

ΔΙΑΔΙΚΤΥΟ & ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ (XML – P2P SYSTEMS - ONTOLOGIES)

GoToggle: ΕΞΥΠΙΝΗ ΜΗΧΑΝΗ ΑΝΑΖΗΤΗΣΗΣ ΜΕ ΧΡΗΣΗ ΟΝΤΟΛΟΓΙΩΝ	2
ΜΟΝΤΕΛΟ ΟΡΓΑΝΩΣΗΣ ΚΑΙ ΑΠΟΘΗΚΕΥΣΗΣ ΣΗΜΑΣΙΟΛΟΓΙΚΩΝ ΣΧΗΜΑΤΩΝ ΤΟΥ ΠΑΓΚΟΣΜΙΟΥ ΙΣΤΟΥ ΠΛΗΡΟΦΟΡΙΩΝ	4
V: ΔΙΑΔΙΚΤΥΑΚΗ ΥΠΗΡΕΣΙΑ ΙΣΤΟΥ ΣΥΝΑΛΛΑΓΗΣ ΑΝΤΙΚΕΙΜΕΝΩΝ ΚΑΙ ΥΠΗΡΕΣΙΩΝ	5
ΥΛΟΠΟΙΗΣΗ ΠΡΑΚΤΟΡΩΝ (Agents) ΓΙΑ ΔΙΚΤΥΟ ΚΙΝΟΥΜΕΝΩΝ ΟΜΟΤΙΜΩΝ ΒΑΣΕΩΝ (Peer-2-Peer Mobile Databases).....	6
ΥΛΟΠΟΙΗΣΗ ΜΗΧΑΝΙΣΜΟΥ ΕΡΩΤΑΠΟΚΡΙΣΕΩΝ ΓΙΑ ΔΙΚΤΥΟ ΚΙΝΟΥΜΕΝΩΝ ΟΜΟΤΙΜΩΝ ΒΑΣΕΩΝ (Peer-2-Peer Databases).....	7
ΠΡΟΤΥΠΟ ΣΥΣΤΗΜΑ ΟΛΟΚΛΗΡΩΣΗΣ ΣΧΗΜΑΤΩΝ ΜΕ ΥΠΟΣΤΗΡΙΞΗ ΟΝΤΟΛΟΓΙΩΝ ΓΙΑ ΚΑΤΑΝΕΜΗΜΕΝΑ ΠΕΡΙΒΑΛΛΟΝΤΑ ΑΥΤΟΝΟΜΩΝ ΚΟΜΒΩΝ .	8
Partialist: ΣΥΣΤΗΜΑ ΕΠΕΞΕΡΓΑΣΙΑΣ ΕΡΩΤΗΣΕΩΝ ΓΙΑ ΔΕΝΤΡΙΚΑ ΔΕΔΟΜΕΝΑ.....	10
ΥΛΟΠΟΙΗΣΗ ΨΗΦΙΑΚΗΣ ΒΑΣΗΣ ΦΩΤΟΕΡΜΗΝΕΥΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΣΤΟ ΔΙΑΔΙΚΤΥΟ	12
ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΥΛΟΠΟΙΗΣΗ ΔΙΑΔΙΚΤΥΑΚΗΣ ΕΦΑΡΜΟΓΗΣ ΓΙΑ ΤΗΝ ΥΠΟΒΟΛΗ ΔΗΛΩΣΕΩΝ ΙΔΙΟΚΤΗΣΙΑΣ.....	13

GoNToggle: ΕΞΥΠΝΗ ΜΗΧΑΝΗ ΑΝΑΖΗΤΗΣΗΣ ΜΕ ΧΡΗΣΗ ΟΝΤΟΛΟΓΙΩΝ

ΠΛΗΡΟΦΟΡΙΕΣ: Α. Δημητρίου, 210 7721402, angela@dblab.ntua.gr, Θ. Δαλαμάγκας, 210 7721402, dalamag@dblab.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στον σχεδιασμό μιας έξυπνης μηχανής αναζήτησης που θα χρησιμοποιεί οντολογίες. Η μηχανή θα μπορεί να χρησιμοποιηθεί για την αναζήτηση κειμένων διαφόρων μορφών (π.χ. doc, ppt, pdf, txt, ps, κ.λ.π.) σε προσωπικούς υπολογιστές, συλλογές κειμένων εταιρειών, κ.λ.π., που διασυνδέονται μέσω ενός δικτύου ομοτίμων (P2P)

ΑΤΟΜΑ: 2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Java

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:

Οι μηχανές αναζήτησης είναι δημοφιλές εργαλείο εύρεσης πληροφοριών στον Ιστό. Αρκούν κάποιες λέξεις κλειδιά π.χ. στο διάσημο Google για να βρεί κάποιος ιστοσελίδες και αρχεία σχετικά με το θέμα του. Ειδικότερα, το scholar google ειδικεύεται στην αναζήτηση κειμένων χρήσιμων σε σπουδαστές (π.χ. επιστημονικές εργασίες, αναφορές, κ.λ.π.). Η τεχνολογία της μηχανής αναζήτησης Google, μπορεί να χρησιμοποιηθεί ήδη και ως αυτόνομο πρόγραμμα αναζήτησης αρχείων σε προσωπικούς υπολογιστές (<http://desktop.google.com/>). Τα προγράμματα αυτά λέγονται *desktop search engines*.

