνει, όπως στο μοντέλο του καταρράκτη, ο σκοπός είναι να δώσουμε στα άτομα την ευκαιρία να περάσουν μέσα από μια διαδικασία δοκιμών και λαθών, όπου ότι προκύπτει θα εκλαμβάνεται ως προδιαγραφή. Η εστίαση γίνεται σε μεθοδολογίες συμμετοχικής σχεδίασης που δεν επικεντρώνονται στη θεωρία, αλλά στη χρησιμοποίηση «σχεδιαστικών κατασκευών» όπως τα πρωτότυπα και οι μακέτες.

Εργασίες όπως το UTOPIA, έχουν αναπτύξει μια μεθοδολογία για την κοινωνική διαδικασία σχεδιασμού, όπου οι σχεδιαστές γίνονται εργαζόμενοι και οι εργαζόμενοι σχεδιαστές, σαν ένα τμήμα μιας μακράς διαδικασίας εκτός της οποίας αναπτύσσονται τα σχέδια. Το επίκεντρο δεν βρίσκεται στη θεωρητική ανάλυση της σχεδιαστικής διαδικασίας, αλλά σε πρακτικά πολιτικά και κοινωνικά θέματα που οδηγούν στο σχηματισμό κατάλληλων ομάδων, σε ανάπτυξη εμπιστοσύνης, στη συνεργασία, στη δημιουργία κατάλληλων κατασκευών και στην ανακοίνωση των συμπερασμάτων.

Ο αναπροσδιορισμός του τι είναι χρήσιμο και τι όχι, καθώς και ο επανακαθορισμός των στόχων που θα πρέπει να επιτυγχάνει ένα υπολογιστικό σύστημα, αναπόφευκτα επηρεάζουν τις διαδικασίες και τις μεθόδους σχεδιασμού του. Η κυριαρχία του «μοντέλου του καταρράκτη» (δηλαδή μιας προκαθορισμένης ακολουθίας βημάτων που συνοπτικά είναι: ο καθορισμός των απαιτήσεων από τους χρήστες – πελάτες, η ανάλυση των λειτουργιών του, και τελικά η υλοποίηση και ο έλεγχος ανταπόκρισής του στις θεσπισμένες προδιαγραφές) υποχωρεί. Ο χρήστης μπορεί να κατανοεί τις ανάγκες του και ο

σχεδιαστής τις δυνατότητες του, ωστόσο αυτό δεν αποτελεί ικανή και αναγκαία συνθήκη ενός επιτυχημένου σχεδιασμού, αφού θα απαιτηθεί οπωσδήποτε μια κοινή δίοδος επικοινωνίας. Τα κύρια ζητήματα που αντιμετωπίζουμε σε μια σχεδίαση είναι:

- Η ενημέρωση του σχεδιαστή πάνω στο αντικείμενο
- Η ισορροπία μεταξύ παραδοσιακών και σύγχρονων μεθόδων.

Δίνοντας μια σύγχρονη απάντηση σε αυτά καταλήγουμε στην ενεργό συμμετοχή του χρήστη στο σχεδιασμό. Η σχεδίαση έτσι δεν ακολουθεί διακριτά βήματα, αλλά μεταβάλλεται σε μια διαδικασία αλληλεπιδράσεων, όπου ο χρήστης μπορεί να επιλέξει ποιες από τις δυνατότητες που του προσφέρει η τεχνολογία είναι κατάλληλες γι' αυτόν, αλλά και ο σχεδιαστής να κατανοήσει καλύτερα την φύση των αναγκών του χρήστη.

1.13. Χρήση συστημάτων πολλαπλών πρακτόρων

Τα τελευταία χρόνια όλο και περισσότερες εφαρμογές διαχείρισης γνώσης (knowledge management systems) και συστήματα πολλαπλών πρακτόρων (multi-agent systems) έχουν προταθεί για να διαχειριστούν τον τεράστιο όγκο πληροφορίας που υπάρχει στα διάφορα και ετερογενή επιχειρησιακά συστήματα. Παράλληλα, η είσοδος των οντολογιών επηρεασμένες από το πεδίο της Περιγραφικής Λογικής (Descriptive Logic) προσφέρουν επαρκείς μηχανισμούς για την σημασιολογική περιγραφή της γνώσης που χαρακτηρίζει τις διαδικασίες και την πληροφορία σε μια επιχείρηση. Επίσης, η ανάπτυξη αρκετών υλοποιήσεων (Jena, Racer, Pellet, Fact++) επιτρέπουν την

εκμετάλλευση της περιγραφικότητας επιτρέποντας την εξαγωγή συμπερασμάτων από την σημασιολογική πληροφορία.

Η χρήση συστημάτων πολλαπλών πρακτόρων για την αυτόματη και έξυπνη διαχείριση πολύπλοκων διεργασιών τόσο στο διαδίκτυο όσο και σε μικρότερα δίκτυα θεωρείται πλέον μια από τις πιο υποσχόμενες λύσεις σε πολύπλοκα προβλήματα που εμπλέκουν διαφορετικά συστήματα. Αφού και η αρχιτεκτονική των συστημάτων πολλαπλών πρακτόρων έχει κατά κάποιο τρόπο παγιωθεί από την FIPA (IEEE Foundation for Intelligent Agents), το πρόβλημα πλέον έγκειται στην μοντελοποίηση της πληροφορίας και της συμπεριφοράς των πρακτόρων ώστε να αναπτυχθούν συστήματα τα οποία θα σχεδιάζουν και να παραμετροποιούν εύκολα τέτοια συστήματα. Μια προσπάθεια έχει γίνει από το Ontology Bean Generator Plugin του Protégé , που επιτρέπει την μοντελοποίηση σε οντολογία (OWL) ενός συστήματος πολλαπλών πρακτόρων και την αυτόματη παραγωγή κώδικα που με μικρές μετατροπές μπορεί να φορτωθεί στο JADE (Java Agent Development Framework).

Τα προγράμματα-πράκτορες (Software Agents) είναι προγράμματα που εκτελούν κάποια λειτουργία (συνήθως χωρίς άμεση επίβλεψη από τον χρήστη) και παράγουν αποτελέσματα με το πέρας της εκτέλεσης αυτής. Συνήθως, τα προγράμματα-πράκτορες περιδιαβαίνουν το Διαδίκτυο και επεξεργάζονται τις πληροφορίες που βρίσκουν στις σελίδες που επισκέπτονται. Προγράμματα-πράκτορες χρησιμοποιούνται ήδη για λειτουργίες όπως ανεύρεση, ταξινόμηση και επιλογή δεδομένων: π.χ. σύγκριση τιμών του ίδιου προϊόντος σε πολλά ηλεκτρονικά καταστήματα ή ειδοποίηση για την εμφάνιση νέου περιεχομένου σε δικτυακούς τόπους ειδήσεων και ενημέρωσης. Οι τεχνολογίες του Semantic Web θα οδηγήσουν σε μια έκρηξη στην αριθμό και τις δυνατότητες των

προγραμμάτων αυτών καθώς το περιεχόμενο του Web θα γίνεται όλο και πιο προσιτό και εύκολα επεξεργάσιμο από τα προγράμματα-πράκτορες.

Ένα πεδίο εφαρμογής του Semantic Web είναι οι δικτυακές υπηρεσίες: (Web services). Δικτυακές υπηρεσίες είναι εφαρμογές υπολογιστών που είναι προσπελάσιμες μέσω του Διαδικτύου (π.χ. η δυνατότητα κράτησης και έκδοσης εισιτηρίου μέσω κάποιας σελίδας στο Διαδίκτυο). Το Semantic Web θα οδηγήσει στην ανάπτυξη γλωσσών και προτύπων για την περιγραφή τέτοιων υπηρεσιών. Μέσω αυτών των γλωσσών, τα προγράμματα πράκτορες θα είναι σε θέση να ανευρίσκουν αυτόματα τέτοιες υπηρεσίες να καταλαβαίνουν τις δυνατότητες και τους περιορισμούς τους και τελικά να κάνουν χρήση αυτών των υπηρεσιών για να εκπληρώσουν τον στόχο τους.

Τα πλεονεκτήματα του Semantic Web θα επεκταθούν σταδιακά και στον φυσικό κόσμο μέσω συσκευών που θα συνδέονται στο Διαδίκτυο (π.χ. οικιακές συσκευές, κινητά τηλέφωνα, αυτοκίνητα κλπ). Όλο και περισσότερες δραστηριότητες της καθημερινής μας ζωής θα γίνονται απλούστερες και ευκολότερες χάρη στις δυνατότητες αυτοματισμού του νέου Διαδικτύου.

1.14. Οντολογική και σημασιολογική προσέγγιση στην εκπαίδευση

Οι σύγχρονες τεχνολογίες της Πληροφορικής και των Επικοινωνιών (Information Technologies –ICT) λειτουργούν ως άριστο μέσο μετάδοσης και διάδοσης υπηρεσιών εκπαίδευσης από απόσταση, συνθέτοντας έτσι εκπαιδευτικές διαδικασίες και υπηρεσίες οι οποίες περιλαμβάνουν δυο πολύ σημαντικές προσεγγίσεις :

- Την παιδαγωγική προσέγγιση προς τον εκπαιδευόμενο
- Την τεχνολογική προσέγγιση μέσω της οποίας προσφέρεται η υπηρεσία.

Ένας βέλτιστος συνδυασμός και των δυο, θα έχει σαν αποτέλεσμα μία ιδανική υπηρεσία μάθησης από απόσταση. Η παραπάνω διαδικασία της προσαρμοστικής μάθησης, πηγάζει από τις σύγχρονες ερευνητικές πρακτικές και την εισαγωγή της οντολογικής και σημασιολογικής μοντελοποίησης στις σύγχρονες διαδικασίες εκπαίδευσης.

1.15. Σημασιολογικός ιστός

Τα τελευταία δυο χρόνια γίνεται πολύς λόγος για το Semantic Web (Σημασιολογικός Ιστός), ενός διαδικτύου, δηλαδή στο οποίο η πληροφορία είναι καλύτερα ορισμένη επιτρέποντας με αυτόν τον τρόπο την ολοκλήρωση, αυτοματοποίηση και επαναχρησιμοποίηση των δεδομένων. Ο Ιστός (Web) του μέλλοντος προβλέπεται να αποτελεί μια παγκόσμια βάση δεδομένων και γνώσης με πληροφορίες οι οποίες θα είναι “κατανοητές” από μηχανές (machine – understandable information). Οι κύριες τεχνολογίες για την υλοποίηση του Σημασιολογικού Ιστού είναι ο σημασιολογικός εμπλουτισμός και η χρήση των οντολογιών. Επίσης μπορούμε να πούμε ότι ο Ιστός αποτελεί το μεγαλύτερο σε παγκόσμιο επίπεδο έργο ευφυής ενσωμάτωσης συστημάτων ώστε να συνεργάζονται δια-λειτουργικά.

Το κέντρο βάρους του περιεχομένου του Ιστού μετατοπίζεται συνεχώς από τον άνθρωπο προς τα δεδομένα. Για να φτάσει το μέγιστο των δυνατοτήτων του, πρέπει να εξελιχθεί σε ένα Σημασιολογικό Ιστό, ο οποίος παρέχει μια διεθνώς προσβάσιμη πλατφόρμα που επιτρέπει σε

αυτοματοποιημένα εργαλεία αλλά και σε ανθρώπους να μοιράζονται και να επεξεργάζονται δεδομένα.

Ο Σημασιολογικός Ιστός αποτελεί πρωτοβουλία της Κοινοπραξίας του Παγκόσμιου Ιστού (W3C), η οποία αναφέρεται λεπτομερώς στο επόμενο κεφάλαιο, και η σχετική Δραστηριότητα (W3C Semantic Web Activity) έχει δημιουργηθεί για να εξυπηρετήσει έναν ηγετικό ρόλο, τόσο στο σχεδιασμό προδιαγραφών, όσο και στην ανοικτή ανάπτυξη της τεχνολογίας μέσω της συνεργασίας.

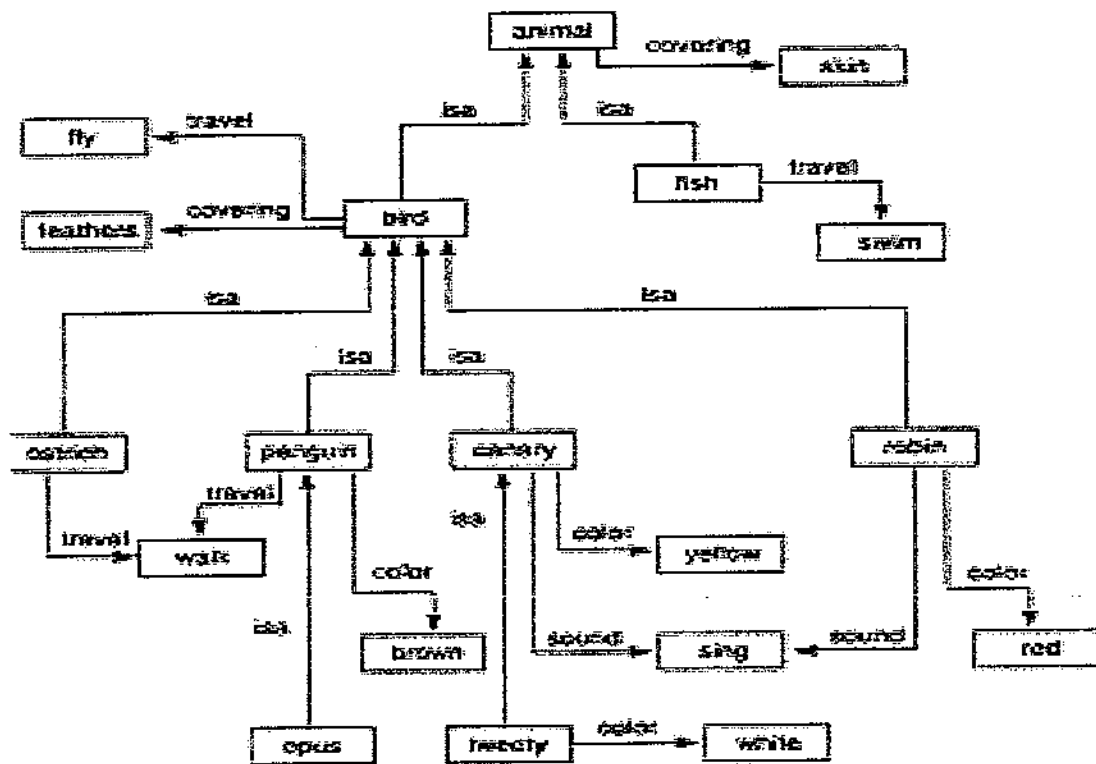
1.16. Μεταδεδομένα και Σημασιολογικός Χαρακτηρισμός

