



**«Ι.Π.ΣΥ.Π.»**

**ΙΝΣΤΙΤΟΥΤΟ ΠΛΗΡΟΦΟΡΙΑΚΩΝ  
ΣΥΣΤΗΜΑΤΩΝ & ΠΡΟΣΟΜΟΙΩΣΗΣ**

**Διπλωματικές Εργασίες Ακαδ. Έτους 2008-2009**

## *Περιεχόμενα*

<b>ΕΙΣΑΓΩΓΗ .....</b>	<b>2</b>
<b>ΠΡΟΣΤΑΣΙΑ ΤΗΣ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΣΕ ΔΕΔΟΜΕΝΑ ΑΝΑΖΗΤΗΣΗΣ ΣΤΟΝ ΙΣΤΟ.....</b>	<b>3</b>
<b>ΠΡΟΣΤΑΣΙΑ ΤΗΣ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΣΕ ΔΙΑΡΚΕΙΣ ΕΡΩΤΗΣΕΙΣ .....</b>	<b>4</b>
<b>MOBILE BILLBOARD.....</b>	<b>5</b>
<b>ΜΕΤΑΤΡΟΠΗ ΕΑΔ ΜΕΤΑΔΕΔΟΜΕΝΩΝ ΣΕ ΜΟΡΦΗ ΟΝΤΟΛΟΓΙΑΣ CIDOC CRM.....</b>	<b>7</b>
<b>ΔΙΑΧΕΙΡΙΣΗ ΤΟΠΟΘΕΣΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΜΕΣΩ ΚΙΝΗΤΩΝ ΣΥΣΚΕΥΩΝ .....</b>	<b>8</b>
<b>ΔΙΑΧΕΙΡΙΣΗ MXML ΚΑΙ MXPATΗ ΜΕΣΩ ΣΧΕΣΙΑΚΩΝ ΒΔ.....</b>	<b>9</b>
<b>ΕΥΦΥΕΣ ΣΥΣΤΗΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ ΑΤΟΜ/ RSS.....</b>	<b>10</b>
<b>ΥΠΗΡΕΣΙΕΣ ΑΝΑΖΗΤΗΣΗΣ ΣΤΟΝ ΙΣΤΟ ΜΕ ΧΡΗΣΗ ΕΡΓΑΛΕΙΩΝ ΔΙΑΧΕΙΡΙΣΗΣ ΣΚΕΨΗΣ .....</b>	<b>12</b>
<b>ΠΡΟΗΓΜΕΝΗ ΕΦΑΡΜΟΓΗ ΙΣΤΟΥ ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ ΒΙΟΕΠΙΣΤΗΜΩΝ .....</b>	<b>14</b>
<b>MOBILE INTERFACE ΓΙΑ ΣΥΛΛΟΓΗ ΚΑΙ ΟΠΤΙΚΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ ΚΥΚΛΟΦΟΡΙΑΚΗΣ ΚΙΝΗΣΗΣ .....</b>	<b>15</b>
<b>GEO-BLOGGING – ΔΗΜΙΟΥΡΓΙΑ ΕΝΟΣ ΣΗΜΑΣΙΟΛΟΓΙΚΟΥ ΠΑΓΚΟΣΜΙΟΥ ΧΑΡΤΗ.....</b>	<b>16</b>

## ΕΙΣΑΓΩΓΗ

Το [Ινστιτούτο Πληροφοριακών Συστημάτων και Προσομοίωσης \(ΙΠΣΥΠ\)](#) είναι το νεότερο ινστιτούτο του [Ερευνητικού Κέντρου Καινοτομίας στις Τεχνολογίες της Πληροφορίας, των Επικοινωνιών και της Γνώσης "Αθηνά"](#). Ξεκίνησε την λειτουργία του τον Μάρτιο του 2007, με διευθυντή τον Τίμο Σελλή, Καθηγητή της Σχολής Ηλεκτρολόγων Μηχ. και Μηχ. Υπολογιστών Ε.Μ.Π.

Το ΙΠΣΥΠ διεξάγει έρευνα και συμμετέχει σε αναπτυξιακά έργα στον τομέα της διαχείρισης και επεξεργασίας δεδομένων σε πληροφοριακά συστήματα μεγάλης κλίμακας. Οι περιοχές ερευνητικής δραστηριότητας του ΙΠΣΥΠ περιλαμβάνουν:

- Συστήματα διαχείρισης γεω-πληροφορίας και υπηρεσίες εντοπισμού κινούμενων αντικειμένων.
- Επιστημονικές βάσεις δεδομένων Ιστού και διαχρονικές βάσεις δεδομένων.
- Προχωρημένες τεχνικές αναζήτησης πληροφορίας στον Ιστό με έμφαση στην προσωποποίηση δεδομένων στο προφίλ του χρήστη.
- Προστασία ιδιωτικότητας δεδομένων Ιστού.

Στα πλαίσια των περιοχών αυτών, το ΙΠΣΥΠ ανακοινώνει μια σειρά από διπλωματικές εργασίες. Οι εργασίες θα εκπονηθούν στο ΙΠΣΥΠ σε συνεργασία με το [Εργαστήριο Συστημάτων Βάσεων Γνώσεων και Δεδομένων της Σχολής Ηλεκτρολόγων Μηχ. και Μηχ. Υπολογιστών Ε.Μ.Π \(ΕΣΒΓΔ\)](#).



## ΠΡΟΣΤΑΣΙΑ ΤΗΣ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΣΕ ΔΕΔΟΜΕΝΑ ΑΝΑΖΗΤΗΣΗΣ ΣΤΟΝ ΙΣΤΟ

ΠΛΗΡΟΦΟΡΙΕΣ: Μανώλης Τερροβίτης, 210 6990522 (101), mter@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι να μελετηθεί το πρόβλημα της προστασίας των ευαίσθητων προσωπικών δεδομένων, σε δεδομένα που προέρχονται από αναζητήσεις χρηστών στον παγκόσμιο ιστό.