Τα δίκτυα ομοτίμων (Peer-2-Peer networks) είναι δίκτυα αυτόνομων κόμβων (peers) δικτύου που επικοινωνούν μεταξύ τους για ανταλλαγή πληροφορίας. Οι κόμβοι είναι ελεύθεροι να συνδέονται και να αποσυνδέονται στο δίκτυο κατά βούληση. Σε καθέναν υπάρχει καταχωρημένη πληροφορία η οποία είναι διαθέσιμη στο δίκτυο από τη στιγμή της σύνδεσης του κόμβου μέχρι την αποσύνδεσή του. Η πληροφορία αυτή έχει μία δομή, που πρέπει ο κάθε κόμβος να διαφημίσει στους γειτονικούς του, έτσι ώστε εκείνοι να είναι σε θέση να την αντλήσουν διατυπώνοντας κατάλληλα ερωτήματα. Τα ερωτήματα ταξιδεύουν μέσα στο δίκτυο και όλοι μαζί οι κόμβοι που διατηρούν σχετική πληροφορία, τη στέλνουν πίσω στον κόμβο που έθεσε την ερώτηση ως μία συνολική απάντηση από το δίκτυο.

Η διπλωματική εργασία στοχεύει στη βελτίωση και επέκταση του *desktop search engine GoNTogle*. Το *GoNTogle* είναι μία μηχανή αναζήτησης σε προσωπικό υπολογιστή. Χρησιμοποιεί οντολογίες για να δώσει τη δυνατότητα στους χρήστες να χαρακτηρίσουν σημασιολογικά (semantic annotation) έγγραφα, που είναι αποθηκευμένα στον προσωπικό τους υπολογιστή, αλλά και για να παρέχει σύνθετες επιλογές αναζήτησης σε αυτά και όχι μόνο με χρήση λέξεων κλειδιών. Δύο είναι οι κύριοι σκοποί αυτής της διπλωματικής:

(α) Να βελτιωθεί ο μηχανισμός χαρακτηρισμού των εγγράφων του *GoNTogle*, έτσι ώστε να μπορεί να γίνεται και με αυτόματο τρόπο. Το σύστημα εκμεταλλευόμενο πληροφορίες όπως ο

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

τίτλος του εγγράφου, η θέση του στο δίσκο κ.α. αλλά και την παρουσία λέξεων κλειδιών εντός του κειμένου θα μπορεί να προτείνει μία ή περισσότερες έννοιες της οντολογίας για σημασιολογικό χαρακτηρισμό.

(β) Να εγκατασταθεί το GoNTogle σε δίκτυο ομοτίμων. Οι κόμβοι του δικτύου θα διαθέτουν ένα σύνολο από έγγραφα σημασιολογικά χαρακτηρισμένα. Οι χρήστες θα έχουν τη δυνατότητα να αναζητούν κάποιο έγγραφο στο δίκτυο συνολικά, διατυπώνοντας την ερώτηση αναζήτησης σε ένα μόνο κόμβο. Η αναζήτηση σε γειτονικούς κόμβους θα μπορεί να γίνεται είτε όταν οι κόμβοι χρησιμοποιούν όμοιες οντολογίες, είτε όταν χρησιμοποιούν διαφορετικές αλλά έχουν οριστέι αντιστοιχίσεις μεταξύ των οντοτήτων τους.

Εκτός από τα παραπάνω, το GoNTogle θα επεκταθεί, έτσι ώστε να δίνει τη δυνατότητα χαρακτηρισμού περισσότερων τύπων αρχείων, όπως email, εικόνων κ.α. Επίσης θα δοθούν δυνατότητες για πιο πλούσια σημασιολογικά ερωτήματα αναζήτησης.

Η διπλωματική αυτή επομένως, θα ασχοληθεί με τα παρακάτω θέματα:

1. Σχεδίαση και υλοποίηση αυτόματου χαρακτηρισμού εγγράφων.
2. Προσθήκη δυνατότητας χαρακτηρισμού νέων τύπων αρχείων στο GoNTogle.
3. Εμπλουτισμός δυνατοτήτων διατύπωσης ερωτημάτων αναζήτησης με νέα χαρακτηριστικά.
4. Επέκταση GoNTogle για επικοινωνία με ομότιμους απομακρυσμένους κόμβους.
 - a. Ανάπτυξη μηχανισμού διαφήμισης τοπικών οντολογιών
 - b. Προώθηση ερωτημάτων μεταξύ δύο γειτονικών κόμβων
 - c. Αντληση αρχείων γειτονικών κόμβων
 - d. Υλοποίηση χαρακτηρισμού εγγράφων γειτονικών κόμβων με τοπική οντολογία
 - e. Μηχανισμός αντιστοίχισης οντοτήτων δύο διαφορετικών οντολογιών (χειρονακτικά και αυτόματα)
 - f. Χρήση των τελευταίων για μετάφραση ερωτημάτων από τον ένα κόμβο στον άλλο

**ΜΟΝΤΕΛΟ ΟΡΓΑΝΩΣΗΣ ΚΑΙ ΑΠΟΘΗΚΕΥΣΗΣ ΣΗΜΑΣΙΟΛΟΓΙΚΩΝ ΣΧΗΜΑΤΩΝ
ΤΟΥ ΠΑΓΚΟΣΜΙΟΥ ΙΣΤΟΥ ΠΛΗΡΟΦΟΡΙΩΝ**

ΠΛΗΡΟΦΟΡΙΕΣ: Π. Μπούρος, 210 772 1446, pbour@dblab.ece.ntua.gr, Θ. Δαλαμάγκας, 210 772 1402, dalamag@dblab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στο σχεδιασμό και στην ανάπτυξη ενός μοντέλου αποθήκευσης σημασιολογικών σχημάτων του παγκόσμιου ιστού, που θα ευνοεί τη γρήγορη αποτίμηση σημασιολογικών ερωτημάτων.