Τα μεταδεδομένα σε διάφορα πρότυπα, χρησιμοποιούνται για να περιγράψουν πληροφοριακά αντικείμενα. Ωστόσο η περιγραφή που προσφέρουν τα μεταδεδομένα από μόνα τους δεν μπορεί να χρησιμοποιηθεί στο Σημασιολογικό Ιστό, καθότι δεν παρέχει επαρκή συσχέτιση μεταξύ των πληροφοριακών αντικειμένων και δεν περιγράφει τη δομή αυτών. Σκοπός της εργασίας είναι η δημιουργία ενός συστήματος που να δίνει την δυνατότητα σε μεταδεδομένα τα οποία είναι αποθηκευμένα σε σχεσιακές βάσεις δεδομένων να αναπαρασταθούν σε κάποιο υπάρχων σημασιολογικό μοντέλο, ώστε αυτά να μπορούν να περιγράφονται σημασιολογικά. Επίσης τα μεταδεδομένα σε διάφορα πρότυπα, χρησιμοποιούνται για να περιγράψουν πληροφοριακά αντικείμενα. Ωστόσο η περιγραφή που

Εκτός από τα μετα-δεδομένα τα οποία προκύπτουν από τις προτιμήσεις των χρηστών ή την δομή ενός περιβάλλοντος hypertext, υπάρχουν και άλλα τα οποία σχετίζονται με αντικείμενα, ιδέες, διαδικασίες καθώς και των σχέσεων που τις διέπουν. Αυτά τα δεδομένα έχουν σημασιολογικό νόημα το οποίο με την

σειρά του είναι δυνατόν να περιγραφεί από φορμαλισμούς ή λεξιλόγια που καλούνται οντολογίες. Ο σημασιολογικός χαρακτήρας της πληροφορίας και η χρήση οντολογιών για την αναπαράσταση της είναι σχετικά ένας καινούργιος τομέας στην περιοχή της διαχείρισης γνώσης. Η διαδικασία της μετατροπής της πληροφορίας σε σημασιολογική είναι δυνατή μέσω ενός πλαισίου περιγραφής, του RDF (Resource Description Framework) και των RDF schemas. Το World Wide Web Consortium (W3C) έχει δημιουργήσει μία ομάδα για την δημιουργία της γλώσσας OWL, η οποία αναφέρεται στο επόμενο κεφάλαιο, και η οποία θα περιγράψει τις οντολογίες που υπάρχουν στο διαδίκτυο.

Για την περιγραφή, διαχείριση και εισαγωγή μεταδεδομένων σε κοινά έγγραφα του διαδικτύου έχουν κάνει την εμφάνισή τους τα τελευταία χρόνια αρκετά πρότυπα όπως περιγραφές τύπων κειμένου (Document Type Definitions, XML-DTDs), το νέο πρότυπο περιγραφής πολυμέσων MPEG-7 και γενικά η χρήση της γλώσσας περιγραφής XML (Extensible Markup Language), Topic Maps (σε μορφή XTM) και Semantic nets. Τέλος αυτό το κεφάλαιο θα δείξουμε με σχεδιάγραμμα την περιγραφή των μεταδεδομένων, στα οποία αναφερθήκαμε παραπάνω.



ΚΕΦΑΛΑΙΟ 2

OWL :Γλώσσα Οντολογίας του Παγκόσμιου Ιστού

2.1. Εισαγωγή

Η OWL είναι γλώσσα για τον ορισμό δομημένων οντολογιών, βασισμένων στον Παγκόσμιο Ιστό, που δίνουν την δυνατότητα για πλουσιότερη ενσωμάτωση και δια-λειτουργικότητα δεδομένων σε εύρος εφαρμογών. Οι πρώτοι που υιοθέτησαν αυτά τα πρότυπα (standards) περιλαμβάνουν κοινότητες βιο-πληροφορικής (bioinformatics) και ιατρικής, ιδιωτικές επιχειρήσεις και κυβερνήσεις. Η OWL δίνει την δυνατότητα πραγματοποίησης εύρους περιγραφικών εφαρμογών, συμπεριλαμβανομένων της διαχείρισης portals του Παγκόσμιου Ιστού, αναζητήσεων που βασίζονται στο περιεχόμενο, διαχείρισης συλλογών, ενδυναμώνοντας τους έξυπνους πράκτορες, τις υπηρεσίες του Παγκόσμιου Ιστού και γενικά όλες τις περιοχές υλοποίησης εφαρμογών των υπολογιστών.

«Η γλώσσα OWL είναι ένα σημαντικό βήμα για να γίνουν τα δεδομένα του Παγκόσμιου Ιστού πιο επεξεργάσιμα από τις μηχανές και επαναχρησιμοποιήσιμα σε εύρος εφαρμογών. Ενθαρρυνόμαστε να την δούμε να χρησιμοποιείται ήδη ως ανοιχτό πρότυπο (standard) για την ανάπτυξη οντολογιών μεγάλης κλίμακας στον Παγκόσμιο Ιστό.», είπε ο Tim Berners-Lee, Διευθυντής του W3C.

2.2. Η OWL Αποφέρει Οντολογίες που Δουλεύουν στον Παγκόσμιο Ιστό

Η OWL είναι γλώσσα Οντολογίας του Παγκόσμιου Ιστού όπως αναφέρθηκε παραπάνω. Ενώ οι προηγούμενες γλώσσες χρησίμευαν στην ανάπτυξη εργαλείων και οντολογιών για συγκεκριμένες κοινότητες χρηστών (ειδικά στις επιστήμες και σε συγκεκριμένες εφαρμογές ηλεκτρονικού

εμπορίου εταιρειών), δεν ορίζονταν ώστε να είναι συμβατές με την αρχιτεκτονική του Παγκόσμιου Ιστού γενικά, ούτε με το Σημασιολογικό Ιστό ειδικότερα. Η γλώσσα OWL το επανορθώνει αυτό, χρησιμοποιώντας URIs για ονοματοδοσία και τις διασυνδέσεις μέσω του RDF για την προσθήκη των παρακάτω δυνατοτήτων στις οντολογίες:

- Δυνατότητα κατανομής σε εύρος συστημάτων
- Επεκτασιμότητα, ανάλογα με τις ανάγκες στον Παγκόσμιο Ιστό
- Συμβατότητα με τα πρότυπα (standards) του Παγκόσμιου Ιστού για προσβασιμότητα και διεθνοποίηση
- Είναι ανοιχτές και επεκτάσιμες

Η OWL παρέχει μια γλώσσα για τον ορισμό δομημένων οντολογιών που βασίζονται στον Παγκόσμιο Ιστό, που αποφέρει πλουσιότερη ενσωμάτωση και δια-λειτουργικότητα δεδομένων ανάμεσα σε περιγραφικές κοινότητες. Επίσης χτίζει πάνω στο Μοντέλο και Σχήμα RDF, προσθέτοντας πλουσιότερο λεξιλόγιο για την περιγραφή ιδιοτήτων και τάξεων: μεταξύ άλλων σχέσεις μεταξύ τάξεων (για παράδειγμα ασυναρτησία), αριθμητικές σχέσεις (για παράδειγμα "ακριβώς ένα"), ισότητα, πλουσιότερη πληκτρολόγηση ιδιοτήτων, χαρακτηριστικά ιδιοτήτων (για παράδειγμα συμμετρία) και αριθμημένες τάξεις.

2.3. Τα Έγγραφα OWL που Παράγονται από το W3C

Η Ομάδα Εργασίας Οντολογίας του Παγκόσμιου Ιστού του W3C παρήγαγε έξι έγγραφα OWL. Κάθε ένα από αυτά προορίζεται για διαφορετικές ομάδες αυτών που επιθυμούν να μάθουν, να χρησιμοποιήσουν, να εφαρμόσουν ή να καταλάβουν τη γλώσσα OWL.

Η OWL ορίζεται μέσω 6 εγγράφων:

- Την **Επισκόπηση (OWL Overview)** που εξηγεί εν συντομία τα χαρακτηριστικά της OWL και το πώς μπορούν να χρησιμοποιηθούν

- Τη **Σημασιολογία και το Αφηρημένο Συντακτικό (OWL Semantics and Abstract Syntax)** που παρουσιάζει τη σημασιολογία και τις λεπτομέρειες αντιστοίχησης από την OWL στο RDF
- Τις **Απαιτήσεις και Σενάρια Χρήσης (OWL Use Cases and Requirements)** που αποτέλεσε κίνητρο για την OWL
- Τις **Περιπτώσεις Δοκιμών (OWL Test Cases)** που παρέχουν περισσότερα από εκατό τεστ που μπορούν να χρησιμοποιηθούν ώστε να γίνει βέβαιη η συμβατότητα των εφαρμογών της OWL με το σχεδιασμό της γλώσσας
- Τον **Οδηγό (OWL Guide)** που περιηγείται ανάμεσα στα χαρακτηριστικά της OWL με πολλά παραδείγματα της χρήσης των χαρακτηριστικών της OWL
- Την **Αναφορά (OWL Reference)** που παρέχει πληροφορίες για κάθε χαρακτηριστικό της OWL

Το στάδιο κατά το οποίο τα έγγραφα της OWL αποτελούν Υποψήφια Σύσταση εκτιμάται να διαρκέσει τουλάχιστον τέσσερις εβδομάδες, οπότε και η Ομάδα Εργασίας θα εκτιμήσει νέες εφαρμογές και σχόλια για τα προσχέδια.

2.4. Η Θέση της OWL στην Αρχιτεκτονική του Σημασιολογικού Ιστού :XML, RDF

Πολλά έχουν γραφτεί για το Σημασιολογικό Ιστό, σαν να αποτελεί αντικαταστάτρια τεχνολογία του Παγκόσμιου Ιστού που ξέρουμε στις μέρες μας. Στην πραγματικότητα δημιουργήθηκε μέσα από στοιχειώδεις αλλαγές, φέρνοντας περιγραφές, αναγνώσιμες από τις μηχανές, για τα δεδομένα και τα έγγραφα που ήδη υπάρχουν στον Παγκόσμιο Ιστό. Με τις περιγραφές αλλά και τους τρόπους σύνδεσης, σύγκρισης και αντιπαράθεσης τους, είναι δυνατή η δημιουργία εφαρμογών, εργαλείων, μηχανών αναζήτησης, πρακτόρων και όλα αυτά χωρίς προφανείς αλλαγές στις ιστοσελίδες του Παγκόσμιου Ιστού.

Η Δραστηριότητα του Σημασιολογικού Ιστού του W3C χτίζει πάνω στη δουλειά που έχει πραγματοποιηθεί μέσα σε άλλες Δραστηριότητες του W3C, όπως τη Δραστηριότητα XML. Η έμφασή της εστιάζεται στην ανάπτυξη

τεχνολογιών που είναι πρότυπα (standards) πάνω από την XML, που υποστηρίζουν την ανάπτυξη του Σημασιολογικού Ιστού. Στη βάση της, η XML παρέχει ένα σύνολο κανόνων για τη δημιουργία λεξιλογίων XML. Η XML είναι ένα ισχυρό, ευέλικτο, επιφανειακό συντακτικό για δομημένα έγγραφα, αλλά δεν επιβάλλει σημασιολογικούς περιορισμούς στο νόημα αυτών των εγγράφων.

Το RDF (Resource Description Framework), Πλαίσιο Περιγραφής Πόρων, είναι ένας τρόπος για να γίνουν απλές περιγραφές. Ότι είναι η XML για το συντακτικό, είναι το RDF για τη σημασιολογία ένα ξεκάθαρο σύνολο κανόνων για την παροχή απλών περιγραφικών πληροφοριών. Το RDF Schema παρέχει ένα τρόπο ώστε οι περιγραφές αυτές να μπορούν να συνδυαστούν σε ένα λεξιλόγιο. Αυτό που χρειάζεται μετά είναι ένας τρόπος ανάπτυξης λεξιλογίων, αναφορικά με το αντικείμενο ή την περιοχή. Αυτός είναι ο ρόλος της οντολογίας.

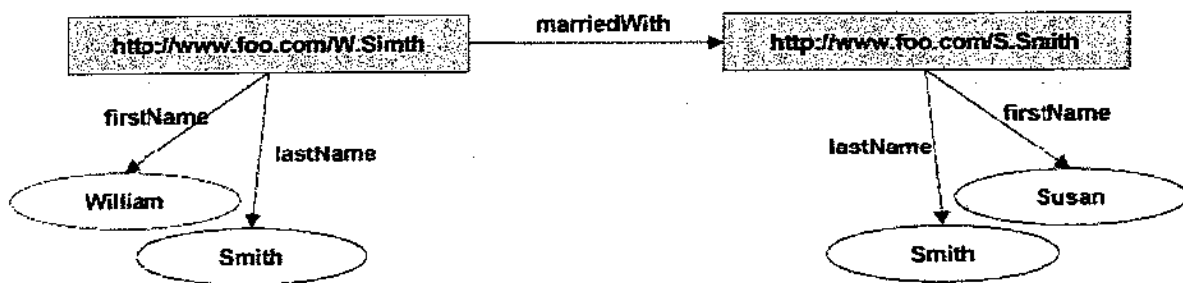
2.5. Σημασιολογικές και Οντολογικές Γλώσσες

Ο παγκόσμιος ιστός είναι μια κατάλληλη περιοχή εφαρμογής καθώς παρέχει πλούσια πληροφορία σε πολλά και διαφορετικά σχήματα τα οποία χρησιμοποιούν διαφορετική σημασιολογία κατά περίπτωση. Στο σημασιολογικό Ιστό (ΣΙ) έχουμε περάσει από την κλασική αναπαράσταση ενός εγγράφου με τη μορφή μιας ομάδας λέξεων που το χαρακτηρίζουν στη σημασιολογική αναπαράστασή του. Σε αυτή, τα δεδομένα που περιέχονται σε ένα έγγραφο αναπαρίστανται ως σύνθετες οντότητες (entities) με χαρακτηριστικά (attributes) που παίρνουν διάφορες τιμές. Οι οντότητες που εμφανίζονται σε ένα έγγραφο συνδέονται μεταξύ τους με σχέσεις (relation). Οντότητες που εμφανίζουν κοινά χαρακτηριστικά ομαδοποιούνται και περιγράφονται με αφηρημένες έννοιες (τάξεις -classes). Ακολουθεί μια επισκόπηση των βασικών μοντέλων που μελετήθηκαν.