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** C/C++

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Οι περισσότερες ανθρώπινες δραστηριότητες τα τελευταία χρόνια προκαλούν κάποια καταγραφή δεδομένων, π.χ. οι αγορές σε ένα σουπερμάρκετ, οι τηλεφωνικές συνδιαλέξεις, οι μετακινήσεις με μέσα μαζικής μεταφοράς στα οποία χρησιμοποιούνται μαγνητικές κάρτες κτλ. Η μελέτη τέτοιων δεδομένων είναι πολύτιμη σε διάφορους τομείς καθώς επιτρέπει να γίνεται πιο αποδοτικός ο σχεδιασμός των υπηρεσιών. Παράλληλα, όμως, έχει δημιουργήσει ένα νέο πρόβλημα, αυτό της προστασίας της ταυτότητας των ατόμων από τα οποία προέρχονται οι καταγραφές. Ακόμη και αν αφαιρεθούν αναγνωριστικά που παραπέμπουν άμεσα στην ταυτότητα ενός προσώπου (π.χ. ΑΦΜ, ονοματεπώνυμο) είναι δυνατόν, μέσω, διασταύρωσης των δημοσιευμένων δεδομένων με άλλες πηγές γνώσης (π.χ. τηλεφωνικός κατάλογος) να αναγνωριστούν μοναδικά ή με μεγάλη βεβαιότητα οι εγγραφές που αφορούν σε έναν συγκεκριμένο άνθρωπο. Η προσπάθεια στην ερευνητική περιοχή της προστασίας της ιδιωτικότητας είναι να αναπτυχθούν μέθοδοι επεξεργασίας των δεδομένων που αποτρέπουν την σύνδεση των δημοσιευμένων εγγραφών με συγκεκριμένα πρόσωπα. Οι πιο συνήθεις εγγυήσεις για την προστασία των προσωπικών δεδομένων είναι αυτές της  $k$ -ανωνυμίας ( $k$ -anonymity) που εγγυάται ότι καμία εγγραφή δεν θα μπορεί να διακριθεί σε σχέση με άλλες  $k-1$  εγγραφές, και η  $l$ -διαφορετικότητα ( $l$ -diversity) που εγγυάται ότι θα υπάρχουν  $l$  διαφορετικές απαντήσεις σε κάθε πιθανή προσπάθεια ανάκτησης προσωπικών δεδομένων. Οι περισσότερες υπάρχουσες προσεγγίσεις υποθέτουν ότι τα δεδομένα είναι οργανωμένα αυστηρά στο σχεσιακό μοντέλο, και οι περισσότερες μέθοδοι ανωνυμοποίησης δουλεύουν μόνο για τέτοια δεδομένα. Στην πράξη όμως τα δεδομένα τα οποία καταγράφουν διάφορες δραστηριότητες είναι αρκετά πιο σύνθετα και απαιτούν νέες λύσεις για την προστασία των προσωπικών δεδομένων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει το πρόβλημα της δημοσίευσης δεδομένων από δικτυακές μηχανές αναζήτησης. Η δημοσίευση από την AOL, το καλοκαίρι του 2008, δεδομένων που προέρχονταν από αναζητήσεις χρηστών στον παγκόσμιο ιστό, δεν κατάφερε να αποτρέψει την αναγνώριση της ταυτότητας των διάφορων χρηστών. Το αποτέλεσμα ήταν να μπορεί να συνδεθεί μία σειρά από όρους αναζήτησης με συγκεκριμένους χρήστες, παραβιάζοντας έτσι την ιδιωτικότητα των χρηστών αυτών.

Στόχος της διπλωματικής είναι η αναγνώριση των πιθανών μοντέλων επίθεσης ενάντια στην ιδιωτικότητα σε δεδομένα από αναζητήσεις στο διαδίκτυο, καθώς και η υλοποίηση αλγορίθμων επεξεργασίας που εγγυώνται την ανωνυμοποίησή τους.



## ΠΡΟΣΤΑΣΙΑ ΤΗΣ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΣΕ ΔΙΑΡΚΕΙΣ ΕΡΩΤΗΣΕΙΣ

ΠΛΗΡΟΦΟΡΙΕΣ: Μανώλης Τερροβίτης, 210 6990522 (101), mter@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι να μελετηθεί το πρόβλημα της προστασίας των ευαίσθητων προσωπικών δεδομένων, σε δεδομένα από διαρκείς ερωτήσεις (continuous queries) και να σχεδιαστεί και υλοποιηθεί ο αλγόριθμος ανωνυμοποίησης τέτοιων δεδομένων, ώστε να είναι κατάλληλα προς δημοσίευση.

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** C/C++

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Οι περισσότερες ανθρώπινες δραστηριότητες τα τελευταία χρόνια προκαλούν κάποια καταγραφή δεδομένων, π.χ. οι αγορές σε ένα σουπερμάρκετ, οι τηλεφωνικές συνδιαλέξεις, οι μετακινήσεις με μέσα μαζικής μεταφοράς στα οποία χρησιμοποιούνται μαγνητικές κάρτες κτλ. Η μελέτη τέτοιων δεδομένων είναι πολύτιμη σε διάφορους τομείς καθώς επιτρέπει να γίνεται πιο αποδοτικός ο σχεδιασμός των υπηρεσιών. Παράλληλα, όμως, έχει δημιουργήσει ένα νέο πρόβλημα, αυτό της προστασίας της ταυτότητας των ατόμων από τα οποία προέρχονται οι καταγραφές. Ακόμη και αν αφαιρεθούν αναγνωριστικά που παραπέμπουν άμεσα στην ταυτότητα ενός προσώπου (π.χ. ΑΦΜ, ονοματεπώνυμο) είναι δυνατόν, μέσω, διασταύρωσης των δημοσιευμένων δεδομένων με άλλες πηγές γνώσης (π.χ. τηλεφωνικός κατάλογος) να αναγνωριστούν μοναδικά ή με μεγάλη βεβαιότητα οι εγγραφές που αφορούν σε έναν συγκεκριμένο άνθρωπο. Η προσπάθεια στην ερευνητική περιοχή της προστασίας της ιδιωτικότητας είναι να αναπτυχθούν μέθοδοι επεξεργασίας των δεδομένων που αποτρέπουν την σύνδεση των δημοσιευμένων εγγραφών με συγκεκριμένα πρόσωπα. Οι πιο συνήθεις εγγυήσεις για την προστασία των προσωπικών δεδομένων είναι αυτές της  $k$ -ανωνυμίας ( $k$ -anonymity) που εγγυάται ότι καμία εγγραφή δεν θα μπορεί να διακριθεί σε σχέση με άλλες  $k-1$  εγγραφές, και η  $l$ -διαφορετικότητα ( $l$ -diversity) που εγγυάται ότι θα υπάρχουν  $l$  διαφορετικές απαντήσεις σε κάθε πιθανή προσπάθεια ανάκτησης προσωπικών δεδομένων. Οι περισσότερες υπάρχουσες προσεγγίσεις υποθέτουν ότι τα δεδομένα είναι οργανωμένα αυστηρά στο σχεσιακό μοντέλο, και οι περισσότερες μέθοδοι ανωνυμοποίησης δουλεύουν μόνο για τέτοια δεδομένα. Στην πράξη όμως τα δεδομένα τα οποία καταγράφουν διάφορες δραστηριότητες είναι αρκετά πιο σύνθετα και απαιτούν νέες λύσεις για την προστασία των προσωπικών δεδομένων.

Μία σημαντική ομάδα δεδομένων είναι αυτή που προέρχεται από διαρκείς ερωτήσεις π.χ. την παρακολούθηση του στίγματος ενός κινούμενου αυτοκινήτου ή την παρακολούθηση των δοσοληψιών μίας πιστωτική κάρτα. Στόχος αυτής της διπλωματικής είναι η ανάπτυξη μίας μεθόδου προστασίας της ιδιωτικότητας σε τέτοιου τύπου δεδομένα, με κύριο άξονα τον ευφυή περιορισμό του χρονικού παραθύρου του διαστήματος αναφοράς κάθε εγγραφής.



## MOBILE BILLBOARD

ΠΛΗΡΟΦΟΡΙΕΣ: Γιάννης Σταύρακας, 210 6990522 (103), yannis@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Ο στόχος της διπλωματικής είναι η ανάπτυξη μιας εφαρμογής για κινητά τηλέφωνα, που να μας επιτρέπει να μοιραζόμαστε με άλλους πληροφορίες που μένουν επίκαιρες για μικρό χρονικό διάστημα.