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Java ή C/C++, XML, SQL

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Ο **Σημασιολογικός Ιστός** αποτελεί τη μετεξέλιξη του σημερινού Παγκόσμιου Ιστού όπου η διαδικτυακή πληροφορία θα είναι κατάλληλα οργανωμένη ώστε να μπορούν μηχανές να τη συλλέγουν, να την κατανοούν και να την επεξεργάζονται αυτόματα. Η διαδικασία αυτή βασίζεται στο **σημασιολογικό εμπλουτισμό** του περιεχομένου της πληροφορίας, δηλαδή τον ορισμό μεταπληροφορίας (πληροφορία για την πληροφορία). Κλειδί για το σημασιολογικό εμπλουτισμό των δεδομένων αποτελεί η κατασκευή και η διαχείριση **οντολογιών**, δηλαδή του λεξιλογίου περιγραφής των μεταδεδομένων.

Με δεδομένο τον τεράστιο όγκο πληροφορίας που οργανώνεται και διακινείται στον Παγκόσμιο Ιστό στην παρούσα μορφή του, είναι δεδομένο ότι για να υλοποιηθεί η ιδέα του Σημασιολογικού Ιστού πρέπει να αντιμετωπιστούν με αποδοτικό τρόπο ζητήματα **οργάνωσης, αποθήκευσης και ευρετηρίασης μεταδεδομένων**. Ταυτόχρονα, τα υπάρχοντα συστήματα **συλλογιστικής ανάλυσης** – reasoners λειτουργούν καλά αλλά όχι για μεγάλο όγκο δεδομένων και ούτε για όλους τους τύπους ερωτημάτων.

Στόχος της προτεινόμενης διπλωματικής είναι η κατασκευή ενός **συστήματος διαχείρισης σημασιολογικών σχημάτων Ιστού** με χρήση συστημάτων Βάσεων Δεδομένων. Ένα τέτοιο περιβάλλον πρέπει να μπορεί να συνεργάζεται με συστήματα συλλογιστικής ανάλυσης, τα οποία θα εμπλουτίζουν τις λειτουργίες ανάλυσης και αποτίμησης ερωτήσεων.

Η διπλωματική αυτή επομένως, θα ασχοληθεί με τα παρακάτω θέματα:

5. Μελέτη και εξοικείωση με γλώσσες περιγραφής σημασιολογικών σχημάτων και κυρίως OWL-DL.
6. Καταγραφή τύπων ερωτήσεων σε σημασιολογικά σχήματα του παγκόσμιου ιστού.
7. Μελέτη υπάρχοντων συστημάτων οργάνωσης και αποθήκευσης σημασιολογικών σχημάτων.
8. Σχεδίαση μοντέλου αποθήκευσης και ευρετηρίασης σημασιολογικών σχημάτων.
9. Υλοποίηση πρότυπου συστήματος που αποθηκεύει και αποτιμά ερωτήσεις σε σχήματα του σημασιολογικού ιστού.
10. Αξιολόγηση των συστημάτων οργάνωσης και αποθήκευσης σημασιολογικών σχημάτων.

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

**V: ΔΙΑΔΙΚΤΥΑΚΗ ΥΠΗΡΕΣΙΑ ΙΣΤΟΥ ΣΥΝΑΛΛΑΓΗΣ ΑΝΤΙΚΕΙΜΕΝΩΝ ΚΑΙ
ΥΠΗΡΕΣΙΩΝ**

ΠΛΗΡΟΦΟΡΙΕΣ: Δ. Σαχαρίδης, 210 772 1402, dsachar@dblab.ece.ntua.gr, Β. Καντερέ, 210 772 1402, verena@dblab.ece.ntua.gr, Θ. Δαλαμάγκας, 210 772 1402, dalamag@dblab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία θα δημιουργήσει μια υπηρεσία Παγκόσμιου Ιστού που θα για συναλλαγή αγαθών μεταξύ των μελών της.

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: υπό συζήτηση!

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:

Στόχος της διπλωματικής εργασίας είναι η δημιουργία ενός συστήματος συναλλαγής αγαθών στον Ιστό. Η οικονομία της ηλεκτρονικής αυτής *κοινωνίας* (κοινωνία V) είναι *μη νομισματική*, δηλαδή, δεν επιτρέπει χρηματικές δοσοληψίες, αλλά βασίζεται ουσιαστικά στην *υπόληψη* των μελών της. Όπως και σε μία αληθινή, τα μέλη της ηλεκτρονικής κοινωνίας σχηματίζουν κοινότητες με βάση κάποια κοινά ενδιαφέροντα ή επιδιώξεις, π.χ., φίλοι κατασκευής γηπέδου καλλιτεχνικού πατινάζ, είτε για διάφορους άλλους λόγους, όπως γεωγραφικούς.