Το μοντέλο RDF

Ένα μοντέλο που χρησιμοποιείται ευρύτατα για την αναπαράσταση

δεδομένων στο σημασιολογικό ιστό είναι το RDF και η επέκτασή του RDF-Schema. Ο βασικός μηχανισμός αναπαράστασης στο RDF είναι η τριάδα υποκείμενο-όρισμα-αντικείμενο (subject-predicate-object). Το αντικείμενο και το υποκείμενο αποτελούν οντότητες (resources) μέσα στο έγγραφο, κάθε ένα χαρακτηρίζεται μοναδικά από ένα URI ενώ έχει μια σειρά από ιδιότητες (properties) με αντίστοιχες τιμές. Η σχέση μεταξύ αντικειμένου και υποκειμένου ορίζεται ως ξεχωριστή οντότητα (ιδιότητα) του υποκειμένου που το συνδέει με το αντικείμενο.



Το RDF υποστηρίζει την κληρονομικότητα μεταξύ των resources (sub-Class), αλλά και μεταξύ σχέσεων (sub-Property) επιτρέποντας έτσι την εξαγωγή υπονοούμενων κανόνων (γενίκευση, μεταβατική ιδιότητα κλπ).

Το μοντέλο RDFS

Οι οντότητες του Semantic Web που περιγράφονται με RDF μεταφέρουν πληροφοριακό περιεχόμενο το οποίο είναι πολύ πιο χρήσιμο όταν εντοπίζεται σε μια θεματική περιοχή. Η επέκταση RDFS (RDF-Schema) μας επιτρέπει να περιγράψουμε τη θεματική περιοχή και το λεξιλόγιο που χρησιμοποιούν οι οντότητες αυτής. Συνεπώς είναι δυνατός ο ορισμός ομάδων οντοτήτων με κοινά χαρακτηριστικά (τάξεις - Classes). Είναι δυνατός επίσης ο ορισμός σχέσεων μεταξύ τάξεων (ως Properties που φέρουν κάποιο όνομα και συνδέουν μια ή περισσότερες τάξεις –domain– με κάποια τάξη – range).

Το RDFS μπορεί να χρησιμοποιηθεί για την περιγραφή μιας βάσης γνώσης κυρίως λόγω της επεκτασιμότητας που εμφανίζει και της απλής

δομής του. Παρόλα αυτά εμφανίζει σημαντικούς περιορισμούς:

- Ορίζει μονόδρομες σχέσεις και συνεπώς δεν υποστηρίζει μεταβατικότητα, αντίθεση και συμμετρία.
- Δεν επιτρέπει τον περιορισμό μιας σχέσης σε ένα πεδίο (π.χ. αν η σχέση hasChild οριστεί για ανθρώπους και σκύλους δεν μπορεί στη συνέχεια να αποκλείσει χιαστή συνδυασμούς
- Δεν επιτρέπει περιορισμούς πλήθους ή ύπαρξης (cardinality or existence)
- Τέλος ένας σημαντικός περιορισμός είναι ότι δεν υποστηρίζει αυτόματη συλλογιστική (automated reasoning).

Το μοντέλο DAML+OIL

Το μοντέλο αναπτύχθηκε πάνω στο RDF για να αναπαριστά γνώση και να υποστηρίζει αυτόματη εξαγωγή συμπερασμάτων από αυτή. Ξεπερνά τους περιορισμούς που αναφέρθηκαν στο RDFS και επιτρέπει την περιγραφική λογική (description logic) και την εξαγωγή συμπερασμάτων. Έχουν αναπτυχθεί αρκετά εργαλεία που χρησιμοποιούν το DAML για την περιγραφή της πληροφορίας και αναπτύσσουν εργαλεία εξαγωγής συμπερασμάτων και κανόνων πάνω σε αυτό.

Το μοντέλο OWL

Το πληρέστερο από τα υπόλοιπα, υποστηρίζει Description Logic και έχει τρεις διαφορετικές εκδόσεις **OWL full** (ένωση **OWL** και **RDF**), **OWL DL** που είναι ισοδύναμη της **DAML-OIL** και **OWL Lite** που αποτελεί υποσύνολο της **OWL DL**.

Ενώ λοιπόν η RDFS είναι αρκετή για τον ορισμό ιεραρχιών από οντότητες και τον καθορισμό των ιδιοτήτων τους, η OWL Lite επιτρέπει τον ορισμό αξιωμάτων (κανόνων) και τον περιορισμό των χαρακτηριστικών σε ορισμένα μέλη μιας τάξης (όπως η DAML OIL). Η OWL DL και OWL Full επιτρέπουν πλήρη εκφραστικότητα (π.χ. λειτουργίες ένωσης, τομής, σύνθεση/ισοδυναμία/ κληρονομικότητα τάξεων κτλ) με την πρώτη να θέτει κάποιους

περιορισμούς που διευκολύνουν την ανάπτυξη εργαλείων συλλογιστικής.
RDF(S) □ OWL Lite □ OWL DL,

Όπου το □ αναφέρεται σε συντακτικά και σημασιολογικά χαρακτηριστικά

Η γλώσσα OWL DL επιτρέπει των ορισμό αξιωμάτων όπως το ακόλουθο:

$\text{Man} \sqcap (\exists \text{CHILD. T}) \sqcap$
 $\forall \text{CHILD. (Doctor} \sqcap \exists \text{FRIEND. (Rich} \sqcup \text{Famous))}$,

π.χ., άντρες που έχουν παιδιά και όλα τους τα παιδιά είναι γιατροί με πλούσιους και διάσημους φίλους Με την προσθήκη αξιωμάτων σε μια οντολογία δημιουργούμε μια βάση γνώσης *knowledge base* η αλλιώς μια ορολογία (terminology - TBox), την οποία μπορεί στη συνέχεια να χρησιμοποιήσει μια μηχανή εξαγωγής συμπερασμάτων (iFact, Racer) και να επεκτείνει τα αξιώματα χρησιμοποιώντας τις λειτουργίες που προσφέρει το μοντέλο OWL DL.

2.6. Μεθοδολογίες Μετάφρασης και Ολοκλήρωσης Σχημάτων

Αρχικά θα επιχειρήσουμε την καταγραφή και μελέτη των υπαρχόντων προσπαθειών μετάφρασης σχημάτων γνώσης καθώς και των τεχνικών για προσθήκη γνώσης σε βάσεις γνώσης. Διακρίνουμε στην προσπάθεια αυτή δύο βασικές κατευθύνσεις: α) τον ορισμό γλωσσών επερωτήσεων σε σχήματα γνώσης, β) την ανάπτυξη συστημάτων για τη διαχείριση και επερώτηση σχημάτων γνώσης.

1. Στην πρώτη κατηγορία συμπεριλαμβάνονται γλώσσες επερωτήσεων προς RDF σχήματα όπως η RQL ή η DQL. Με τις γλώσσες αυτές οντολογική γνώση που έχει αναπαρασταθεί σε RDF μπορεί εύκολα να ερωτηθεί, να αποκοπεί, να μετασχηματιστεί ή να συγκεντρωθεί. Η γνώση αυτή συχνά προέρχεται από κάποια υπηρεσία ιστού και είτε ενημερώνεται από κάποιον ειδικό. Τα συστήματα αυτά λειτουργούν ως wrappers και διευκολύνουν την ολοκλήρωση σχημάτων. Τέλος συμπεριλαμβάνονται και APIs όπως τα Jena, JeSS και KAON που χρησιμοποιούνται από μεγαλύτερα συστήματα διαχείρισης γνώσης για

την υποβολή ερωτήσεων και την εξαγωγή κανόνων (inference engines). Το Jena είναι μια εφαρμογή σε Java για τη διαχείριση OWL και RDF/RDFS αρχείων το οποίο υποστηρίζει reasoning. Το KAON είναι μια ολοκληρωμένη πλατφόρμα για τη διαχείριση οντολογιών σε RDF που όμως δεν υποστηρίζει RDFS και OWL.

2. 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:

- **Sesame**: που αποθηκεύει και διαχειρίζεται RDF και RDF Schema αρχεία (στη μνήμη ή σε βάση δεδομένων) και υποστηρίζει επερωτήσεις σε SeRQL, RQL και RDQL
- **OWL JessKB** :που διαχειρίζεται έγγραφα OWL στη μνήμη χρησιμοποιώντας το Jess API που αναφέραμε.
- **DLDB-OWL** :που αποθηκεύει σε σχεσιακή ΒΔ δεδομένα OWL data. Πριν τη δημιουργία της βάσης χρησιμοποιεί τον FaCT reasoner και εξαγει αυτόματα κανόνες ιεραρχίας τάξεων (subsumption) από τα OWL δεδομένα. Οι κανόνες αυτοί εμπλουτίζουν τη ΒΔ με απόψεις που διευκολύνουν την απάντηση ερωτήσεων που αφορούν στιγμιότυπα. (extensional queries)
- **TRIPLE**: που διαχειρίζεται επερωτήσεις σε RDF αρχεία υποστηρίζοντας μεταβατικές ιδιότητες (subClassOf, subPropertyOf κλπ.). Το TRIPLE επίσης δεν διαθέτει API κάτι που το καθιστά δύσκολα επεκτάσιμο.
- Το **OntoTrack** που υποστηρίζει RDF, DAML+OIL και OWL αρχεία και περιορισμένους τύπους επερωτήσεων και

ενσωματώνει το 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 επερωτήσεων από αυτά που υποστηρίζει (custom queries).

Στόχος των συστημάτων αυτών είναι η αποθήκευση στη μνήμη ή σε μια ΒΔ 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 (reasoners). Σημαντικός παράγοντας για την επιτυχία των παραπάνω συστημάτων είναι η πληρότητα των απαντήσεων (completeness: πόσα από τα στιγμιότυπα εντοπίζουν-αντίστοιχο του recall) και η σαφήνεια (soundness: πόσα από τα στιγμιότυπα που εντοπίζουν είναι σωστά αντίστοιχο του precision).

Στα προηγούμενα συστήματα θα πρέπει να προστεθούν και APIs που στόχο έχουν τη σύγκριση οντολογιών. Το Alignment API επιτρέπει την εξαγωγή αντιστοιχήσεων μεταξύ δύο οντολογιών (ontology concept matching) την μετατροπή τους σε αξιώματα εκφρασμένα σε OWL και το μετασχηματισμό από τη μία στην άλλη (π.χ. με δημιουργία XSLT scripts). Το σύστημα Prompt παρομοίως συγκρίνει δύο διαφορετικές εκδόσεις μιας οντολογίας σε RDF-Schema και εντοπίζει διαφορές στη δομή και στα ονόματα που χρησιμοποιούνται (αλγόριθμος PromptDiff). Στα συστήματα αυτά οι οντολογίες που έρχονται στην είσοδο παρουσιάζουν μεγάλες ομοιότητες και μικρές συντακτικές διαφορές. Στόχος είναι ο εντοπισμός των διαφορών και η έκφρασή τους με τη μορφή κανόνων μετάβασης από τη μία οντολογία στην άλλη. Στην περίπτωση που οι οντολογίες που συγκρίνουμε παρουσιάζουν σημαντικές διαφορές τα συστήματα δεν είναι αποτελεσματικά. Στη συνέχεια θα δούμε συστήματα που αναλαμβάνουν τη διαχείριση οντολογιών και σχημάτων γνώσης από διαφορετικές πηγές και αποσκοπούν στην ολοκλήρωσή τους σε μια ενιαία βάση γνώσης. Στα συστήματα αυτά συχνά υπάρχει μια μεγάλη οντολογία αναφοράς και κανόνες γνώσης που εξάγονται από τα δεδομένα με τη βοήθεια της οντολογίας και στη συνέχεια

προσαρτώνται σε αυτή.

2.7. Σχετικά με την Κοινοπραξία Παγκόσμιου Ιστού [W3C]

Το W3C δημιουργήθηκε για να οδηγήσει τον Ιστό στην πλήρη αξιοποίηση του αναπτύσσοντας κοινά πρωτόκολλα που προάγουν την εξέλιξη του και εξασφαλίζουν τη δια-λειτουργικότητα του. Είναι μία διεθνής βιομηχανική κοινοπραξία που τη διαχειρίζονται σε συνεργασία το MIT Laboratory for Computer Science (MIT LCS) στην Αμερική, το ERCIM-European Research Consortium for Informatics and Mathematics με αρχηγείο στη Γαλλία και το Πανεπιστήμιο Keio στην Ιαπωνία. Οι υπηρεσίες που παρέχονται από την Κοινοπραξία περιλαμβάνουν: μία αποθήκη πληροφοριών σχετικά με τον Ιστό για τους κατασκευαστές και τους χρήστες και ποικίλες εφαρμογές πρωτοτύπων και δειγμάτων για να επιδείξουν τη χρήση της νέας Τεχνολογίας. Μέχρι σήμερα, σχεδόν 400 οργανισμοί είναι Μέλη (Members) της Κοινοπραξίας.