**ΑΤΟΜΑ:** 1 (συνιστάται να υπάρχει προηγούμενη εξοικείωση με Java)

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Java για mobiles, Java, κάποια βάση δεδομένων.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Σε πολλές περιπτώσεις χρειαζόμαστε κάποια πληροφορία που η αξία της χάνεται μετά από μικρό χρονικό διάστημα, τέτοιο που κάνει μη πρακτική την καταγραφή και ανάκτηση της μέσω συμβατικών συστημάτων. Αυτή την πληροφορία γνωρίζουμε ότι κάποιοι την έχουν, αλλά δεν γνωρίζουμε ποιοι είναι αυτοί και πότε την έχουν για να τους ρωτήσουμε. *Παράδειγμα:* κάνω καθημερινά μια διαδρομή με την μηχανή μου μια συγκεκριμένη ώρα, και θέλω να γνωρίζω πριν ξεκινήσω αν κάπου στη διαδρομή αυτή υπάρχει μπλόκο. Φυσικά όποιος περνά από το μπλόκο το αντιλαμβάνεται, εγώ όμως δεν ξέρω ποιοί είναι αυτοί για να τους ρωτήσω. Ο στόχος της διπλωματικής είναι να υλοποιήσει μια εφαρμογή που θα επιτρέπει στον άγνωστο να στείλει τη συγκεκριμένη πληροφορία με το κινητό του, και σε μένα να την λάβω αυτόματα στο δικό μου. Ένας παράλληλος στόχος είναι η υποστήριξη ομαδικών δραστηριοτήτων, όπως για παράδειγμα επικοινωνία για ενημέρωση μιας παρέας για το που βρίσκεται και τι κάνει κάθε μέλος της.

Έτσι η διπλωματική θα υλοποιήσει ένα «πίνακα θεματικών ανακοινώσεων». Η εφαρμογή αποτελείται από δύο μέρη:

1. Το τμήμα που τρέχει στο κινητό.
2. Το τμήμα που τρέχει σε έναν κεντρικό server.

Το τμήμα που τρέχει στο κινητό θα πρέπει να επιτρέπει στον χρήστη:

- Να αιτείται τη δημιουργία ενός «θεματικού πίνακα» με κάποιον τίτλο, περιγραφή, κωδικό.
- Την ανάρτηση ανακοίνωσης (αποστολή μηνύματος) σε συγκεκριμένο πίνακα.
- Τη προαιρετική σύνδεση κάθε πίνακα με ομάδες χρηστών.
- Την εγγραφή του χρήστη σε συγκεκριμένους πίνακες.
- Την διαγραφή του χρήστη από συγκεκριμένους πίνακες.
- Τον ορισμό διαφόρων παραμέτρων, που αποτελούν το query για έναν πίνακα.

Το τμήμα που τρέχει στον server θα πρέπει να:

- Αποθηκεύει ότι απαραίτητη πληροφορία του στέλνουν τα κινητά.
- Επιτρέπει την έγκριση και απόρριψη αιτήσεων για θεματικούς πίνακες.
- Προωθεί τις ανακοινώσεις στους ενδιαφερόμενους (αποστολή μηνυμάτων).
- Παρέχει γενικές συγκεντρωτικές λειτουργίες πάνω στα δεδομένα.

Επεκτάσεις και διαφορετικές προσεγγίσεις στα παραπάνω είναι ευπρόσδεκτες και συζητήσιμες. Μια τέτοια επέκταση είναι η ύπαρξη συστήματος απόρριψης χρηστών που επανειλημμένα έχουν στείλει ακατάλληλες ή ανακριβείς πληροφορίες. Μια άλλη, είναι η εξέταση της δυνατότητας δυναμικού interface στο κινητό με ξεχωριστές φόρμες για κάθε νέο πίνακα.

Η διπλωματική έχει δύο άξονες:

1. Μελέτη / εκμάθηση του περιβάλλοντος προγραμματισμού Java για κινητά, και ανεύρεση εναλλακτικών λύσεων υλοποίησης (εφόσον υπάρχουν πολλές) για τα τεχνολογικά προβλήματα που θα αναλυθεί ότι προκύπτουν από την παραπάνω περιγραφή.
2. Σχεδιασμός και υλοποίηση της εφαρμογής που περιγράφεται. Έμφαση στο user interface σε ότι αφορά το τμήμα στο κινητό, και στην αποδοτικότητα στο τμήμα που αφορά στον server.



## ΜΕΤΑΤΡΟΠΗ EAD ΜΕΤΑΔΕΔΟΜΕΝΩΝ ΣΕ ΜΟΡΦΗ ΟΝΤΟΛΟΓΙΑΣ CIDOC CRM

ΠΛΗΡΟΦΟΡΙΕΣ: Γιάννης Σταύρακας, 210 6990522 (103), yannis@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Ο στόχος της διπλωματικής είναι η ανάπτυξη μιας εφαρμογής για την μετατροπή EAD XML εγγράφων σε σημασιολογικά ισοδύναμες περιγραφικές βασισμένες στην οντολογία CIDOC CRM.

**ΑΤΟΜΑ:** 1

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Οντολογίες, XML τεχνολογίες, Java.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Οι οργανισμοί (δημόσιοι και ιδιωτικοί) παράγουν πληροφορία η οποία τεκμηριώνει τη λειτουργία τους και οργανώνεται συστηματικά με τη μορφή αρχείων (archives). Τα αρχεία αποτελούνται από υλικό διαφορετικών μορφοτύπων (έντυπα/ηλεκτρονικά έγγραφα, φωτογραφίες, ηχητικό, πολυμεσικό υλικό κ.α.) και είναι φορείς πληροφορίας που αποδεικνύει/τεκμηριώνει τις δραστηριότητες οργανισμών. Η διατήρηση της αρχειακής πληροφορία είναι σημαντική τόσο για την εύρυθμη και αποδοτική λειτουργία των οργανισμών, όσο και για την ιστορική έρευνα και τη διατήρηση της πολιτιστικής κληρονομιάς.

Το Encoded Archival Description (EAD) (Library of Congress, ΗΠΑ) είναι ένα διεθνές πρότυπο αρχειακής περιγραφής, το οποίο εκφράζεται ως XML DTD και ως XML Schema. Η υλοποίησή του επιτρέπει τη δημιουργία ηλεκτρονικών εργαλείων έρευνας (*finding aids*), τα οποία ενσωματώνουν την περιγραφή των αρχείων (*archives*). Αναλυτικότερα, τα εργαλεία έρευνας είναι τα πληροφοριακά εργαλεία που περιλαμβάνουν μεταδεδομένα για την ταύτιση, διαχείριση και ερμηνεία ενός αρχείου και, συγχρόνως, αναλύουν το περιβάλλον δημιουργίας του. Το συγκεκριμένο XML σχήμα μεταδεδομένων υλοποιείται στις ψηφιακές βιβλιοθήκες τόσο σε εθνικό (Γενικά Αρχεία του Κράτους, Δημοτικό Κέντρο Ιστορίας και Τεκμηρίωσης Βόλου κλπ) όσο και διεθνές επίπεδο (Library of Congress, California Digital Library, Access to Archives-National Archives, UK- κλπ).