- ✓ Μια συναλλαγή μπορεί να είναι: (α) *ανταλλαγή (exchange/swap)*, (β) *δωρεά (gift)*, (γ) *κοινοχρησία (sharing)* αγαθών.
- ✓ Τα αγαθά μπορεί να είναι: (α) *αντικείμενα*, (β) *υπηρεσίες*.

Ως παράδειγμα, θεωρούμε ένα χρήστη που διαθέτει φορητό υπολογιστή και θέλει να τον (i) ανταλλάξει με τηλεόραση πλάσματος, ή (ii) να το δωρίσει σε κάποιο μέλος ή κοινότητα, ή (iii) να το μοιράζεται με μέλη σε κάποια κοινότητα που ανήκει, π.χ., με τους συναδέλφους του. Επίσης, μπορούν να πραγματοποιηθούν συναλλαγές υπηρεσιών: ένας φοιτητής ικανός με τους υπολογιστές επιθυμεί να προσφέρει την τεχνογνωσία του με αντάλλαγμα φροντιστήριο σε κάποιο μάθημα από συμφοιτητή του. Τέλος συναλλαγές μπορούν να πραγματοποιηθούν μεταξύ υπηρεσιών και αντικειμένων: κάποιος διαθέτει δύο εισιτήρια για μία παράσταση αλλά δεν έχει αυτοκίνητο, οπότε ανταλλάσει το ένα με την μεταφορά του από και προς το συναυλιακό χώρο.

Κύριο συστατικό της κοινωνίας V είναι η *υπόληψη* των μελών της, εφόσον δεν επιτρέπονται οι χρηματικές δοσοληψίες. Η υπόληψη ενός μέλους είναι μια συνολική βαθμολογία που αντικατοπτρίζει πόσο καλός πολίτης της κοινωνίας V είναι. Μπορεί να προκύψει από τη συνάθροιση βαθμολογιών των συναλλαγών του μέλους.

Η διπλωματική αυτή επομένως, θα ασχοληθεί με τα παρακάτω θέματα:

11. Ανάλυση απαιτήσεων συστήματος.
12. Σχεδίαση Βάσης Δεδομένων/Συστήματος
13. Υλοποίηση υποσυστημάτων ορισμού κοινότητων, ανταλλαγής αγαθών και υπηρεσιών και διαχείρισης υπολήψεων.

**ΥΛΟΠΟΙΗΣΗ ΠΡΑΚΤΟΡΩΝ (Agents) ΓΙΑ ΔΙΚΤΥΟ ΚΙΝΟΥΜΕΝΩΝ ΟΜΟΤΙΜΩΝ
ΒΑΣΕΩΝ (Peer-2-Peer Mobile Databases)**

ΠΛΗΡΟΦΟΡΙΕΣ: Βηρένα Καντερέ, 2107721402, ykante@dblabb.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στην υλοποίηση ενός συστήματος κινούμενων ομότιμων όπου κάθε κόμβος αναζητά πληροφορία στο δίκτυο προωθώντας πράκτορες. Οι πράκτορες λειτουργούν ως σκανδαλιστές (triggers) που συλλέγουν πληροφορία από τις τοπικές βάσεις των κινούμενων κόμβων του συστήματος.

ΑΤΟΜΑ: 1

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: JAVA

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Θεωρούμε ένα σύστημα ομοτίμων όπου όμως οι τελευταίοι είναι κινούμενοι και κατέχουν μια μικρή βάση δεδομένων. Οι ομότιμοι επικοινωνούν με ασύρματο τρόπο απευθείας με άλλους ομότιμους που βρίσκονται μέσα στην ακτίνα επικοινωνίας τους. Οι ομότιμοι ζητούν να βρουν άλλους ομότιμους προωθώντας μια ερώτηση στο δίκτυο με τη μορφή ενός πράκτορα (agent). Ο τελευταίος είναι εκφρασμένος σύμφωνα με το ECA (event-condition-action) πρότυπο. Ο πράκτορας-ερώτηση που φτάνει σε έναν κόμβο εκμεταλλεύεται τον ενεργό μηχανισμό της τοπικής βάσης για να αποφασίσει αν ο κόμβος είναι κατάλληλος να του προσφέρει την πληροφορία που αναζητά. Έτσι, εγκαθίσταται στην τοπική βάση ως σκανδαλιστής (trigger) και απεγκαθίσταται όταν αποφασίσει ότι ο κόμβος δεν είναι κατάλληλος ή όταν έχει συλλέξει την πληροφορία που αναζητά.

**ΥΛΟΠΟΙΗΣΗ ΜΗΧΑΝΙΣΜΟΥ ΕΡΩΤΑΠΟΚΡΙΣΕΩΝ ΓΙΑ ΔΙΚΤΥΟ ΚΙΝΟΥΜΕΝΩΝ
ΟΜΟΤΙΜΩΝ ΒΑΣΕΩΝ (Peer-2-Peer Databases)**

ΠΛΗΡΟΦΟΡΙΕΣ: Βηρένα Καντερέ, 2107721402, vkante@dbl-lab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στην υλοποίηση κόμβων ομότιμων βάσεων που μπορούν να συνδέονται μεταξύ τους με αντιστοιχίσεις (mappings) και να θέτουν ερωτήσεις (queries) που να προωθούνται από τη μία στην άλλη. Μέσω των ερωτήσεων που θέτουν οι βάσεις αποσκοπούν στο να βρουν βάσεις με κοινά ενδιαφέροντα και να συνδεθούν μαζί τους.