2.8. Βιομηχανικοί και Ακαδημαϊκοί Ηγέτες Προωθούν την OWL Μπροστά

Η Ομάδα Εργασίας Οντολογίας του Παγκόσμιου Ιστού του W3C εμπεριέχει συμπλήρωμα ακαδημαϊκής και βιομηχανικής τεχνογνωσίας, αποφέροντας το ερευνητικό βάθος και την εμπειρία εφαρμογών που είναι απαραίτητα για το χτίσιμο ενός ισχυρού συστήματος γλώσσας οντολογίας. Οι συμμετέχοντες περιλαμβάνουν τους οργανισμούς: Agfa-Gevaert N.V, Daimler Chrysler Research and Technology, DARPA, Defense Information Systems Agency (DISA), EDS, Fujitsu, Forschungszentrum Informatik (FZI), Hewlett Packard Company, Ibrov, IBM, INRIA, Ivis Group, Lucent, University of Maryland, Mondeca, Motorola, National Institute of Standards and Technology (NIST), Network Inference, Nokia, Philips, University of Southampton, Stanford University, Sun Microsystems, Unicorn Solutions καθώς και

προσκεκλημένους ειδικούς από τους οργανισμούς: German Research Center for Artificial Intelligence (DFKI) GmbH, the Interoperability Technology Association for Information Processing, Japan (INTAP) και University of West Florida.

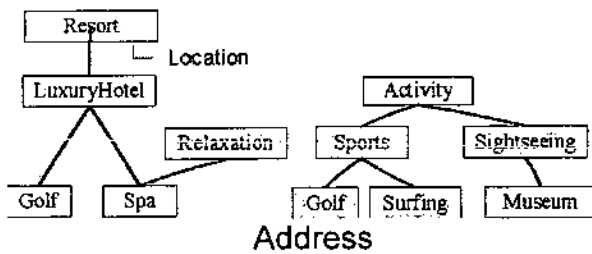
Η OWL φέρνει κοντά την έρευνα ομάδων που δραστηροποιούνται στην ανάπτυξη γλωσσών για να αποδώσουν οντολογικές εκφράσεις στον Παγκόσμιο Ιστό (languages in which to express ontological expressions on the web). Η OWL ξεκίνησε μέσα από δύο σημαντικές ερευνητικές προσπάθειες: από ένα προσχέδιο γλώσσας, γνωστό ως παραστάσεις Οντολογίας της Γλώσσας Σήμανσης Πρακτόρων DARPA (DAML-ONT) και Στρώμα Διεπαφής Οντολογίας (Ontology Interface Layer –OIL) που αναπτύχθηκαν από Ευρωπαίους ερευνητές με την υποστήριξη της Ευρωπαϊκής Επιτροπής. Έκτοτε, μία ειδική ομάδα ερευνητών έχει σχηματίσει την κοινή επιτροπή Ηνωμένων/ Ευρωπαϊκής Ένωσης για Γλώσσες Σήμανσης Πρακτόρων (Joint US/EU committee on Agent Markup Languages) και εξέδωσε μία νέα έκδοση αυτής της γλώσσας που συγχωνεύει τη DAML με την OIL. Τα έγγραφα που εκδόθηκαν σήμερα αντικατοπτρίζουν τη συνεργασία των διεθνών ερευνητών με συμμετέχοντες από τη βιομηχανία που δουλεύουν μαζί στην Κοινοπραξία του Παγκόσμιου Ιστού.

2.9. Περιπτώσεις εμπλουτισμού και συγχώνευσης βάσεων γνώσης

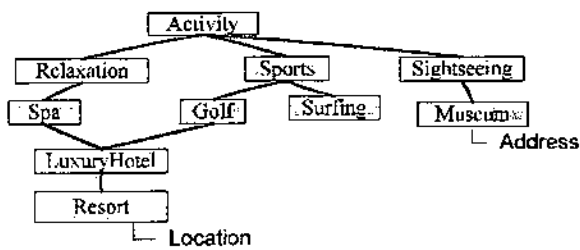
Η ύπαρξη μεγάλων οντολογιών σε συγκεκριμένα πεδία εφαρμογής και η συντήρησή τους από διάφορες ομάδες ειδικών δημιούργησε την ανάγκη για εργαλεία που θα υποστηρίζουν τη συλλογική ανάπτυξη μιας οντολογίας (με συγχώνευση επιμέρους οντολογιών), θα ανιχνεύουν σφάλματα και ασυμβατότητες σε αυτή και θα τη διατηρούν ενήμερη με το χρόνο. Παράλληλα, η ύπαρξη πολλών πηγών δεδομένων και υπηρεσιών γνώσης που καθεμία χρησιμοποιεί τη δική της οντολογία οδήγησε στην ανάπτυξη εργαλείων που κάνουν χρήση της οντολογίας για την εξαγωγή γνώσης αλλά και την αντιστοίχιση και συγχώνευσή της στην οντολογία.

Τα τελευταία χρόνια αναπτύχθηκαν μεγάλες και σύνθετες οντολογίες σε πεδία

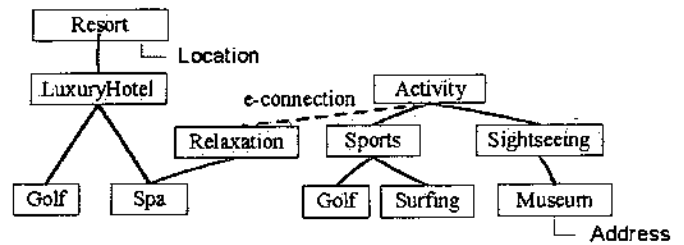
την πολυπλοκότητα των επερωτήσεων καθώς εισάγει πλεονασμούς.



α



β



γ

Σχήμα 1

Ένας μεγάλος αριθμός ερευνητικών προσπαθειών υπάρχει στο χώρο της συγχώνευσης πηγών δεδομένων. Στις προσπάθειες αυτές θεωρείται ότι τα δεδομένα διαφέρουν τόσο στο σχήμα (π.χ. XML έγγραφα με διαφορετικά DTDs) όσο και στα σημασιολογικά χαρακτηριστικά (τα elements και attributes που χρησιμοποιούνται διαφέρουν) και ο στόχος είναι να αναπτυχθεί ένα περιβάλλον που θα ενοποιεί τα διαφορετικά σχήματα και θα επιτρέπει την υποβολή ερωτήσεων. Για την ενοποίηση απαιτείται καταρχήν ταίριασμα των στοιχείων κάθε σχήματος που γίνεται με εκπαίδευση ενός συστήματος μηχανικής μάθησης. Το σύστημα ορίζει κατηγορίες πληροφορίας για το πεδίο ενδιαφέροντος (π.χ. για αγορές ακινήτων ορίζει τις κατηγορίες διεύθυνση, τιμή, περιγραφή, τηλέφωνο) και ταξινομητές που κατατάσσουν ένα δεδομένο σε μια κατηγορία πληροφορίας (ανάλογα με την τιμή του, την ετικέτα στην XML, τη θέση του μέσα στο XML ή HTML έγγραφο κλπ.) με κάποιο βαθμό πίστης (π.χ. τα περιεχόμενα της ετικέτας address, location, οι αριθμοί που ακολουθούνται από τη λέξη ευρώ, που ακολουθούν τη λέξη τηλ. κλπ). Οι επερωτήσεις μας στη συνέχεια γίνονται στο ενοποιημένο σχήμα που περιλαμβάνει τις γενικές κατηγορίες πληροφορίας και εμπλουτίζεται με τα δεδομένα των πηγών (για παράδειγμα τα περιεχόμενα των ετικετών Location και Address στο σχήμα 1 συγχωνεύονται στην κατηγορία Location που αφορά

όπως:

- η ιατρική: SNOMED, Gene (22000 έννοιες) και NCI (27500 έννοιες)

- το εμπόριο: UNSPCS, NAICS αλλά και διαθεματικά όπως το DMOZ που περιλαμβάνει διευθύνσεις του παγκόσμιου ιστού σε πάνω από 200.000 κατηγορίες. Οι οντολογίες αυτές ξεκινούν ορίζοντας μια ή περισσότερες ταξινομίες των εννοιών (με τις σχέσεις ISA και HASA) και στη συνέχεια τις εμπλουτίζουν διασυνδέοντας όρους μεταξύ των ταξινομιών με επιπλέον σχέσεις.

Στο χώρο αυτό θα πρέπει να αναφέρουμε και τις προσπάθειες που έγιναν για την μετατροπή των υπάρχοντων βάσεων γνώσης σε μοντέλα όπως η RDF, η OWL ή η DAML-OIL, ώστε να διευκολύνονται οι συγχωνεύσεις, η αιτιολόγηση (reasoning) των παραγόμενων γνώσεων κλπ. Μια οντολογία που είναι εκφρασμένη σε κάποιο από τα ευρέως διαδεδομένα μοντέλα μπορούμε εύκολα να τη διαχειριστούμε.

Συγχώνευση βάσεων γνώσεων

Οι οντολογίες αυτές εμπλουτίζονται και συντηρούνται από ομάδες ειδικών. Αυτοί δημιουργούν ξεχωριστά κομμάτια οντολογίας τα οποία προσαρτώνται στη συνέχεια στην οντολογία. Για την υποστήριξη αυτής της διαδικασίας αναπτύχθηκαν εργαλεία όπως το Chimera που διευκολύνει τη συγχώνευση πολλών μικρών οντολογιών σε μία μεγαλύτερη. Κάνοντας χρήση reasoners και με κατάλληλο ταίριασμα όρων εντοπίζει σημεία όπου υπάρχουν συγκρούσεις ονομάτων, ελλείψεις στις περιγραφές κλπ., χωρίς όμως να επεμβαίνει στην οντολογία. Άλλες προσπάθειες επεκτείνουν γλώσσες όπως η OWL με έννοιες αντίστοιχες των υπερσυνδέσμων (e-connections) που επιτρέπουν τη σύνδεση όρων μεταξύ δύο οντολογιών παρακάμπτοντας έτσι τη διαδικασία της συγχώνευσης. Αν θελήσουμε να κάνουμε συγχώνευση δύο οντολογιών (Σχήμα 1α) θα πρέπει να εντοπίσουμε σε ποιο σημείο της δεύτερης οντολογίας θα γίνει η εισαγωγή κάθε έννοιας της πρώτης (κάτω από ποια έννοια) και να ελέγξουμε ότι η δομή και τα χαρακτηριστικά της έννοιας δεν έρχονται σε αντίθεση με τη δομή της οντολογίας (Σχήμα 1β). Με τη χρήση υπερσυνδέσμων η σύνδεση μεταξύ οντολογιών είναι πιο «χαλαρή» χωρίς να προκαλεί αλλαγές στις δύο οντολογίες (Σχήμα 1γ), ταυτόχρονα όμως αυξάνει

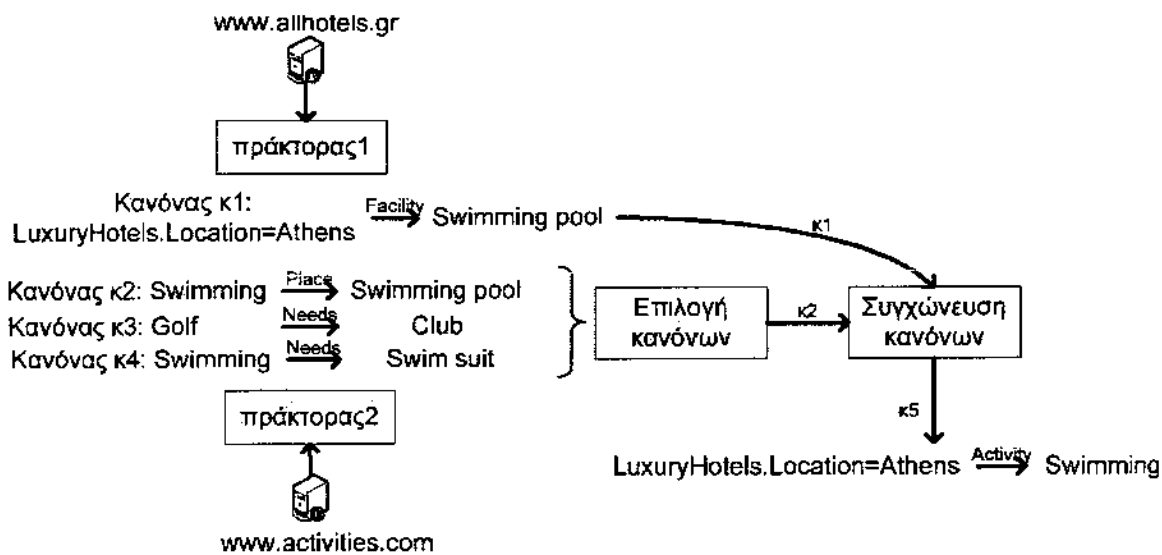
την τοποθεσία ενός κτιρίου).

Στην περίπτωση που οι πηγές χρησιμοποιούν διαφορετική σημασιολογία, εκφρασμένη με τη μορφή οντολογιών, το πρόβλημα που έχουμε να αντιμετωπίσουμε μετατοπίζεται στο πεδίο συγχώνευσης (ή αντιστοίχισης) οντολογιών. Η αντιστοίχιση μπορεί να γίνει με ή χωρίς υποστήριξη από το χρήστη (GLUE). Στόχος είναι να αντιστοιχηθούν μεταξύ τους αρχικά οι έννοιες (κατηγορίες πληροφορίας) της κάθε οντολογίας και στη συνέχεια και τα δεδομένα που εντάσσονται σε αυτές (στιγμιότυπα).