Στα πλαίσια μίας ευρύτερης τάσης για ολοκλήρωση δεδομένων και μεταδεδομένων (*data and metadata integration*) ενός συγκεκριμένου θεματικού τομέα αξιοποιούνται τα τελευταία χρόνια οι οντολογίες ως σχήματα διαμεσολάβησης πηγών (*mediated schema*). Πιο συγκεκριμένα, η οντολογία CIDOC CRM (ICOM/CIDOC and CIDOC CRM SIG) αξιοποιείται ως σχήμα διαμεσολάβησης πηγών μεταδεδομένων του τομέα πολιτιστικής κληρονομιάς (*αρχεία, βιβλιοθήκες μουσεία*). Αξίζει να σημειωθεί ότι η συγκεκριμένη οντολογία εκφράζεται σε σύνταξη RDF, OWL ενώ παράλληλα έχουν γίνει ερευνητικές προσπάθειες αποθήκευσης δεδομένων βασισμένων στην οντολογία σε σχεσιακές βάσεις δεδομένων. Με στόχο τη σημασιολογική ολοκλήρωση πηγών μέσω της συγκεκριμένης οντολογίας έχουν οριστεί σημασιολογικές συσχετίσεις (*semantic mappings*) ανάμεσα σε διάφορα σχήματα μεταδεδομένων, όπως το EAD, το Dublin Core κτλ, και την οντολογία CIDOC CRM.

Στόχος της προτεινόμενης εργασίας είναι η ανάπτυξη εργαλείου για την μετατροπή EAD XML εγγράφων σε σημασιολογικά ισοδύναμες περιγραφικές βασισμένες στην οντολογία CIDOC CRM. Το εργαλείο αυτό είναι δυνατό να αναπτυχθεί σε XSLT (ή σε κάποια γλώσσα προγραμματισμού, π.χ. JAVA, στον βαθμό που θα εκτιμηθεί ότι είναι καταλληλότερη για την ανάπτυξη ενός τέτοιου εργαλείου).

Οι σημασιολογική αντιστοίχιση των ανάμεσα στα πεδία του EAD και τις κλάσεις/ιδιότητες της οντολογίας CIDOC CRM θα βασιστεί σε προηγούμενη εργασία ενώ θα υπάρξει συνεργασία με τα μέλη της συγγραφικής ομάδας της εργασίας αυτής.





## ΔΙΑΧΕΙΡΙΣΗ ΤΟΠΟΘΕΣΙΩΝ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΜΕΣΩ ΚΙΝΗΤΩΝ ΣΥΣΚΕΥΩΝ

ΠΛΗΡΟΦΟΡΙΕΣ: Γιάννης Σταύρακας, 210 6990522 (103), yannis@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Ο στόχος της διπλωματικής είναι η ανάπτυξη μιας εφαρμογής για κινητά τηλέφωνα, που να μας επιτρέπει την εύκολη διαχείριση (εισαγωγή, τροποποίηση, ανεύρεση) τοποθεσιών ενδιαφέροντος.

**ΑΤΟΜΑ:** Ένα άτομο. Συνίσταται να υπάρχει προηγούμενη εξοικείωση με Java.

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Java για mobiles, Java, κάποια βάση δεδομένων.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Σε πολλές περιπτώσεις μετακινούμενοι στη πόλη περνούμε τυχαία από μέρη που θέλουμε να θυμηθούμε να επισκεφτούμε στο μέλλον. Η διαβάζουμε για μέρη που μας κινούν το ενδιαφέρον και τα βάζουμε στη νοητική λίστα «προς εξερεύνηση». Τέτοια σημεία μπορούν να ανήκουν σε διαφορετικές κατηγορίες, όπως εστιατόρια, γκαράζ, καταστήματα, καφέ, συνεργεία, κτλ, όλα όμως έχουν μια θέση στο χώρο. Σε άλλες περιπτώσεις έρχεται η στιγμή που χρειαζόμαστε και προσπαθούμε να θυμηθούμε κάποιο από τα μέρη αυτά, για παράδειγμα επειδή είμαστε σε μια συγκεκριμένη περιοχή και έχουμε χρόνο για ένα κάλο καφέ κάπου κοντά. Η επειδή ψάχνουμε τσαγκάρη αλλά δε θυμόμαστε που τον είχαμε δει.

Ο στόχος της διπλωματικής είναι να αναπτύξει μια εφαρμογή για την εύκολη διαχείριση τέτοιων περιπτώσεων. Η εφαρμογή θα αποτελείται από τρία μέρη:

1. Το τμήμα που τρέχει στο κινητό.
2. Το τμήμα που τρέχει μέσω web browser.
3. Το τμήμα που τρέχει σε έναν κεντρικό server.

Το τμήμα που τρέχει στο κινητό θα πρέπει να επιτρέπει στον χρήστη:

- Την άμεση και εύκολη καταχώρηση τοποθεσίας ενδιαφέροντος, της θέσης της (αντιμετώπιση περίπτωσης που η συσκευή γνωρίζει τις συντεταγμένες της, και περίπτωσης που απαιτείται περιγραφή από το χρήστη), και της κατηγορίας της.
- Την διαγραφή / τροποποίηση τοποθεσιών ενδιαφέροντος.
- Την ανεύρεση τοποθεσιών ενδιαφέροντος με διάφορα κριτήρια (μεταξύ των οποίων και: εγγύτητα στην τρέχουσα θέση, κατηγορία τοποθεσίας).
- Την δημιουργία κοινοτήτων χρηστών που μοιράζονται τις τοποθεσίες ενδιαφέροντός τους.
- Την εμφάνιση των τοποθεσιών σε χάρτη.

Το τμήμα που θα τρέχει στον web server θα δίνει στον χρήστη μεγάλο μέρος από τη λειτουργικότητα που θα προσφέρει το κινητό, αλλά και επιπλέον λειτουργικότητα με έμφαση στην διαχείριση των τοποθεσιών και άλλων παραμέτρων, ώστε να μπορεί να παραμείνει η εφαρμογή του κινητού απλή στη χρήση.

Το τμήμα που τρέχει στον server θα πρέπει να:

- Αποθηκεύει ότι απαραίτητη πληροφορία του στέλνουν τα κινητά.
- Προωθεί τις τοποθεσίες στους ενδιαφερόμενους.
- Παρέχει γενικές συγκεντρωτικές λειτουργίες πάνω στα δεδομένα.

Επεκτάσεις και διαφορετικές προσεγγίσεις στα παραπάνω είναι ευπρόσδεκτες και συζητήσιμες.

Η διπλωματική έχει δύο άξονες:

1. Μελέτη / εκμάθηση του περιβάλλοντος προγραμματισμού Java για κινητά, και ανεύρεση εναλλακτικών λύσεων υλοποίησης (εφόσον υπάρχουν πολλές) για τα τεχνολογικά προβλήματα που θα αναλυθεί ότι προκύπτουν από την παραπάνω περιγραφή.
2. Σχεδιασμός και υλοποίηση της εφαρμογής που περιγράφεται. Έμφαση στο user interface σε ότι αφορά το τμήμα στο κινητό, και στην αποδοτικότητα στο τμήμα που αφορά στον server.



## ΔΙΑΧΕΙΡΙΣΗ MXML ΚΑΙ MXPATH ΜΕΣΩ ΣΧΕΣΙΑΚΩΝ ΒΔ

ΠΛΗΡΟΦΟΡΙΕΣ: Γιάννης Σταύρακας, 210 6990522 (103), yannis@imis.athena-innovation.gr

**ΠΕΡΙΛΗΨΗ:** Υλοποίηση συστήματος αποθήκευσης Multidimensional XML (MXML) και απάντησης ερωτημάτων Multidimensional XPath (MXPath).