ΑΤΟΜΑ: 1

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: JAVA

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Θεωρούμε ένα σύστημα ομοτίμων όπου κάθε ομότιμος κατέχει μια σχεσιακή βάση δεδομένων. Οι ομότιμοι που είναι απευθείας συνδεδεμένοι μεταξύ τους, έχουν ανά δύο αντιστοιχίσεις (mappings) πάνω σε μέρος των σχημάτων τους. Όταν ένας ομότιμος θέτει μια ερώτηση στο δίκτυο ομοτίμων, αυτή προωθείται σε κάποιο μονοπάτι κόμβων. Η ερώτηση πρέπει να μεταφράζεται σε κάθε κόμβο του μονοπατιού στο τοπικό σχήμα, έτσι ώστε να απαντηθεί από τον τοπικό κόμβο, αλλά και για να προωθηθεί παρακάτω. Έτσι όμως η ερώτηση χάνει σε κάθε κόμβο πληροφορία που δεν μπορεί να μεταφραστεί τοπικά, με αποτέλεσμα να μην μπορεί πολλές φορές να φτάσει μακριά. Για το λόγο αυτό, συχνά η ερώτηση δεν φτάνει σε κόμβους που μπορούν να την απαντήσουν ικανοποιητικά, ή όταν φτάνει σε αυτούς έχει χάσει σχετική πληροφορία.

Ο σκοπός της διπλωματικής αυτής είναι να υλοποιηθεί μια υπάρχουσα τεχνική για την αντιμετώπιση του παραπάνω προβλήματος. Συγκεκριμένα, η τεχνική προτείνει μια επαναληπτική μέθοδο μέσω της οποίας ένας κόμβος μαθαίνει σταδιακά για τα περιεχόμενα ενός απομακρυσμένου κόμβου. Κατά τη διαδικασία της μάθησης χτίζονται σταδιακά αντιστοιχίσεις μεταξύ των σχημάτων των απομακρυσμένων κόμβων. Δύο τέτοιοι κόμβοι μπορούν σε κάποιο σημείο να αποφασίσουν ότι έχουν κοινά ενδιαφέροντα – δηλαδή έχουν παρόμοια σχήματα, και να συνδεθούν απευθείας χρησιμοποιώντας τις αντιστοιχίσεις που χτίσανε κατά τη γνωριμία τους.

**ΠΡΟΤΥΠΟ ΣΥΣΤΗΜΑ ΟΛΟΚΛΗΡΩΣΗΣ ΣΧΗΜΑΤΩΝ ΜΕ ΥΠΟΣΤΗΡΙΞΗ
ΟΝΤΟΛΟΓΙΩΝ ΓΙΑ ΚΑΤΑΝΕΜΗΜΕΝΑ ΠΕΡΙΒΑΛΛΟΝΤΑ ΑΥΤΟΝΟΜΩΝ ΚΟΜΒΩΝ**

ΠΛΗΡΟΦΟΡΙΕΣ: Β. Καντερέ, 210 772 1402, vkante@dblab.ece.ntua.gr, Θ. Δαλαμάγκας, 210 772 1402, dalamag@dblab.ece.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στην ανάπτυξη του πρότυπου συστήματος διαχείρισης ολοκλήρωσης σχημάτων βάσεων δεδομένων για κατανεμημένα περιβάλλοντα αυτόνομων κόμβων, π.χ. δίκτυα ομοτίμων (p2p schema-based systems).

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Java ή C++ /XML/RDF

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Τα κατανεμημένα περιβάλλοντα αυτόνομων κόμβων προσφέρουν μια ευέλικτη αρχιτεκτονική υποστήριξης ανταλλαγής δεδομένων στον Ιστό. Δημοφιλές παράδειγμα είναι τα δίκτυα ομοτίμων (peer-2-peer systems). Ένα δίκτυο ομοτίμων κόμβων αποτελείται από ένα σύνολο αυτόνομων υπολογιστικών κόμβων στο Διαδίκτυο, οι οποίοι συνεργάζονται με σκοπό την ανταλλαγή δεδομένων. Τα συστήματα ομοτίμων κόμβων που χρησιμοποιούνται ευρέως σήμερα κυρίως για την ανταλλαγή αρχείων μουσικής, έχουν πολύ μικρές δυνατότητες διαχείρισης δεδομένων. Η αναζήτηση πληροφορίας γίνεται με χρήση λέξεων κλειδιών (keyword-based search). Η ανάγκη λοιπόν για πιο εκφραστικές λειτουργίες, σε συνδυασμό με την ανάπτυξη του Σηματολογικού Ιστού, οδήγησε στα συστήματα ομοτίμων κόμβων βασισμένα σε σχήμα (schema-based peer-to-peer systems).