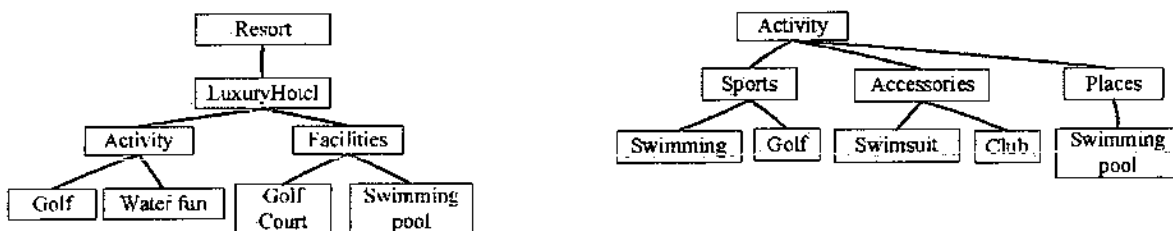
Αν η σύνδεση των διαφόρων οντολογιών γίνει με επιτυχία, τότε η αντιστοίχιση μπορεί να χρησιμοποιηθεί για να συγχωνευτούν κανόνες και άλλα σχήματα γνώσης που έχουν προκύψει από την κάθε πηγή, δίνοντας έτσι ακόμη μεγαλύτερο ενδιαφέρον στο τελικό αποτέλεσμα (π.χ. καλύτερες περιγραφές κατηγοριών πληροφορίας). Περιγράφεται μια διαδικασία συγχώνευσης γνώσης (κανόνες) που προέρχεται από πολλούς πράκτορες που χρησιμοποιούν διαφορετικά λεξικά. Στη φάση της αυτόνομης μάθησης κάθε πράκτορας λειτουργεί ανεξάρτητα και παράγει κανόνες. Στη φάση της ολοκλήρωσης που ακολουθεί έχουμε μια δεύτερη διαδικασία επαγωγικής μάθησης στην οποία από τα παραδείγματα που επαληθεύουν κάθε κανόνα προκύπτουν οι αντιστοιχήσεις όρων μεταξύ των διαφόρων λεξικών. Οι αντιστοιχήσεις προκύπτουν έπειτα από πολλές μετεγγραφές και επανεκτιμήσεις των κανόνων (πόσες περιπτώσεις τους ικανοποιούν), οι οποίες περιλαμβάνουν γενικεύσεις, ειδικεύσεις όρων κ.ά. Ένα πρόβλημα που πρέπει να αντιμετωπιστεί στη φάση αυτή είναι το ποιοι κανόνες θα συγχωνευτούν και ποιοι θα απορριφθούν από την επόμενη φάση. Αρχικά ορίζεται η κατηγορία στην οποία ανήκει ο κανόνας (από το δεξί μέρος του κανόνα) και προφανώς η συγχώνευση γίνεται μεταξύ κανόνων της ίδιας κατηγορίας. Στη συνέχεια ορίζονται μέτρα αξιολόγησης των κανόνων όπως: η ακρίβεια (το ποσοστό όλων των περιπτώσεων που τους ικανοποιούν) και η κάλυψη που εμφανίζουν ως προς την κατηγορία (σχετίζεται με το ποσοστό των περιπτώσεων της κατηγορίας που ικανοποιούν τον κανόνα). Στην τελική φάση γίνονται διαδοχικές συγχωνεύσεις (με χρήση αξιωμάτων κατηγορηματικής λογικής) των κανόνων με τη μεγαλύτερη κάλυψη και

ακρίβεια για να προκύψουν πληρέστεροι κανόνες και να εξαλειφθούν οι πλεονασμοί.

Το πλεονέκτημα της χρήσης πολλών πρακτόρων για την εξαγωγή και συγχώνευση γνώσης από πολλές πηγές είναι ότι συγκεκριμένοι πράκτορες μπορούν να προσαρμοστούν στις ιδιαιτερότητες συγκεκριμένων πηγών και να δώσουν πολύ καλύτερα αποτελέσματα από ένα γενικό πράκτορα. Η συγχώνευση των αποτελεσμάτων έρχεται στη συνέχεια να δώσει πληρέστερη γνώση αποφεύγοντας τους πλεονασμούς. Ένα παράδειγμα που περιγράφει την προηγούμενη διαδικασία φαίνεται στο σχήμα 2. Χρησιμοποιώντας την 1^η οντολογία του σχήματος 3, ένας πράκτορας αντλεί από κάποιο δικτυακό τόπο με τουριστικές πληροφορίες δεδομένα και εξάγει τον κανόνα κ1. Ένας δεύτερος πράκτορας χρησιμοποιώντας τη δεύτερη οντολογία και στοιχεία από ένα άλλο δικτυακό τόπο εξάγει τρεις κανόνες (κ2, κ3, κ4). Στο πρώτα στάδιο της συγχώνευσης θα επιλεγεί ποιος από τους κανόνες κ2, κ3, κ4 θα συγχωνευτεί με τον κ1. Εδώ μπορεί να χρησιμοποιηθεί μια γενικότερη οντολογία, κάποιοι λεκτικοί κανόνες ή η ίδια η δομή της οντολογίας (οι κ3, κ4 απορρίπτονται γιατί αφορούν εξαρτήματα ενώ ο κ1 αφορά χώρο και άθλημα - σχήμα 4α).

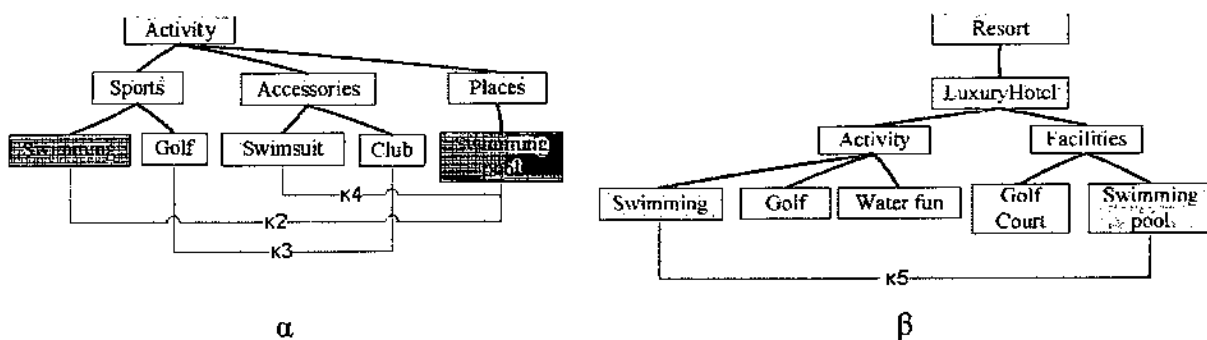


Σχήμα 2



Σχήμα 3

Από τη συγχώνευση των κανόνων κ1 (τα πολυτελή ξενοδοχεία στην Αθήνα έχουν πισίνα) και κ2 (η κολύμβηση γίνεται σε πισίνα) προκύπτει ο κανόνας κ5 (τα πολυτελή ξενοδοχεία στην Αθήνα προσφέρουν τη δραστηριότητα κολύμβηση).



Σχήμα 4

Εμπλουτισμός βάσεων γνώσης

Με στόχο την ανάπτυξη οντολογιών από έγγραφα του Ιστού έχουν αναπτυχθεί συστήματα που επεξεργάζονται τα κείμενα και εξαγάουν γνώση που μπορεί στη συνέχεια να ενταχθεί ημι-αυτόματα σε μια οντολογία ή μια ορολογία. Αυτές οι προσπάθειες αντλούν όρους, σχέσεις, κανόνες κλπ και

προσπαθούν να εντάξουν την παραγόμενη γνώση σε μια βάση γνώσης. Ακολουθώντας το παράδειγμα, οι γνώσεις που εξάγονται είτε περιγράφουν στιγμιότυπα τάξεων της οντολογίας (π.χ. Hilton:LuxuryHotel, Hilton.Location-Athens, Palace:LuxuryHotel, Palace.Location- Piraeus), είτε τα συνδέουν με στιγμιότυπα άλλων τάξεων της οντολογίας. Δημιουργούνται έτσι περιπτώσεις στιγμιότυπων που επιβεβαιώνουν ή διαψεύδουν κανόνες σαν κι αυτούς του παραδείγματος facilities (Hilton,Swimming pool), facilities (Palace, Swimming pool), activities (Hilton, water fun) κλπ.).

Υπάρχουν οι εξής τεχνικές που περιλαμβάνουν συντακτική ανάλυση των κειμένων και συσταδοποίηση των εξαγόμενων όρων, ανάλυση της συχνότητας εμφάνισης των όρων (MetaMap) ή παραγωγή κανόνων συσχέτισης.

- Συσταδοποίηση όρων: με τη χρήση κάποιου λεξικού ή θησαυρού όρων και σε συνδυασμό με τη σύνταξή τους μέσα στα κείμενα οι όροι οργανώνονται σε ιεραρχίες (ιεραρχική συσταδοποίηση). Αν η ίδια διαδικασία εφαρμοστεί στο κείμενο ερμηνευτικών λεξικών εντοπίζει συσχετίσεις μεταξύ όρων του λεξικού και όρων που εμφανίζονται στις περιγραφές.
- Ανάλυση της συχνότητας εμφάνισης: όροι μιας οντολογίας που εμφανίζονται συχνά μαζί σε ένα σώμα κειμένων συσχετίζονται μεταξύ τους. Χαρακτηριστικό είναι το παράδειγμα των όρων Raynaud και Fish Oil που συσχέτισε την ασθένεια με τη θεραπεία της μέσα από ανάλυση ιατρικών κειμένων.
- Κανόνες συσχέτισης: η γνώση που προκύπτει από τη συσταδοποίηση ή συνεμφάνιση όρων μπορεί να αναπαρασταθεί με τη μορφή κανόνων, της μορφής $\{wi\} \square \{wj\}$, που ερμηνεύονται ως εξής: αν υπάρχουν οι όροι w_i σε ένα κείμενο τότε θα υπάρχουν και οι όροι w_j . ή αντίστοιχα οι σύνθετοι όροι w_i και w_j συνδέονται συντακτικά και σημασιολογικά μεταξύ τους.

Σε όλες αυτές τις περιπτώσεις η συντακτική και στατιστική ανάλυση των κειμένων δίνει ομάδες λέξεων βασιζόμενη στις συνεμφάνσεις. Στη συνέχεια

προσπαθεί να εντοπίσει το είδος της σχέσης που συνδέει τις λέξεις (π.χ. σύνθετοι όροι, αντίθετοι, συνώνυμοι, κλπ), τη σχέση τους με κάποιο όρο της οντολογίας (π.χ. στιγμιότυπο, γενίκευση, ειδίκευση) και κάποιο βαθμό ισχύος. Ο βαθμός ισχύος καθορίζει το πόσο σίγουροι είμαστε για τη σχέση αυτή και μπορεί να μας βοηθήσει να αποδεχθούμε ή να απορρίψουμε ένα κανόνα βάζοντας κάποιο κατώφλι.

Συνεχίζοντας το παράδειγμα της προηγούμενης ενότητας, ο κανόνας κ5 που προέκυψε εξετάζεται ως προς την υποστήριξη που έχει χωρίς τον περιορισμό στην τοποθεσία. Στην περίπτωση που ο βαθμός υποστήριξης (περιπτώσεις που τον επιβεβαιώνουν/συνολικές περιπτώσεις όπως προκύπτουν από τον πράκτορα¹) είναι μεγάλος τότε μπορούμε τον κανόνα κ5 να τον εντάξουμε στην οντολογία (σχήμα 4β). Η διαδικασία αυτή μπορεί να γίνει ημιαυτόματα, εμφανίζοντας τους κανόνες και το βαθμό υποστήριξης και επιτρέποντας στο διαχειριστή της οντολογίας να την ενημερώσει. Μπορεί επίσης να γίνει αυτόματα αν ορίσουμε ένα ελάχιστο βαθμό υποστήριξης ώστε ένας κανόνας να απεικονίζεται ως σχέση στην οντολογία και μόνο για πολύ απλές περιπτώσεις. Συνήθως έχουμε πολύ σύνθετα προβλήματα να διαχειριστούμε όπως η διαχείριση των σχέσεων που έχει ο όρος προς εισαγωγή με άλλους όρους, αν θα εισαχθεί ως τάξη ή ως στιγμιότυπο στην οντολογία, με ποιους όρους θα συνδεθεί, με ποιες σχέσεις κλπ. Σημαντικό είναι επίσης στις περιπτώσεις αυτές να καθοριστεί και ένα μέτρο για την αξιολόγηση της καινοτομίας ενός κανόνα καθώς είναι πολύ πιθανό να προκύψουν κανόνες κοινά αποδεκτοί και γνωστοί που δεν έχουν καμία αξία. Στην περίπτωση των κανόνων συσχέτισης μεταξύ σύνθετων όρων έχουν προταθεί τεχνικές αξιολόγησης της σημαντικότητας των κανόνων με χρήση οντολογίας. Μια προσπάθεια αξιολόγησης της καινοτομίας ενός κανόνα με χρήση του wordnet. Η σημασιολογική απόσταση των δύο συνόλων μετρούμενη στην οντολογία είναι ένα μέτρο της καινοτομίας του κανόνα, όσο πιο μακριά βρίσκονται οι όροι που συσχετίζονται τόσο πιο καινοτόμος θεωρείται ο κανόνας. Και πάλι με αναφορά στο παράδειγμα ένας κανόνας που σχετίζει το γκολφ (golf) με ένα ξενοδοχείο (hotel) θεωρείται πιο καινοτομικός από ένα κανόνα που σχετίζει το γκολφ (golf) με το μπαστούνι του γκολφ (club) καθώς η απόσταση στο wordnet για το πρώτο ζευγάρι είναι

πολύ μεγαλύτερη. Στα πλαίσια της ερευνητικής εργασίας θα προσπαθήσουμε να καθορίσουμε τεχνικές για την αξιολόγηση της καινοτομίας ενός κανόνα συσχέτισης.

Οι συσχετίσεις μεταξύ όρων που προκύπτουν από την ανάλυση των κειμένων εντάσσονται στη συνέχεια σε μια μεγαλύτερη βάση γνώσης, συνήθως μια οντολογία. Στο σημείο αυτό χρησιμοποιούνται *reasoners* που εντοπίζουν τα σημεία στα οποία μπορούν να εισαχθούν οι σχέσεις (π.χ. ταιριάζοντας τους όρους με αυτούς της οντολογίας) και τα προτείνουν στους ειδικούς. Η γνώση συνεπώς που εξάγεται μπορεί να υπάρξει αυτόνομα χωρίς να συνδέεται με κάποια οντολογία, αν όμως αυτό χρειαστεί τότε απαιτείται και η ανθρώπινη παρέμβαση για να εξασφαλιστεί η ευθυγράμμιση της οντολογίας. Συνεχίζοντας στο ίδιο παράδειγμα, μπορούμε από ανάλυση πολλών ιστοσελίδων να βρούμε περιπτώσεις όπου πολλά ξενοδοχεία έχουν *spa* (η λέξη *spa* συνεμφανίζεται με το όνομα των ξενοδοχείων). Από τις περιπτώσεις αυτές (αν τα ονόματα ξενοδοχείων αντιστοιχούν σε *luxury_hotels*) προκύπτει ο κανόνας που λέει ότι τα *luxury_hotels* έχουν *spa*. Στην περίπτωση αυτή ο *reasoner* θα πρέπει να εξετάσει τον κανόνα και την οντολογία και να προτείνει αν η οντότητα *spa* θα μπει ως τιμή κάποιου χαρακτηριστικού του *luxury_hotel* ή ως ξεχωριστή τάξη στην οντολογία και επίσης με ποια τάξη θα συνδεθεί (απευθείας με το *luxury_hotel*, στα *facilities* ή κάτω από την τάξη *relaxation_activities*).