**ΑΤΟΜΑ:** 1 έως 2. Συνίσταται προηγούμενη εξοικείωση με Java.

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Java, XML, XPath, SQL.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Θυμίζω ότι η XML είναι μια γλώσσα αναπαράστασης (κωδικοποίησης) πληροφορίας για το Web, και το XPath είναι μια γλώσσα αναζήτησης δεδομένων μέσα σε XML κείμενα. Η MXML από την άλλη, είναι μια επέκταση της XML που αναπαριστά πληροφορία η οποία μπορεί να παρουσιάζει διαφορετικές εναλλακτικές μορφές ανάλογα με το **context**, όπου σαν context θεωρούμε μια σειρά «εξωτερικών» συνθηκών. Για παράδειγμα, το ίδιο βιβλίο ανάλογα με την γλώσσα που είναι γραμμένο (πχ. ελληνικά, αγγλικά) μπορεί να έχει διαφορετικό τίτλο, και μπορεί να έχει ή να μην έχει μεταφραστή. Έτσι η MXML είναι κατάλληλη για τον χειρισμό δεδομένων των οποίων η δομή και η τιμή αλλάζουν ανάλογα με τις συνθήκες (context). Αντίστοιχα, το MXPath είναι μια επέκταση του XPath που ενσωματώνει context, και είναι κατάλληλο για αναζήτηση πληροφορίας σε MXML κείμενα. Προηγούμενες εργασίες έχουν χρησιμοποιήσει σχεσιακές ΒΔ για την αποθήκευση XML εγγράφων, και για την απάντηση XPath ερωτημάτων (μέσω μετάφρασής τους σε SQL). Η διπλωματική αυτή θα βασιστεί σε προτάσεις που έχουν γίνει για την αποθήκευση MXML σε σχεσιακές βάσεις δεδομένων, και υλοποιήσει ένα σύστημα για MXML και MXPath, το οποίο:

- Θα αποθηκεύει σε σχεσιακούς πίνακες με συγκεκριμένο (ήδη προδιαγεγραμμένο) τρόπο οποιοδήποτε κείμενο MXML
- Θα απαντάει σε context-aware ερωτήματα MXPath πάνω στα αποθηκευμένα MXML κείμενα, μεταφράζοντας τα ερωτήματα αυτά σε ερωτήματα SQL πάνω στους σχετικούς πίνακες.

Ο τρόπος που προτείνεται να υλοποιηθούν τα παραπάνω έχει ως εξής:

- Σχεδιασμός και υλοποίηση αλγορίθμου που μετατρέπει και αποθηκεύει MXML σε συγκεκριμένη σχεσιακή δομή.
- Σχεδιασμός και υλοποίηση αλγορίθμου που μεταφράζει MXPath σε «ισοδύναμη» ερώτηση SQL.
- Σχεδιασμός και υλοποίηση αλγορίθμου που μετατρέπει περιεχόμενα των σχεσιακών πινάκων (που ερχονται σαν απάντηση στην ερώτηση SQL) πίσω σε MXML (που είναι και η τελική απάντηση στο αρχικό MXPath).
- Σχεδιασμός και υλοποίηση εφαρμογής που θα ενσωματώνει τους παραπάνω αλγορίθμους κάτω από ένα εννιαίο user interface, και που θα μπορεί να εισάγει και να διαγράφει MXML κείμενα, καθώς και να θέτει MXPath ερωτήματα και να παρουσιάζει τα αποτελέσματα.

Μια εναλλακτική υλοποίηση θα μπορούσε να αποτελείται από αλγόριθμους μετατροπής της MXML σε XML και του MXPath σε XPath, μαζί με χρήση κάποιου συστήματος διαχείρισης XML μέσω σχεσιακής βάσης. Η τελική προσέγγιση θα αποφασιστεί στην αρχή της διπλωματικής.

Η διπλωματική είναι προγραμματιστική και θα δώσει μια καλή πρακτική εξάσκηση σε Java, έχει όμως και αρκετό ερευνητικό υπόβαθρο, και είναι πιθανό να οδηγήσει σε κάποια δημοσίευση.

## ΕΥΦΥΕΣ ΣΥΣΤΗΜΑ ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ ATOM/ RSS

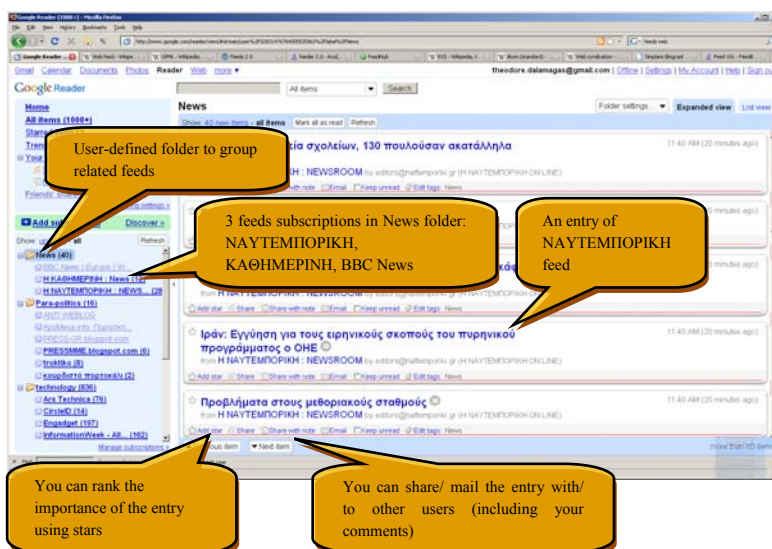
ΠΛΗΡΟΦΟΡΙΕΣ: Θεοδωρής Δαλαμάγκας, 210 6990522 (102), [dalamag@imis.athena-innovation.gr](mailto:dalamag@imis.athena-innovation.gr)  
Ανάσσης Βεργούλης, 210 6990522 (102), [bergoulis@dblabece.ntua.gr](mailto:bergoulis@dblabece.ntua.gr)

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη μιας ενός συστήματος διαχείρισης δεδομένων RSS/Atom (intelligent Web RSS/Atom feeder) που θα δίνει έμφαση στη διαχείριση περιεχομένου με μορφή ιστοριών.

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Ανοικτή

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Ένα Web feed είναι μια μορφή αρχείου δεδομένων, η οποία χρησιμοποιείται για να προσφέρει συχνά ανανεωνόμενο ψηφιακό περιεχόμενο (π.χ. εγγραφές ιστολογίων (blogs), πρωτοσέλιδα ειδήσεων κτλ) με εύκολο τρόπο στους χρήστες. Το RSS και το Atom είναι μορφές Web feed.



The screenshot shows the Google Reader interface with several annotations in yellow callout boxes:

- User-defined folder to group related feeds:** Points to a folder named 'News' in the left sidebar.
- 3 feeds subscriptions in News folder: ΝΑΥΤΕΜΠΟΡΙΚΗ, ΚΑΘΗΜΕΡΙΝΗ, BBC News:** Points to the 'News' folder.
- An entry of ΝΑΥΤΕΜΠΟΡΙΚΗ feed:** Points to a news entry from the 'News' folder.
- You can rank the importance of the entry using stars:** Points to a star icon next to a news entry.
- You can share/ mail the entry with/ to other users (including your comments):** Points to share and mail icons next to a news entry.