Στα συστήματα αυτά κάθε κόμβος χρησιμοποιεί ένα σχήμα με βάση το οποίο οργανώνει τα τοπικά διαθέσιμα δεδομένα. Το ζήτημα λοιπόν που προκύπτει είναι πώς θα μπορούν οι κόμβοι να αναζητούν και να ανταλλάσσουν δεδομένα, διατηρώντας όμως την αυτονομία τους. Μια προσέγγιση απαιτεί να υπάρχει ένα κεντρικό σχήμα το οποίο θα χρησιμοποιούν όλοι οι κόμβοι. Οι ερωτήσεις διατυπώνονται και αποτιμούνται με βάση το ίδιο σχήμα. Μια τέτοια λύση θα ήταν καλή για περιβάλλοντα με καθορισμένα όρια, όπως για παράδειγμα το τοπικό δίκτυο ενός οργανισμού. Όμως σε ένα ανοιχτό περιβάλλον όπως είναι ο Παγκόσμιος Ιστός χρειάζεται ένα πιο ευέλικτο μοντέλο που να επιτρέπει την χρήση πολλών σχημάτων.

Η διπλωματική αυτή θα μελετήσει τρόπους *εξαγωγής ενός κεντρικού σχήματος G* για ένα σύνολο κόμβων που ο καθένας έχει το δικό του (τοπικό) σχήμα βάσης δεδομένων (ΒΔ). Τα τοπικά αυτά σχήματα θα θεωρούνται *υποσύνολα-όψεις (views)* του κεντρικού σχήματος. Δεν θα απαιτείται δηλαδή να υπάρχει εκ των προτέρων ένα κεντρικό σχήμα το οποίο θα χρησιμοποιούν όλοι οι κόμβοι.

Για την εξαγωγή του σχήματος αυτού θα ακολουθηθούν τα παρακάτω βήματα:

1. Χρήση μοντέλου γράφου για την αναπαράσταση τοπικών σχημάτων ΒΔ

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

2. Χειρωνακτική προσθήκη κανόνων αντιστοίχισης (*mapping rules*) που συσχετίζουν συστατικά των σχημάτων ΒΔ. Π.χ. ο κανόνας $\text{Person}(\text{id}, \text{age}, \text{salary}) \leftrightarrow \text{Faculty}(\text{id}, \text{age}, \text{salary})$ μας λέει ότι όλες οι εγγραφές *Person* είναι και *Faculty*.
3. Αυτόματη εξαγωγή σημασιολογικών σχέσεων από τα σχήματα ΒΔ και τους κανόνες αντιστοίχισης. Π.χ. από τον κανόνα $\text{Person}(\text{id}, \text{age}, \text{salary}) \leftrightarrow \text{Faculty}(\text{id}, \text{age}, \text{salary})$ μπορούμε να εξάγουμε την σημασιολογική σχέση *Person IS_A Faculty* (ότι δηλαδή οι δύο έννοιες σημασιολογικά ταυτίζονται).
4. Αναπαράσταση των μοντέλων γράφων τοπικών σχημάτων και των σημασιολογικών σχέσεων σε *οντολογίες* τύπου RDF/OWL.
5. *Ενοποίηση* (*merging*) των οντολογιών σε μια κεντρική οντολογία.
6. Εξαγωγή του κεντρικού σχήματος *G* που αντιστοιχεί στην κεντρική οντολογία. Το *G* είναι η *ολοκλήρωση* (*integration*) των τοπικών σχημάτων ΒΔ των κόμβων.

Η διπλωματική θα πρέπει επίσης να συγκρίνει την ποιότητα του κεντρικού σχήματος που προκύπτει από τα προηγούμενα βήματα με το κεντρικό σχήμα που προκύπτει από την απ'ευθείας ένωση των τοπικών σχημάτων ΒΔ, χωρίς την μεσολάβηση του σημασιολογικού εμπλουτισμού που δίνουν οι οντολογίες. Η σύγκριση μπορεί να γίνει, για παράδειγμα, μετρώντας το ποσοστό των συστατικών των τοπικών σχημάτων (σχέσεις και ιδιότητες) που μπορούν να ταιριάξουν με κάποιο συστατικό του κεντρικού σχήματος.

Η διπλωματική αυτή επομένως, θα ασχοληθεί με τα παρακάτω θέματα:

14. Μελέτη τεχνικών ολοκλήρωσης σχημάτων.
15. Μελέτη αναπαράστασης γνώσης με οντολογίες RDF/OWL.
16. Ανάπτυξη τεχνικής ολοκλήρωσης σχημάτων με υποστήριξη οντολογιών.
17. Πειραματική αξιολόγηση ποιότητας σχημάτων με χρήση γεννήτριας σχημάτων.

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

Partialist: ΣΥΣΤΗΜΑ ΕΠΕΞΕΡΓΑΣΙΑΣ ΕΡΩΤΗΣΕΩΝ ΓΙΑ ΔΕΝΤΡΙΚΑ ΔΕΔΟΜΕΝΑ

ΠΛΗΡΟΦΟΡΙΕΣ: Σ. Σουλδάτος, 210 7721402, stef@dblab.ece.ntua.gr, Θ. Δαλαμάγκας, 210 7721402, dalamag@dblab.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική εργασία στοχεύει στο σχεδιασμό και την υλοποίηση συστήματος επεξεργασίας ερωτήσεων για δεντρικά δεδομένα XML. Βασικό χαρακτηριστικό των ερωτήσεων είναι ότι επιτρέπουν μερικό προσδιορισμό δομής.