Στην ίδια κατηγορία περιλαμβάνονται και τα συστήματα από ημιδομημένα ή σχεσιακά σχήματα δεδομένων. Αυτά χρησιμοποιούν *Naive Bayesian* ή *Data correlation* τεχνικές και προσφέρουν ακόμη πιο πλούσια γνώση (π.χ. εξαγωγή του εννοιολογικού σχήματος από μια σχεσιακή ΒΔ, εξαγωγή σημασιολογικής περιγραφής για μια πηγή δεδομένων ή για μια υπηρεσία Ιστού). Στις περιπτώσεις αυτές δεδομένα μιας άγνωστης πηγής συγκρίνονται με αυτά μιας γνωστής πηγής του ίδιου πεδίου. Για παράδειγμα η περιγραφή του καιρού σε μια υπηρεσία περιέχει το όνομα της περιοχής και ένα ζεύγος συντεταγμένων, δύο θερμοκρασίες και ένα βαθμό έντασης ανέμων. Συγκρίνοντας με την περιγραφή που δίνει μια άλλη υπηρεσία (ταχ. κώδικας, όνομα περιοχής, μέση θερμοκρασία, ζεύγος ελάχιστης-μέγιστης

θερμοκρασίας, ένταση ανέμων) παίρνουμε μια αντιστοίχιση των περιγραφών που υποδεικνύει ότι οι συντεταγμένες είναι προσδιοριστικό της περιοχής (location), η μικρότερη θερμοκρασία αντιστοιχεί στην ελάχιστη και η άλλη στη μέγιστη θερμοκρασία και ο βαθμός έντασης στην ένταση του ανέμου. Η γνώση αυτή είναι αυστηρά συνδεδεμένη με την αρχική πηγή και δεν μπορεί να γενικευθεί.

Σε πιο απλές περιπτώσεις (π.χ. ανάλυση συναλλαγών) προκύπτουν κανόνες συσχέτισης που συνδέουν στιγμιότυπα (π.χ. τσιπς και μπίρες ή φιστίκια και ούισκι) και όχι γενικότερες έννοιες (σνακ και ποτά). Αυτό είναι ένα σημείο στο οποίο θα εστιάσει η παρούσα ερευνητική εργασία.

ΚΕΦΑΛΑΙΟ 3

Μελέτη περίπτωσης : Ανάπτυξη Συστήματος Αεροπορικής Εταιρείας

3.1. Εισαγωγή

Για την καλύτερη κατανόηση της υπάρχουσας εργασίας, θα δείξουμε με ένα παράδειγμα την λειτουργία της γλώσσας OWL. Πρώτα θα γίνει ανάλυση του προβλήματος που αναφέρεται στο παράδειγμα μας, μέσω του διαγράμματος ER.

Η air airlines είναι μια νέα αεροπορική εταιρεία που δραστηριοποιείται σε Ελλάδα και εξωτερικό. Η διοίκηση της air airlines, στα πλαίσια των ευρύτερων προσπαθειών της για την εδραίωσή της στο κλάδο των αερομεταφορών, αποφάσισε την εφαρμογή ενός ολοκληρωμένου πληροφοριακού συστήματος.

Η εταιρεία έχει ξεκινήσει εντατική προσπάθεια εισαγωγής της πληροφορικής σε όλες τις επιμέρους λειτουργίες της. Ο βασικός στόχος της air airlines που επιδιώκεται να επιτευχθεί μέσω της υλοποίησης του Ολοκληρωμένου Πληροφοριακού Συστήματος είναι η εύκολη πρόσβαση στις υπηρεσίες που θα προφέρει

Αναλυτικά, έχει να κάνει με το σχεδιασμό ενός πληροφοριακού συστήματος για τις ημερομηνίες και ώρες πτήσεων, τη διαθεσιμότητα και κράτηση θέσεων, το κόστος εισιτηρίων, τις προμήθειες πρακτόρων, τον τρόπο πληρωμής, ειδικές πληροφορίες για τη πτήση (γεύματα, καιρός κ.τ.λ.), επιβεβαίωση και ακύρωση κράτησης και εκτύπωση εισιτηρίων

Ειδικότερα να μπορεί εύκολα ο χρήστης, είτε ο πράκτορας είτε ο πελάτης μέσω internet, να μπορεί να πάρει πληροφορίες για τη διαθεσιμότητα των θέσεων στην πιθανή ημερομηνία που θέλει να ταξιδέψει ο πελάτης, το είδος των θέσεων (θέση προσφοράς, οικονομική ή θέση business) και η επιλογή θέσης (παράθυρο ή εσωτερικός διάδρομος).

Σε ο,τι αφορά το κόστος να υπάρχει ενημέρωση του χρήστη για τις τιμές των εισιτηρίων, τις εκπτώσεις ή τις επιβαρύνσεις στην τιμή ανάλογα με το είδος της θέσης και του εισιτηρίου (απλό ή με επιστροφή) καθώς και την επιβάρυνση από τους φόρους αεροδρομίου. Επιπλέον να γνωστοποιούνται τα

στοιχεία του πράκτορα και το ποσοστό που του αντιστοιχεί από την κράτηση. Επίσης να αναφέρεται η έκπτωση που θα υπάρχει αν η κράτηση γίνεται από το internet (δηλαδή η πληρωμή θα γίνεται μέσω πιστωτικής) ή μέσω κάποιου πρακτορείου μετρητοίς χωρίς τη συγκεκριμένη έκπτωση. Ακόμα για να γίνει η κράτηση πρέπει να δίνονται τα στοιχεία του πελάτη(ονοματεπώνυμο, αριθμός ταυτότητας κ.τ.λ). Επιπλέον , σε περίπτωση ακύρωσης του εισιτηρίου έως μια συγκεκριμένη χρονική στιγμή θα υπάρχει κάποια επιβάρυνσή σε αναλογία με την τιμή για τον πελάτη.

Επίσης, να μπορεί να χρησιμοποιείται από αρκετά μεγάλο αριθμό χρηστών, οι οποίοι θα είναι on line συνδεδεμένοι με το κεντρικό σύστημα της επιχείρησης, να μπορούν να έχουν πρόσβαση οποιαδήποτε χρονική στιγμή και σε σύντομο χρονικό διάστημα.

3.2.Περιγραφή της αεροπορικής εταιρίας σε ER διάγραμμα

Θα κάνουμε πρώτα ανάλυση του προβλήματος μας μέσω ενός διαγράμματος ER. Θα αναλύσουμε τις οντότητες από τις οποίες αποτελείται το διάγραμμα καθώς και τις συσχετίσεις μεταξύ τους.

ΔΙΑΓΡΑΜΜΑ ER

Οντότητες :

- Θέση (οικονομική, προσφορά,business)
- Εισιτήριο (απλό, κωδικό, επιστροφή)
- Πτήση (κωδικός, ώρα, ημέρα)
- Πελάτης (αριθμός ταυτότητας, όνομα, επώνυμο)
- Πράκτορας (κωδικός, όνομα, επώνυμο)

Συσχετίσεις:

1. Αντιστοιχεί (εισιτήριο ,θέση)
2. Διαθεσιμότητα (θέση , πτήση)
3. Κράτηση (εισιτήριο , πελάτης)
- 4.Πληρωμή (πράκτορας , πελάτη)

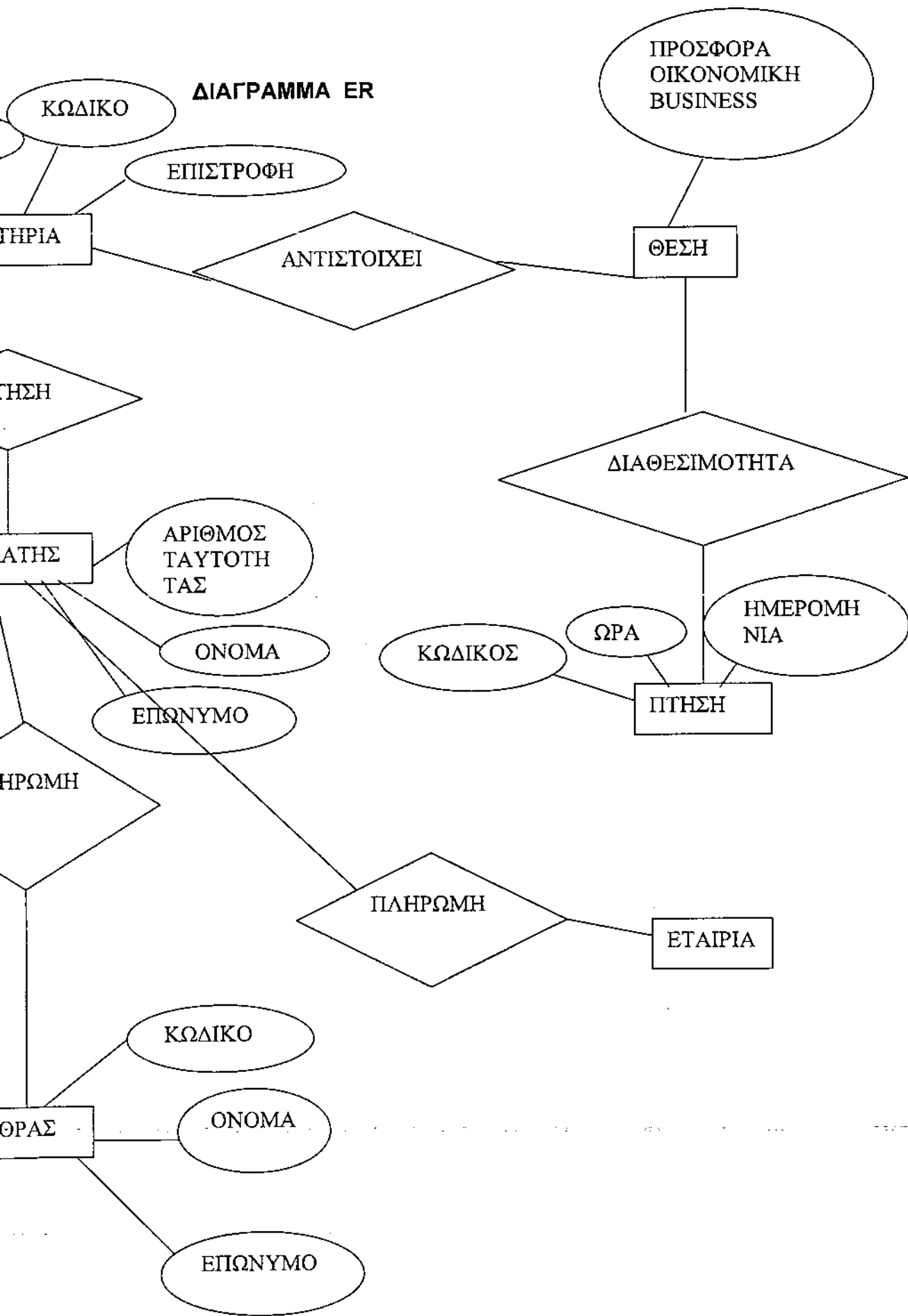
ΛΕΚΤΙΚΗ ΠΕΡΙΓΡΑΦΗ

Αντιστοιχεί :Κάθε εισιτήριο αντιστοιχεί μόνο σε μία θέση και κάθε θέση αντιστοιχεί μόνο σε ένα εισιτήριο.

Διαθεσιμότητα : Κάθε πτήση έχει διαθέσιμες πολλές θέσεις και κάθε θέση είναι διαθέσιμη μόνο σε μία πτήση

Κράτηση : Κάθε πελάτης μπορεί να κάνει κράτηση σε πολλά εισιτήρια αλλά κάθε εισιτήριο μπορεί να αντιστοιχεί σε ένα μόνο πελάτη.

Πληρωμή : Κάθε πράκτορας μπορεί να πληρώνεται από πολλούς πελάτες αλλά κάθε πελάτης μπορεί να πληρώνει το πολύ έναν πράκτορα.



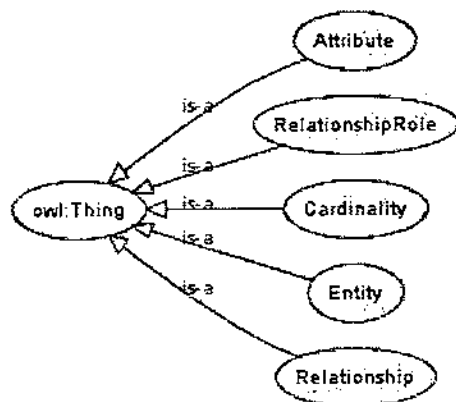
3.3. Ανάπτυξη οντολογίας για το διάγραμμα Οντοτήτων Σχέσεων (ER)

Το πρότυπο σχέσης και χρόνου του διαγράμματος ER περιλαμβάνει τις κατηγορίες αντικειμένων και τις κατηγορίες σχέσης. Αυτό το πρότυπο επεκτείνεται έτσι ώστε οι τρέχουσες απαιτήσεις, στην επιχειρησιακή διαμόρφωση, να μπορούν να ενσωματωθούν. Ο βασικός μηχανισμός ER, καθορίζει τους παραγώγους που εμφανίζονται μέσα στην εφαρμογή.