Οι χρήστες προκειμένου να λαμβάνουν το περιεχόμενο ενός feed χρησιμοποιούν ένα συσσωρευτή Web feed (Web feed aggregator/reader) με τον οποίο και εγγράφονται (subscribe) στην RSS υπηρεσία. Δείτε για παράδειγμα τον συσσωρευτή Google Reader στη διπλανή εικόνα.

Ο συσσωρευτής παρακολουθεί τα εγγεγραμμένα σε αυτόν feeds για το πότε έχουν νέο περιεχόμενο και ενημερώνει σχετικά το χρήστη ώστε να το δει.

Τα τελευταία χρόνια έχουν εμφανιστεί στην αγορά συσσωρευτές Web feed νέας γενιάς. Οι συσσωρευτές αυτοί δεν περιορίζονται απλώς στην εμφάνιση των περιεχομένων του feed, αλλά επιχειρούν επίσης να τα βαθμολογήσουν (ranking) με βάση τις προτιμήσεις του χρήστη. Οι προτιμήσεις του χρήστη μπορούν να εξαχθούν (α) άμεσα, π.χ. ρωτώντας το χρήστη εκ των προτέρων για το ποια είναι τα ενδιαφέροντά του και ποια άρθρα του αρέσουν ή (β) έμμεσα, π.χ. καταγράφοντας τις επιλογές του και ποια άρθρα έχει διαβάσει στο παρελθόν. Παραδείγματα τέτοιων συστημάτων είναι το FreeHub (<http://www.freehub.com>) και το Feeds 2.0 (<http://www.feeds2.com>).

Η διπλωματική εργασία στοχεύει στη σχεδίαση και ανάπτυξη ενός προηγμένου συσσωρευτή Web feed με τα ακόλουθα χαρακτηριστικά:

- **Βαθμολόγηση των περιεχομένων** του Web feed με βάση τις προτιμήσεις του χρήστη.
- **Αυτόματη ομαδοποίηση περιεχομένου** με βάση θεματική κατηγορία (topic-based aggregation). Η θεματική αυτή κατηγορία θα μπορεί να ορίζεται από το χρήστη.
- Έξυπνος χειρισμός **διπλοτύπων** (δηλαδή του ίδιου περιεχομένου που φτάνει στον συσσωρευτή από δύο διαφορετικές πηγές).

- **Διαχείριση περιεχομένου με μορφή ιστοριών.** Ο χρήστης θα μπορεί να ορίζει θέματα (topics) που αφορούν μια ιστορία. Ο συσσωρευτής θα μπορεί να ανιχνεύει περιεχόμενο σχετικό με τα θέματα αυτά και να το παρουσιάζει ως μια ιστορία, δηλαδή μια χρονική αλληλουχία γεγονότων. Οι ιστορίες θα διαχειρίζονται ως αυτόνομες οντότητες. Θα υπάρχουν δηλαδή μηχανισμοί αποθήκευσης, ανανέωσης, αναζήτησης παρόμοιων ιστοριών και αναζήτησης κομβικών σημείων των ιστοριών.



## ΥΠΗΡΕΣΙΕΣ ΑΝΑΖΗΤΗΣΗΣ ΣΤΟΝ ΙΣΤΟ ΜΕ ΧΡΗΣΗ ΕΡΓΑΛΕΙΩΝ ΔΙΑΧΕΙΡΙΣΗΣ ΣΚΕΨΗΣ

ΠΛΗΡΟΦΟΡΙΕΣ: *Θοδωρής Δαλαμάγκας, 210 6990522 (102), dalamag@imis.athena-innovation.gr*  
*Γιώργος Γιαννόπουλος, 210 6990522 (102), giann@dblab.ece.ntua.gr*

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη μιας εφαρμογής μετα-αναζήτησης πληροφορίας στον Ιστό που θα βασίζεται στην οργάνωση της αναζήτησης μέσω μοντέλων διαχείρισης σκέψης (mindmaps).

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Freemind/Java

### ΠΕΡΙΓΡΑΦΗ:

Οι μηχανές αναζήτησης, όπως το Google, είναι το βασικό εργαλείο αναζήτησης πληροφορίας στο Web. Η δημοφιλία στη χρήση τους οφείλεται σε δύο παράγοντες: (α) στην απλή μορφή της γλώσσας ερώτησης (keyword-based search), και (β) στην ωριμότητα των τεχνολογιών αναζήτησης κειμένων. Το απλό μοντέλο ερώτησης είναι σημαντικό πλεονέκτημα για αναζητήσεις σε θέμα καλά ορισμένο εκ των προτέρων. Αν για παράδειγμα θέλετε να βρείτε κριτικές για τις παραστάσεις “Sticky & Sweet tour” της Madonna, τότε πληκτρολογώντας απλά τις λέξεις-κλειδιά “sticky sweet madonna review” όλα τα πρώτα αποτελέσματα ικανοποιούν συνήθως πλήρως τις ανάγκες σας.

Συχνά όμως οι ανάγκες αναζήτησης πληροφορίας είναι πιο σύνθετες. Σκεφτείτε ένα μεταπτυχιακό φοιτητή που ψάχνει πληροφορίες για τις τρέχουσες τεχνολογίες, τις δημοσιεύσεις, τις ερευνητικές ομάδες, κ.λ.π. για μια ερευνητική θεματική περιοχή. Ο φοιτητής έχει στο μυαλό του μια **αφαιρετική περιγραφή του πεδίου γνώσης που θέλει να εξερευνήσει**.

✓ Για παράδειγμα, γνωρίζει ότι τον ενδιαφέρει το θέμα *κατευθυνόμενοι γράφοι* σε σχέση με το θέμα *βάσεις δεδομένων*. Η πληροφορία που αναζητά είναι δημοσιεύσεις τύπου *survey/review papers*, δηλαδή σχετικά άρθρα που συνοψίζουν την κατάσταση προόδου μιας συγκεκριμένης ερευνητικής περιοχής.

✓ Εκτός από αυτό, τον ενδιαφέρουν οι *αλγόριθμοι αποτίμησης ερωτήσεων σε γράφους*, και πιο συγκεκριμένα οι ερωτήσεις τύπου *reachability* που απαντούν αν δύο κόμβοι είναι στο ίδιο μονοπάτι. Γνωρίζει ότι οι αλγόριθμοι αυτοί είναι δύο κατηγοριών: αυτοί που χρησιμοποιούν κάποια μορφή αριθμητικής κωδικοποίησης (*labeling scheme*) για τους κόμβους και αυτοί που δεν τη χρησιμοποιούν. Και για τις δύο περιοχές θα ήθελε να βρει *σχετικά άρθρα από συνέδρια και επιστημονικά περιοδικά*.

✓ Γνωρίζει επίσης, ότι οι παλιοί αλγόριθμοι της δεύτερης κατηγορίας γνώρισαν μια δεύτερη άνθιση, γιατί πολλοί ερευνητές προσπάθησαν να αντιμετωπίσουν τους “ανταγωνιστές” αλγόριθμους της πρώτης κατηγορίας που επίσης εμφανίστηκαν λίγο πριν. Θέλει επομένως, τα άρθρα της δεύτερης κατηγορίας που θα αναζητηθούν να είναι *πιο νέα από τα άρθρα της πρώτης*.