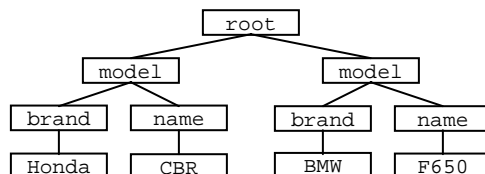
ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: C++/Linux

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:

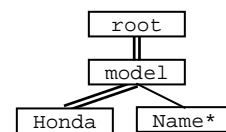
Η γλώσσα XML είναι πλέον πρότυπο ανταλλαγής δεδομένων σε εφαρμογές του Παγκόσμιου Ιστού. Η χρήση **ετικετών** (tags/elements) για το μαρκάρισμα και τον χαρακτηρισμό των δεδομένων σε ένα XML αρχείο διευκολύνει την επεξεργασία του από προγράμματα. Δείτε ένα πολύ απλό παράδειγμα XML αρχείου που αποθηκεύει πληροφορίες για μοντέλα μηχανών.

```
<?xml version="1.0">
<root>
  <model>
    <brand>Honda</brand>
    <name>CBR</name>
  </model>
  <model>
    <brand> BMW </brand>
    <name>F650</name>
  </model>
</root>
```



Το XML αρχείο του παραδείγματος αντιστοιχεί σε μια **δεντρική δομή**, αυτή που φαίνεται στο διπλανό σχήμα.

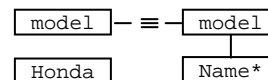
Γλώσσες ερωτήσεων, όπως η XPath, χρησιμοποιούνται ευρέως για την ανάκτηση δεδομένων από τέτοιες δεντρικές δομές. Για παράδειγμα, η ερώτηση `//model[//Honda]/Name`, ζητά τα ονόματα των μοντέλων της εταιρίας Honda. Η απάντηση της ερώτησης είναι η εικόνα της (mapping) πάνω στο δέντρο, εδώ το μοντέλο CBR.



Οι γλώσσες αυτές απαιτούν από το χρήστη να ξέρει τη δομή του XML αρχείου. Για παράδειγμα, στην προηγούμενη ερώτηση, θα πρέπει ο χρήστης να ξέρει ότι το Honda βρίσκεται κάτω από το model. Αν ο χρήστης δε γνωρίζει τις σχετικές θέσεις τους, θα πρέπει να φτιάξει ερωτήσεις για όλες τις πιθανές θέσεις των κόμβων, δηλ. μια ερώτηση που θα έχει το Honda πάνω από το model και μια που θα τα έχει ανάποδα.

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

Για το λόγο αυτό, έχουμε ορίσει μια **γλώσσα μερικώς ορισμένων δενδρικών ερωτήσεων (PTPQ)**, παρόμοια με την XPath, η οποία όμως επιτρέπει σε κόμβους να μείνουν ξεκρέμαστοι, όπως ο κόμβος Honda στο παράδειγμα. Επίσης, έχουμε ξεκινήσει την υλοποίηση αρκετών μηχανισμών, όπως αποτίμηση ερωτήσεων, έλεγχος ικανοποιησιμότητας, εξαγωγή πλήρους μορφής, έλεγχος περιεκτικότητας, κ.λ.π. Βασικό κομμάτι των μηχανισμών αυτών αποτελεί ο Γράφος Διαστάσεων που αποτελεί περίληψη της δομής του XML αρχείου.



Το ζητούμενο σύστημα θα εκτελεί τις παρακάτω λειτουργίες:

- *Επεξεργασία Δέντρων*: θα υπάρχει η δυνατότητα γραφικής αποτύπωσης, επεξεργασίας και εκ νέου αποθήκευσης δενδρικών δεδομένων. Επιπλέον, θα υποστηρίζεται η ημιαυτόματη εξαγωγή Γράφου Διαστάσεων.
- *Επεξεργασία Γράφων*: θα υπάρχει η δυνατότητα γραφικής αποτύπωσης, επεξεργασίας και αποθήκευσης γράφων.
- *Επεξεργασία Ερωτήσεων*: θα υπάρχει η δυνατότητα γραφικής αποτύπωσης, επεξεργασίας και αποθήκευσης ερωτήσεων. Επιπλέον, θα ενσωματωθούν υπάρχουσες λειτουργίες (υλοποιημένες), όπως η εξαγωγή ερώτησης πλήρους μορφής, έλεγχος ικανοποιησιμότητας, έλεγχος περιεκτικότητας. Τέλος, θα υλοποιηθούν αλγόριθμοι επεξεργασίας και αποτίμησης ερωτήσεων με χρήση μηχανής XQuery.

Η διπλωματική αυτή επομένως, θα ασχοληθεί με τα παρακάτω θέματα:

18. Μελέτη μερικώς ορισμένων δενδρικών ερωτήσεων.
19. Ανάλυση απαιτήσεων συστήματος.
20. Σχεδίαση συστήματος.
21. Υλοποίηση συστήματος και σύνδεση με τις υλοποιημένες λειτουργίες
22. Υλοποίηση αλγορίθμων επεξεργασίας και αποτίμησης ερωτήσεων με χρήση μηχανής XQuery.