Πιο κάτω βλέπουμε πως συσχετίζονται μεταξύ τους οι κλάσεις και τι δηλώνει η κάθε μια. Οι κλάσεις μας είναι οι εξής:

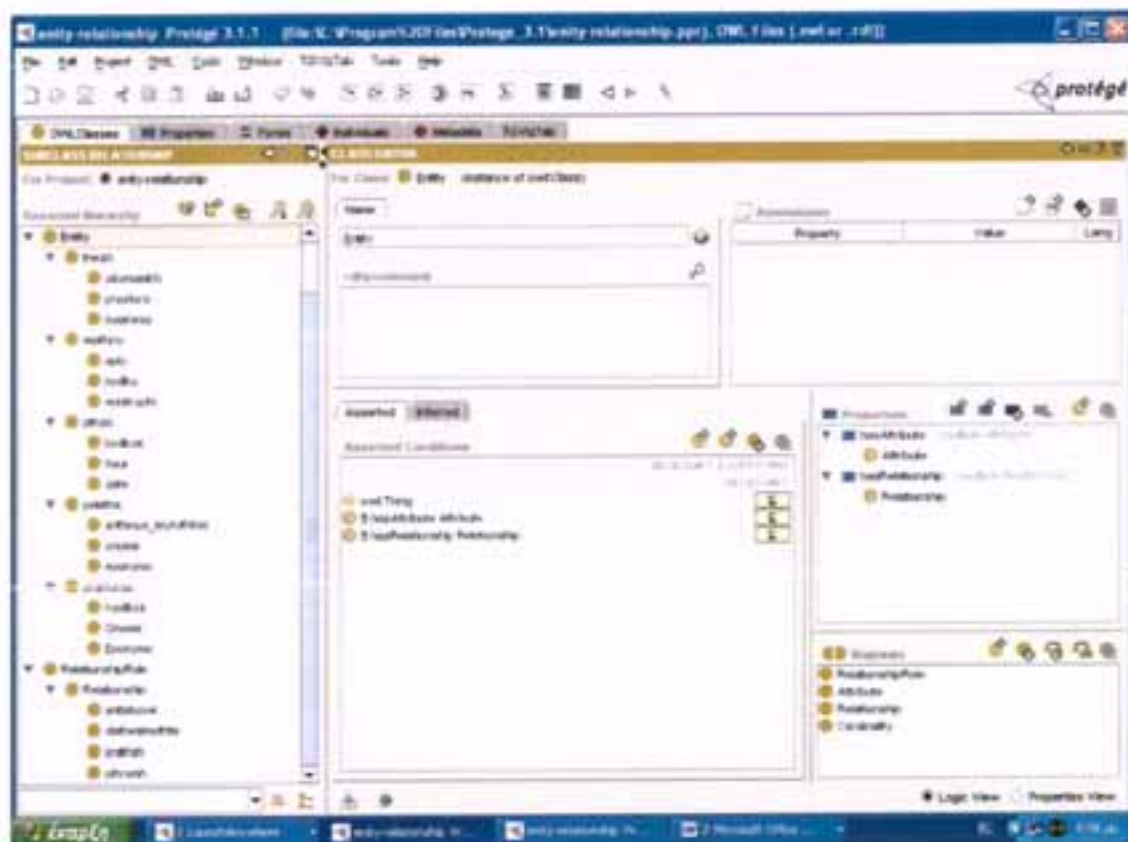
- Entity: όπου περιλαμβάνει τις οντότητες που αναφέρονται σε κάθε πρόβλημα
- Attribute: όπου αντιπροσωπεύει το όνομα κάθε σχέσης
- Relationship: όπου αναφέρει την συσχέτιση μεταξύ δύο οντοτήτων
- Cardinality: όπου μας δείχνει τον τύπο της κάθε σχέσης
- RelationshipRole: όπου περιλαμβάνει το όνομα της συσχέτισης

Το διάγραμμα της γλώσσας OWL είναι ως εξής :



Σχήμα 3.1

Μετά την εισαγωγή των στοιχείων στο πρόγραμμα Protégé φαίνονται οι κλάσεις.



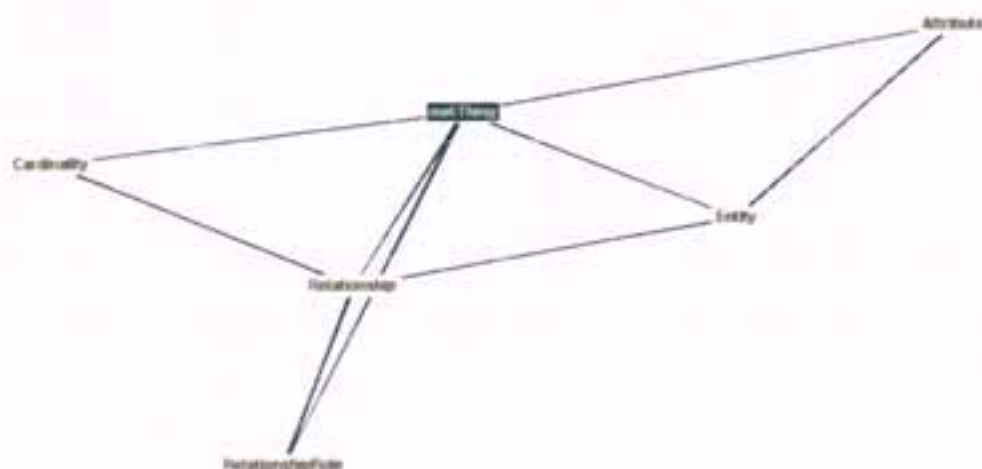
Σχήμα 3.2

Οι σχέσεις μεταξύ των κλάσεων περιγράφονται παρακάτω ως εξής:

- Has Attribute – is AttributeOf
Domain_Entity
Range_Attribute
- Has Cardinality – is CardinalityOf
Domain_Relationship
Range_Cardinality

- Has Relationship – is RelationshipOf
Domain_Entity
Range_Relationship
- Has Role – is RoleOf
Domain_relationship
Range_RelationshipRole
- Is AttributeOf – has Attribute
Domain_Attribute
Range_Entity
- Is CardinalityOf – has Cardinality
Domain_Cardinality
Range_Relationship
- Is RelationshipOf – has Relationship
Domain_Relationship
Range_Entity
- Is RoleOf – has Role
Domain_RelationshipRole
Range_Relationship

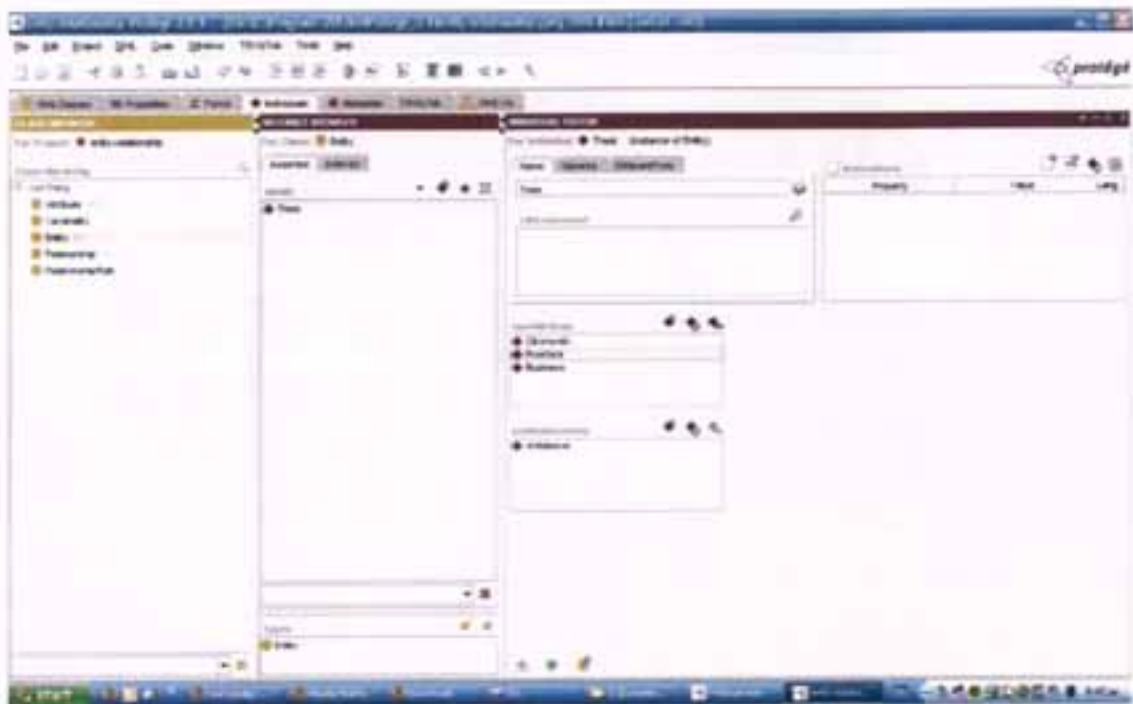
Στο παρακάτω σχήμα αναλύονται οι κλάσεις μεταξύ τους και βλέπουμε τις σχέσεις μεταξύ τους.



Σχήμα 3.3

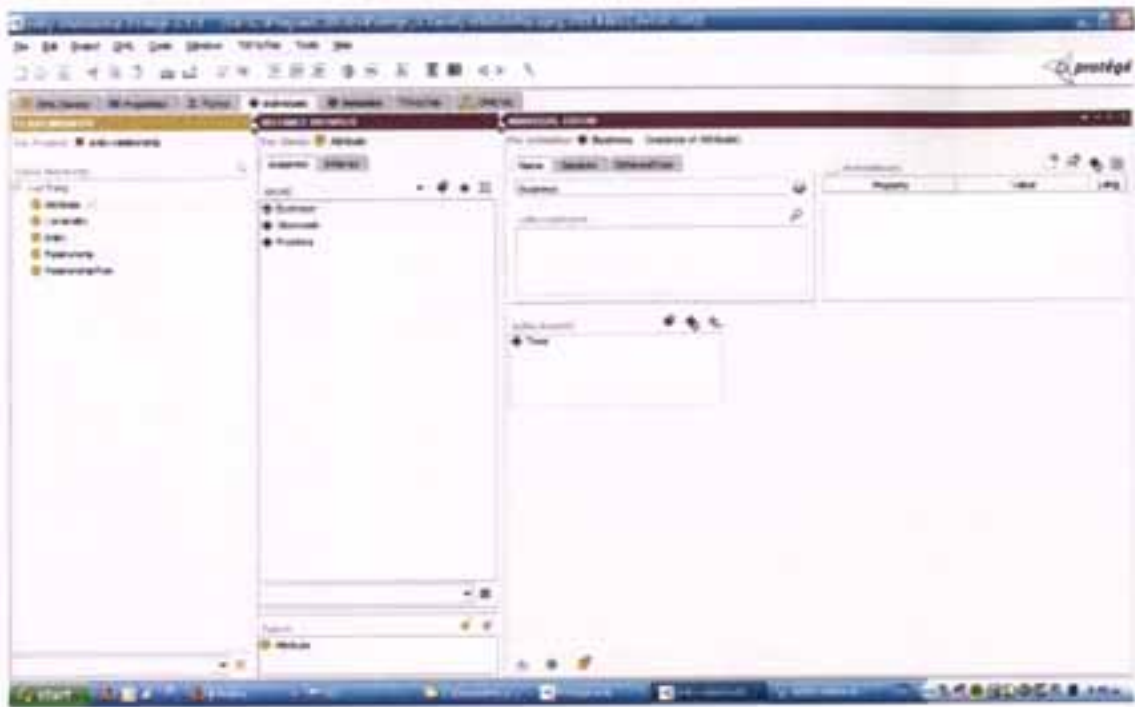
3.4. Περιγραφή της αεροπορικής εταιρίας σε OWL

Παραπάνω έγινε απεικόνιση του ER διαγράμματος . Στην τελευταία φάση της εργασίας μας βλέπουμε την αναπαράσταση του διαγράμματος ER σε γλώσσα προγραμματισμού, την OWL, για την αεροπορική εταιρία air airlines που περιγράψαμε παραπάνω. Πρώτα θα περιγράψουμε τα individuals και τα properties των οντοτήτων (entity) που φαίνονται ως εξής :



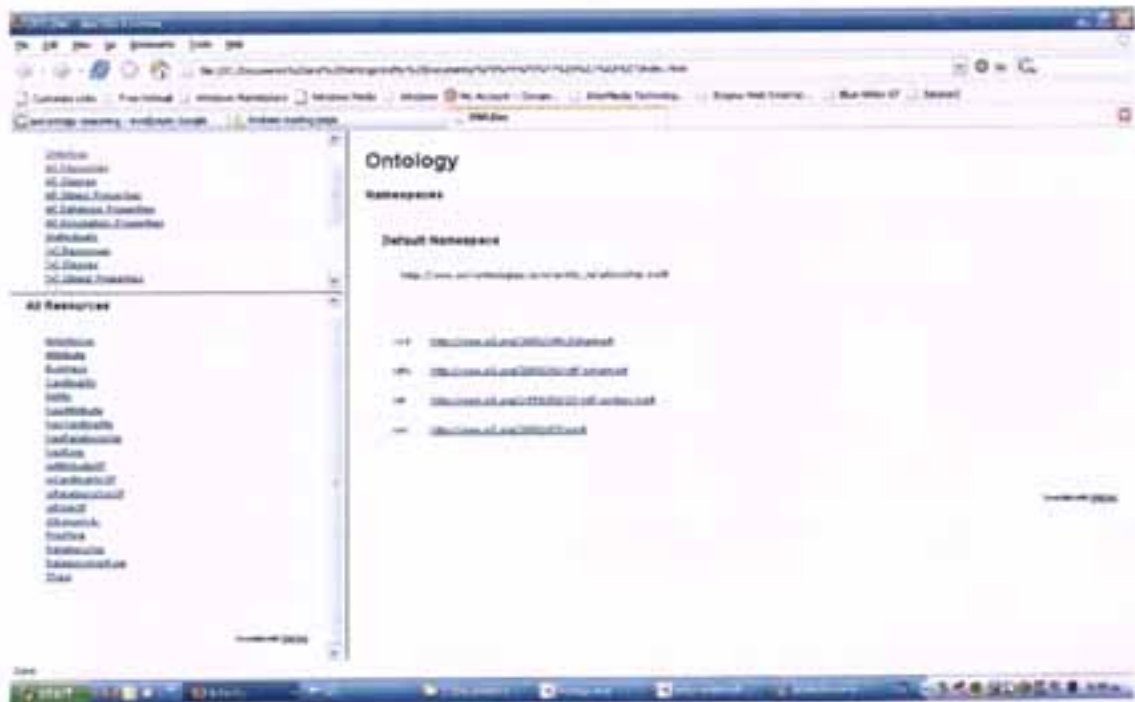
Σχήμα 3.4

Δηλαδή σ' αυτό το διάγραμμα βλέπουμε ότι για τα entities περνάμε την θέση και στα individual προσθέτουμε οικονομική, προσφορά και business. Στα properties περνάμε για παράδειγμα ότι μία θέση αντιστοιχεί σε ένα εισιτήριο. Το ίδιο ισχύει και για τα attribute όπως φαίνεται στο παρακάτω σχήμα :