Οι αφαιρετικές περιγραφές του πεδίου γνώσης είναι δημοφιλείς χαρακτηριστικό των **εργαλείων διαχείρισης της σκέψης (mind manager tool)**. Τέτοια εργαλεία είναι γνωστά στο χώρο της εκπαιδευτικής κοινότητας (δείτε σχετικά [http://en.wikipedia.org/wiki/Mind\\_map](http://en.wikipedia.org/wiki/Mind_map)). Χρησιμοποιούν συνήθως διαγράμματα αναπαράστασης ιδεών και συσχετίσεων μεταξύ τους, ώστε να βοηθήσουν τον εκπαιδευόμενο να κατανοήσει τις βασικές ιδέες τις οποίες αργότερα θέλει να αναλύσει και να εξειδικεύσει.

Πολλές φορές λοιπόν έχουμε στο μυαλό μας κάποιες **έννοιες (concepts)** που περιγράφουν το γενικότερο **πεδίο γνώσης** που θέλουμε να εξερευνήσουμε και να αναζητήσουμε πληροφορία. Για κάθε μια τέτοια έννοια, μπορεί να αναζητούμε διαφορετικά πράγματα: π.χ. κείμενα σχετικά με την έννοια Α, εικόνες σχετικές με την έννοια Β, blogs και άρθρα εφημερίδων για την έννοια Γ. Επίσης, οι έννοιες μπορεί να σχετίζονται



μεταξύ τους (**concept relationships**), και οι σχέσεις αυτές μπορεί να θέλουμε να υπόκειται σε **εξαρτήσεις (constraints)**.

Το τρέχον μοντέλο λειτουργικότητας των μηχανών αναζήτησης αδυνατεί να ικανοποιήσει τις ανάγκες χρηστών όπως ο παραπάνω φοιτητής. Συγκεκριμένα:

1. Το μοντέλο ερώτησης δεν μπορεί να εκφράσει ενιαία τις παραπάνω ανάγκες.
2. Η αφαιρετική περιγραφή του πεδίου γνώσης που περιγράψαμε, παρέχει από μόνη της σημασιολογική πληροφορία χρήσιμη για τη βελτίωση της αναζήτησης. Η βελτίωση αυτή μπορεί να είναι α) σε επίπεδο ποιότητας ταξινόμησης αποτελεσμάτων (**ranking**), αλλά και β) σε επίπεδο παρουσίασης (π.χ. **συσταδοποίηση – clustering** - αποτελεσμάτων με κοινό θεματικό περιεχόμενο). Ο μηχανισμός αποτίμησης ερωτήσεων στο τρέχον μοντέλο λειτουργικότητας των μηχανών αναζήτησης αγνοεί τη σημασιολογική αυτή πληροφορία.
3. Η αφαιρετική περιγραφή του πεδίου γνώσης παρέχει επίσης πληροφορία με την οποία μπορεί κάποιος να **συντονίσει την αναζήτηση**, κατευθύνοντας τις ερωτήσεις σε συγκεκριμένες μηχανές (π.χ. technorati για blogs, google images για εικόνες), πάλι έχοντας ως στόχο τη βελτίωση των αποτελεσμάτων αναζήτησης. Και πάλι ο μηχανισμός αποτίμησης ερωτήσεων στο τρέχον μοντέλο λειτουργικότητας των μηχανών αναζήτησης αδυνατεί να πραγματοποιήσει το συντονισμό αυτό.

Στα πλαίσια των δραστηριοτήτων του ΠΣΥΠ, αναπτύσσεται μια εφαρμογή που υποστηρίζει **προηγμένες τεχνικές αναζήτησης στον Ιστό**. Οι τεχνικές αυτές βασίζονται σε μοντέλο ερώτησης που χρησιμοποιεί αφαιρετικές περιγραφές και συσχετίσεις των εννοιών προς αναζήτηση. Το μοντέλο αυτό έχει κοινά σημεία με αυτό των **εργαλείων διαχείρισης σκέψης (mindmaps)**. Οι δύο βασικοί στόχοι του έργου είναι οι εξής: (α) επέκταση του απλού μοντέλου αναζήτησης με χρήση λέξεων κλειδιών, ενσωματώνοντας ιδέες **μοντέλων διαχείρισης σκέψης** στη διαδικασία της αναζήτησης και (β) κατασκευή ενός συνόλου υπηρεσιών σε ένα ενδιάμεσο επίπεδο μεταξύ χρήστη και μηχανών αναζήτησης που θα παρέχουν προηγμένες λειτουργίες αναζήτησης. Συγκεκριμένα, οι υπηρεσίες αυτές:

1. θα **συντονίζουν την αναζήτηση** σε πολλές μηχανές αναζήτησης και άλλες πηγές παροχής πληροφορίας στο Web ανάλογα με την περίπτωση. Το σύστημα θα επικεντρωθεί στην αναζήτηση κειμένων επιστημονικών συνεδρίων και περιοδικών, άρθρων εφημερίδων, κειμένων από blogs και εικόνων.
2. θα παρεμβαίνουν στην **ταξινόμηση των αποτελεσμάτων (ranking)** για να βελτιώσουν την ποιότητα των αποτελεσμάτων των μηχανών αναζήτησης. Αυτό μπορεί να γίνει λαμβάνοντας υπ' όψη (α) το **ιστορικό του χρήστη**, και (β) την αφαιρετική περιγραφή του πεδίου γνώσης για αναζήτηση.
3. θα παρεμβαίνουν στην παρουσίαση των αποτελεσμάτων, εφαρμόζοντας δυναμικά **τεχνικές συσταδοποίησης – clustering** – με βάση διάφορες ιδιότητες, π.χ. θεματικές περιοχές (π.χ. clustering ανά θέμα, τιμή πεδίου κ.λ.π.).

Προηγούμενη διπλωματική εργασία είχε ως θέμα της το 1. Η παρούσα διπλωματική εργασία θα ασχοληθεί κυρίως με το 3. Σε κάθε περίπτωση βέβαια, οι λειτουργικές απαιτήσεις είναι ενδεικτικές γιατί η εφαρμογή έχει πλήθος από ενδιαφέροντα λειτουργικά χαρακτηριστικά. Ο καθορισμός τους στα πλαίσια της διπλωματικής εξαρτάται και από τη φαντασία και τη δημιουργικότητά σας.



## ΠΡΟΗΓΜΕΝΗ ΕΦΑΡΜΟΓΗ ΙΣΤΟΥ ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ ΒΙΟΕΠΙΣΤΗΜΩΝ

**ΠΛΗΡΟΦΟΡΙΕΣ:** *Θοδωρής Δαλαμάγκας, 210 6990522 (102), [dalamag@imis.athena-innovation.gr](mailto:dalamag@imis.athena-innovation.gr)  
Θανάσης Βεργούλης, 210 6990522 (102), [bergoulis@dblab.ece.ntua.gr](mailto:bergoulis@dblab.ece.ntua.gr)*

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη μιας εφαρμογής Ιστού για την υποστήριξη ερευνητικών διεργασιών για ερευνητές βιοεπιστημών. Η εφαρμογή θα δίνει δυνατότητες αποθήκευσης, διαχείρισης και αναζήτησης βιολογικών δεδομένων, πειραμάτων και σχετιζόμενων πληροφοριών από βάσεις βιομορίων και ψηφιακές βιβλιοθήκες επιστημονικών δημοσιεύσεων.