ΥΛΟΠΟΙΗΣΗ ΨΗΦΙΑΚΗΣ ΒΑΣΗΣ ΦΩΤΟΕΡΜΗΝΕΥΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ ΣΤΟ ΔΙΑΔΙΚΤΥΟ

ΠΛΗΡΟΦΟΡΙΕΣ: Καθ. Δημήτριος Αργιαλάς, 2107722595, argialas@central.ntua.gr

ΠΕΡΙΛΗΨΗ: Η διπλωματική αυτή εργασία στοχεύει στην υλοποίηση μιας ψηφιακής βάσης φωτοερμηνευτικών δεδομένων στο διαδίκτυο για τη Σχολή Αγρονόμων-Τοπογράφων Μηχανικών του ΕΜΠ. Πρόκειται για ένα σύστημα γεωγνώσης στο οποίο υπάρχει η δυνατότητα παρουσίασης / διαχείρισης αεροφωτογραφιών και ψηφιακών τηλεπισκοπικών απεικονίσεων ποικίλων ειδών με συγκεκριμένη δομή. Απαραίτητη είναι και η δυνατότητα διασυνδέσεων ώστε να χρησιμοποιηθεί η βάση αυτή και ως βάση φωτοερμηνευτικών κλειδιών.

ΑΤΟΜΑ: 1

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Ανοικτό λογισμικό Βάσεων Δεδομένων: MySQL, Postgres

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Μια ψηφιακή βάση φωτοερμηνευτικών δεδομένων είναι μια βάση δεδομένων γεωγραφικού περιεχομένου. Αποτελείται κυρίως από δεδομένα τύπου raster τα οποία είτε είναι απεικονίσεις της φυσικής γήινης επιφάνειας (μονοχρωματικές, έγχρωμες, πολυφασματικές εικόνες) είτε είναι γεωγραφικά δεδομένα (θεματικοί χάρτες, ψηφιακά μοντέλα εδάφους, κ.α.). Για να χρησιμοποιηθεί η βάση αυτή και ως βάση φωτοερμηνευτικών κλειδιών για ακαδημαϊκή χρήση είναι απαραίτητα τα ακόλουθα:

Ιεραρχική δομή

Σύστημα διασυνδέσεων

Δυνατότητα αναζήτησης με πολλαπλά κριτήρια

Δυνατότητες πρόσβασης και διαχείρισης στους χρήστες μέσω διαδικτύου με παράλληλη καταγραφή αυτών καθώς και δυνατότητα διαβάθμισης

Η δομή των δεδομένων αυτών θα περιλαμβάνει τίτλο, ιεράρχηση, template περιγραφής της κάθε εικόνας (με γεωμετρικά, ραδιομετρικά, φασματικά και χρονικά χαρακτηριστικά), ψηφιακές εικόνες και πιθανόν συνοδευτικά ψηφιακά αρχεία. Απαραίτητο είναι το σύστημα των διασυνδέσεων ώστε να είναι δυνατή η με κάθε τρόπο πλοήγηση σε δεδομένα με ομοειδή ή αντίθετα χαρακτηριστικά. Παράλληλα θα πρέπει να υποστηρίζεται η δυνατότητα αναζήτησης με βάση τόσο γεωγραφικά όσο και φυσικά κριτήρια. Η πρόσβαση στην βάση θα γίνεται μέσω διαδικτύου και θα είναι διαβαθμισμένη ως προς την δυνατότητα χρήσης δηλαδή να μπορούν καθηγητές, μεταπτυχιακοί και προπτυχιακοί φοιτητές να προσθέτουν σελίδες-κλειδιά με παράλληλη καταγραφή ποιός πρόσθεσε τι και πότε.

Τέλος είναι επιθυμητή η χρήση ανοικτού λογισμικού ώστε να είναι δυνατή η συντήρηση και επαύξηση των δυνατοτήτων της στο μέλλον χωρίς δεσμεύσεις.

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργαστήριο Συστημάτων Βάσεων Γνώσεων & Δεδομένων

**ΣΧΕΔΙΑΣΜΟΣ ΚΑΙ ΥΛΟΠΟΙΗΣΗ ΔΙΑΔΙΚΤΥΑΚΗΣ ΕΦΑΡΜΟΓΗΣ ΓΙΑ ΤΗΝ
ΥΠΟΒΟΛΗ ΔΗΛΩΣΕΩΝ ΙΔΙΟΚΤΗΣΙΑΣ**

ΠΛΗΡΟΦΟΡΙΕΣ: Παπουτσής, krap@ktimatologio.gr

ΑΤΟΜΑ: 2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: .NET, Oracle

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Στόχος της διπλωματικής εργασίας είναι ο σχεδιασμός και η ανάπτυξη μίας διαδικτυακής εφαρμογής για την ηλεκτρονική υποβολή δήλωσης ιδιοκτησίας. Η εφαρμογή θα υποστηρίζει τόσο την δήλωση της ιδιοκτησίας όσο και ένα σύνολο από υποστηρικτικές λειτουργίες όπως π.χ. διαχείριση στοιχείων χρήστη, εισαγωγή αιτήσεων για διόρθωση στοιχείων κλπ. Η υλοποίηση του περιβάλλοντος διεπαφής με το χρήστη θα βασιστεί στην .NET πλατφόρμα και η διαχείριση των πληροφοριών θα βασιστεί στο Σύστημα Διαχείρισης Βάσεων Δεδομένων Oracle.