Σχήμα 3.5

Την οντολογία που περιγράψαμε μπορούμε να την εξηγήσουμε μέσα από ένα web browser ως εξής:



Τέλος ένα παράδειγμα της περιγραφής OWL είναι ο παρακάτω κώδικας :

```
<Rdf: RDF
  Xmlns: rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  Xmlns="http://www.owl-ontologies.com/entity_relationship.owl#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
xml:base="http://www.owl-ontologies.com/entity_relationship.owl">
<owl:Ontology rdf:about=""/>
<owl:Class rdf:ID="Attribute">
  <owl:disjointWith>
    <owl:Class rdf:ID="Cardinality"/>
  </owl:disjointWith>
  <owl:disjointWith>
    <owl:Class rdf:ID="Entity"/>
  </owl:disjointWith>
  <owl:disjointWith>
    <owl:Class rdf:ID="Relationship"/>
  </owl:disjointWith>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <owl:disjointWith>
    <owl:Class rdf:ID="RelationshipRole"/>
  </owl:disjointWith>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:allValuesFrom>
        <owl:Class rdf:about="#Entity"/>
      </owl:allValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="isAttributeOf"/>
      </owl:onProperty>
    </owl:Restriction>
```

```

</rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:about="#RelationshipRole">
  <owl:disjointWith>
    <owl:Class rdf:about="#Relationship"/>
  </owl:disjointWith>
  <owl:disjointWith>
    <owl:Class rdf:about="#Cardinality"/>
  </owl:disjointWith>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:allValuesFrom>
        <owl:Class rdf:about="#Relationship"/>
      </owl:allValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="isCardinalityOf"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <owl:disjointWith rdf:resource="#Attribute"/>
  <owl:disjointWith>
    <owl:Class rdf:about="#Entity"/>
  </owl:disjointWith>
</owl:Class>
<owl:Class rdf:about="#Relationship">
  <owl:disjointWith rdf:resource="#Attribute"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="isRelationshipOf"/>
      </owl:onProperty>
      <owl:allValuesFrom>
        <owl:Class rdf:about="#Entity"/>
      </owl:allValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>

```

```

    </owl:allValuesFrom>
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
<owl:disjointWith rdf:resource="#RelationshipRole"/>
<owl:disjointWith>
  <owl:Class rdf:about="#Cardinality"/>
</owl:disjointWith>
<owl:disjointWith>
  <owl:Class rdf:about="#Entity"/>
</owl:disjointWith>
</owl:Class>
<owl:Class rdf:about="#Entity">
  <owl:disjointWith rdf:resource="#RelationshipRole"/>
  <owl:disjointWith rdf:resource="#Relationship"/>
  <owl:disjointWith rdf:resource="#Attribute"/>
  <owl:disjointWith>
    <owl:Class rdf:about="#Cardinality"/>
  </owl:disjointWith>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="hasRelationship"/>
      </owl:onProperty>
      <owl:someValuesFrom rdf:resource="#Relationship"/>
    </owl:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:someValuesFrom rdf:resource="#Attribute"/>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="hasAttribute"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>

```

```

    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
<owl:Class rdf:about="#Cardinality">
  <owl:disjointWith rdf:resource="#Entity"/>
  <owl:disjointWith rdf:resource="#RelationshipRole"/>
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:allValuesFrom rdf:resource="#Relationship"/>
      <owl:onProperty>
        <owl:ObjectProperty rdf:about="#isCardinalityOf"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
  <owl:disjointWith rdf:resource="#Relationship"/>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>
  <owl:disjointWith rdf:resource="#Attribute"/>
</owl:Class>
<owl:ObjectProperty rdf:ID="hasCardinality">
  <rdfs:range rdf:resource="#Cardinality"/>
  <rdfs:domain rdf:resource="#Relationship"/>
  <owl:inverseOf>
    <owl:ObjectProperty rdf:about="#isCardinalityOf"/>
  </owl:inverseOf>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#hasAttribute">
  <owl:inverseOf>
    <owl:ObjectProperty rdf:about="#isAttributeOf"/>
  </owl:inverseOf>
  <rdfs:range rdf:resource="#Attribute"/>
  <rdfs:domain rdf:resource="#Entity"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="isRoleOf">
  <owl:inverseOf>

```

```

    <owl:ObjectProperty rdf:ID="hasRole"/>
</owl:inverseOf>
<rdfs:range rdf:resource="#Relationship"/>
<rdfs:domain rdf:resource="#RelationshipRole"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#hasRelationship">
  <rdfs:domain rdf:resource="#Entity"/>
  <rdfs:range rdf:resource="#Relationship"/>
  <owl:inverseOf>
    <owl:ObjectProperty rdf:about="#isRelationshipOf"/>
  </owl:inverseOf>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#isAttributeOf">
  <owl:inverseOf rdf:resource="#hasAttribute"/>
  <rdfs:range rdf:resource="#Entity"/>
  <rdfs:domain rdf:resource="#Attribute"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#isCardinalityOf">
  <rdfs:range rdf:resource="#Relationship"/>
  <owl:inverseOf rdf:resource="#hasCardinality"/>
  <rdfs:domain rdf:resource="#Cardinality"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#hasRole">
  <rdfs:range rdf:resource="#RelationshipRole"/>
  <rdfs:domain rdf:resource="#Relationship"/>
  <owl:inverseOf rdf:resource="#isRoleOf"/>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#isRelationshipOf">
  <owl:inverseOf rdf:resource="#hasRelationship"/>
  <rdfs:domain rdf:resource="#Relationship"/>
  <rdfs:range rdf:resource="#Entity"/>
</owl:ObjectProperty>
</rdf:RDF>

```


<!-- Created with Protege (with OWL Plugin 2.1, Build 284)
http://protege.stanford.edu -->

ΚΕΦΑΛΑΙΟ 4

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στον ρυθμό παραγωγής πληροφοριών , παρατηρείται μια αλματώδη αύξηση, που είναι αποτέλεσμα των εξελίξεων σε διάφορα επιστημονικά πεδία αλλά και των απαιτήσεων των σύγχρονων εταιριών για αυξημένη παραγωγικότητα και αποτελεσματικότητα. Αποτέλεσμα αυτής της πληροφοριακής έκρηξης είναι το γεγονός ότι η εξαγωγή γνώσης από την πληροφορία είναι εξίσου βαρυσήμαντη όσο και η ίδια η πληροφορία.

Η οντολογική προσέγγιση για την υποστήριξη πληροφοριακών συστημάτων έχει ήδη παρουσιάσει ενθαρρυντικά αποτελέσματα μέσω της κωδικοποίησης και εκμετάλλευσης της σημασιολογίας της πληροφορίας. Τα σημαντικότερα προβλήματα που παρουσιάζονται σε αυτήν την προσέγγιση αφορούν την δυσκολία της διαχείρισης των μετα-δεδομένων και των οντολογιών. Ένα άλλο σημαντικό πρόβλημα είναι η απουσία ενιαίας σημασιολογικής κωδικοποίησης αφού κάθε πεδίο (ή κάθε οργανισμός) αντιλαμβάνεται με διαφορετικό τρόπο την πληροφορία.

Οι περισσότερες προσεγγίσεις στην κατασκευή πληροφοριακών συστημάτων δεν έχουν επικεντρώσει την προσοχή τους σε παράγοντες ποιότητας στο θέμα της ζήτησης /ανάκτησης πληροφορίας, κυρίως γιατί τα κριτήρια απόδοσης αποτελούσαν το κύριο μέλημα των σχεδιαστών. Μία πρώτη προσέγγιση για την λύση του προβλήματος της ποιοτικής αναζήτησης / ανάκτησης πληροφορίας αποτελεί η ανάκτηση πληροφορίας με βάση το περιεχόμενο (content based retrieval). Η προσέγγιση αυτή στηρίζεται στη ανάλυση του περιεχομένου της πληροφορίας μέσω στατικών ή και υπολογιστικών μεθόδων για την επίτευξη όσο το δυνατόν καλύτερων αποτελεσμάτων. Παρόλη την πρόοδο που έχει επιτευχθεί στον τομέα της αναζήτησης /ανάκτησης πληροφορίας με βάση το περιεχόμενο, τα αποτελέσματα έχουν δείξει ότι υπάρχουν σημαντικοί περιορισμοί σε αυτήν την προσέγγιση. Αυτό οφείλεται στο γεγονός ότι η συνολική πληροφορία δεν περιορίζεται μόνο στο περιεχόμενο ή στους χρήστες αλλά αποτελείται και από

διαδικασίες και σχέσεις που διέπουν την ίδια την πληροφορία. Η δυναμική που έχει αναπτυχθεί επιβεβαιώνει την σημασία των οντολογιών για τα σύγχρονα πληροφοριακά συστήματα. Ειδικά στο πεδίο της ανάκτησης πληροφορίας, η χρήση όλης της διαθέσιμης πληροφορίας (περιεχόμενο και γνώση) έχει ήδη δώσει πολύ ενθαρρυντικά αποτελέσματα αλλά έχει παρουσιάσει και νέα προβλήματα.

Ο Ιστός (Web) του μέλλοντος, όπως αποκαλούμε τον Σημασιολογικό Ιστό, προβλέπεται να αποτελεί μια παγκόσμια βάση δεδομένων και γνώσης με πληροφορίες οι οποίες θα είναι "κατανοητές" από μηχανές (machine – understandable information). Οι κύριες τεχνολογίες για την υλοποίηση του Σημασιολογικού Ιστού είναι ο σημασιολογικός εμπλουτισμός και η χρήση των οντολογιών. Επίσης μπορούμε να πούμε ότι ο Ιστός αποτελεί το μεγαλύτερο σε παγκόσμιο επίπεδο έργο ευφυής ενσωμάτωσης συστημάτων ώστε να συνεργάζονται δια-λειτουργικά.

Ο σημασιολογικός χαρακτήρας της πληροφορίας και η χρήση οντολογιών για την αναπαράσταση της είναι σχετικά ένας καινούργιος τομέας στην περιοχή της διαχείρισης γνώσης. Η διαδικασία της μετατροπής της πληροφορίας σε σημασιολογική είναι δυνατή μέσω ενός πλαισίου περιγραφής, του RDF (Resource Description Framework) και των RDF schemas. Το World Wide Web Consortium (W3C) έχει δημιουργήσει μία ομάδα για την δημιουργία οντολογιών χρησιμοποιώντας το πλαίσιο αυτό και να ενσωματώσουν τα πρόσθετα, καθορισμένα από την χρήση, πρότυπα. Κατά αυτόν τον τρόπο, οποιοσδήποτε τύπος πληροφοριών απαίτησης μπορεί να συμπεριληφθεί στο W3C.

Η ανάπτυξη πληροφοριακών συστημάτων με οντολογικές προσεγγίσεις είναι μία νέα περιοχή έρευνας η οποία όμως βρίσκεται ακόμα σε πρώιμο ερευνητικό στάδιο. Στην πτυχιακή εργασία αυτή προσπαθήσαμε την εφαρμογή μιας τέτοιας οντολογικής ανάπτυξης με την μελέτη περίπτωσης μιας αεροπορικής εταιρείας. Στο παράδειγμα αυτό απλά μελετήθηκε η διαδικασία ανάπτυξης και η κωδικοποίηση στη γλώσσα owl. Για να φανούν τα οφέλη από μια τέτοια διαδικασία χρειάζεται η ανάπτυξη περισσότερων οντολογιών μεθόδων ανάπτυξης όπως π.χ. διαγράμματα ροής δεδομένων και η μελέτη και άλλων περιπτώσεων χρήσης ώστε να γίνει η δημιουργία μιας

ευρύτερης οντολογίας η οποία να ενσωματώνει ολόκληρη τη γνώση ανάπτυξης. Κατ' αυτόν τον τρόπο θα έχουμε την δυνατότητα να ερευνούμε τη σχετική γνώση (reasoning) μέσα από μία ενιαία οντολογία. Με άλλα λόγια θα ξεπεράσουμε τα εμπόδια χρήσης διαφορετικών μεθοδολογιών και προτύπων αλλά θα βασιζόμαστε στην αποδοχή του πρότυπου owl. Επιπλέον η δυνατότητα ενσωμάτωσης τέτοιων οντολογιών στον σημασιολογικό ιστό, ιδιαίτερα με τη χρήση υπηρεσιών owl-s θα μπορούσε να δώσει μια νέα διάσταση στα εργαλεία ανάπτυξης συστημάτων και στη πρόσβαση της γνώσης που απαιτείται για την ανάπτυξη αυτή.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Ora Lassila and Ralph Swick. Resource Description Framework (RDF). Model and Syntax Srecification W3C recommendation, feb 1999.
<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>
2. Dan Brickley and R.V.Guha.Resource Description Framework (RDF) Schema Specification. W3C <http://www.w3.org/TR/1999/PR-rdf-schema-19990303>
3. OWL Web Ontology Language Guide, <http://www.w3.org/TR/OWL-guide/>
4. KAON: The Karlsruhe Ontology and Semantic Web tool suite.
<http://kaon.semanticweb.org/>
5. DQL, <http://www.daml.org/dq1/DQL>
6. DMOZ-Open Directory Project <http://www.dmoz.org>