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Ανοικτή

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Τα μόρια μικρο-RNA (micro-RNA) είναι βιομόρια πολύ μικρού μήκους, τα οποία προσκολλώνται επάνω σε μόρια «αγγελιοφόρου» RNA (mRNA) εμποδίζοντας την παραγωγή πρωτεϊνών. Για την έρευνα στις βιοεπιστήμες, είναι χρήσιμο σε κάποιον ερευνητή να γνωρίζει σε ποια mRNA προσκολλάται ένα συγκεκριμένο μόριο μικρο-RNA. Τα σημεία προσκόλλησης αποκαλούνται στόχοι (targets). Η πληροφορία αυτή είναι πολύτιμη για την εξαγωγή συμπερασμάτων σχετικά με τα αίτια ασθενειών. Επειδή ο ρόλος των μικρο-RNA ανακαλύφθηκε μόλις το 2001, η περιοχή είναι την περίοδο αυτή πολύ δημοφιλής στην ερευνητική κοινότητα των βιοεπιστημών.

Το **DIANA microT** (<http://diana.cslab.ece.ntua.gr/microT/>) είναι μια εφαρμογή Ιστού, που έχει αναπτυχθεί από το ΠΣΥΠ σε συνεργασία με το Ερευνητικό Κέντρο Βιοϊατρικών Επιστημών "Αλέξανδρος Φλέμινγκ". Η εφαρμογή προσφέρει στον ερευνητή τη δυνατότητα να αναζητήσει στόχους για διάφορα είδη μικρο-RNA. Οι στόχοι έχουν προϋπολογιστεί με διάφορες υπολογιστικές τεχνικές από την ερευνητική ομάδα του "Αλέξανδρος Φλέμινγκ" και έχουν αποθηκευτεί σε μια σχεσιακή βάση δεδομένων. Η πρόσβαση γίνεται μέσω φορμών αναζήτησης. Προσφέρονται διάφορα είδη αναζητήσεων και φιλτραρίσματος των αποτελεσμάτων. Επίσης, παρέχονται δυνατότητες αναζήτησης πληροφορίας σχετικής με τα εμπλεκόμενα βιομόρια πληροφορίας ή βιβλιογραφίας.

Στα πλαίσια της διπλωματικής εργασίας οι υποψήφιοι θα ασχοληθούν με διάφορα θέματα που αφορούν (α) τη βελτίωση της λειτουργικότητας της εφαρμογής DIANA microT, αλλά και (β) την επέκταση των δυνατοτήτων της. Ενδεικτικά αναφέρουμε τις ακόλουθες επεκτάσεις:

- Προσθήκη προσωπικού χώρου χρήστη (personal space) για την αποθήκευση και διαχείριση δεδομένων αναζήτησης.
- Μοντελοποίηση της πειραματικής διαδικασίας σε μια μορφή οντολογίας και χρήση της γνώσης για βελτίωση της ακρίβειας στην αναζήτηση σχετικών δημοσιεύσεων.
- Αναζήτηση δημοσιεύσεων με περιεχόμενο σχετικό με τα αποτελέσματα ενός πειράματος ανεύρεσης στόχου μικρο-RNA.

Οι επεκτάσεις είναι ενδεικτικές γιατί η εφαρμογή έχει πλήθος από ενδιαφέροντες λειτουργικές απαιτήσεις. Ο καθορισμός τους στα πλαίσια της διπλωματικής εξαρτάται και από τη φαντασία και τη δημιουργικότητά σας.



**MOBILE INTERFACE ΓΙΑ ΣΥΛΛΟΓΗ ΚΑΙ ΟΠΤΙΚΟΠΟΙΗΣΗ ΔΕΔΟΜΕΝΩΝ  
ΚΥΚΛΟΦΟΡΙΑΚΗΣ ΚΙΝΗΣΗΣ**

ΠΛΗΡΟΦΟΡΙΕΣ: Dieter Pfoser, [pfoser@imis.athena-innovation.gr](mailto:pfoser@imis.athena-innovation.gr)

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι ο σχεδιασμός και η υλοποίηση ενός φορητού συστήματος καταγραφής και οπτικοποίησης πληροφόρησης κίνησης (traffic) για την περιοχή της Αθήνας.

**ΑΤΟΜΑ:** 1++

**ΠΕΡΙΒΑΛΛΟΝ ΕΡΓΑΣΙΑΣ:** iPhone SDK, Unix (OS-X), C (Objective-C)

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Η διπλωματική θα σχεδιάσει και θα υλοποιήσει ένα φορητό σύστημα καταγραφής και οπτικοποίησης πληροφόρησης κίνησης (traffic) για την περιοχή της Αθήνας. Χρησιμοποιώντας τις εγγενείς δυνατότητες προσδιορισμού θέσης της φορητής συσκευής (iPhone) και πληροφορίες για το υπάρχον οδικό δίκτυο, γίνεται καταγραφή της τωρινής θέσης του χρήστη και υπολογισμός του χρόνου που απαιτείται (και άρα και της ταχύτητας) για τη μετάβαση από ένα σημείο του οδικού σημείου σε ένα άλλο. Οι πληροφορίες που καταγράφονται από τη συσκευή με αυτόν τον τρόπο, συλλέγονται σε ένα κεντρικό εξυπηρετητή (server) και είναι διαθέσιμες σε όλες τις συνδεδεμένες κινητές συσκευές (πελάτες). Ο χρήστης της κινητής συσκευής, συνεπώς θα έχει τα δυνατότητα οπτικής παρουσίασης των δεδομένων του εξυπηρετητή, μέσω μιας φιλικής οπτικής διασύνδεσης (GUI). Πρέπει να τονισθεί πως ο κεντρικός εξυπηρετητής συλλογής δεδομένων κίνησης έχει ήδη υλοποιηθεί (και βρίσκεται σε πλήρη λειτουργία) στα πλαίσια του Ευρωπαϊκού προγράμματος Track & Trade ([www.trackandtrade.org](http://www.trackandtrade.org)). Συνεπώς, η παρούσα εργασία θα εστιάσει αποκλειστικά στην υλοποίηση της client εφαρμογής που θα παρέχει την προτεινόμενη λειτουργικότητα στη φορητή συσκευή (iPhone).

Σημαντική έννοια για την παρούσα εργασία είναι η έννοια των Floating Car Data (FCD), δηλαδή η χρήση των δεδομένων που συλλέγονται από ένα όχημα – συσκευή ως δείγμα αξιολόγησης της κυκλοφοριακής κίνησης στο οδικό δίκτυο. Οι αλγόριθμοι για την λύση που θα βασίζεται σε FCD είναι διαθέσιμοι. Η εργασία θα αντιμετωπίσει το πρόβλημα της συλλογής των δεδομένων από τη φορητή συσκευή, την αποστολή – μεταφόρτωσή τους στον (ήδη υλοποιημένο) κεντρικό εξυπηρετητή, καθώς και την οπτικοποίηση των δεδομένων του εξυπηρετητή στη φορητή συσκευή, μέσω ενός φιλικού interface. Το αποτέλεσμά της θα είναι ένα πρωτότυπο (prototype) διαθέσιμο στο κοινό.

Το αντικείμενο της εργασίας είναι η υλοποίηση μιας εφαρμογής (με χρήση του iPhone SDK) που θα παρέχει

- (i) Μία «ζωντανή» (live) εικόνα της κατάστασης κίνησης στην Αθήνα, βασισμένη στα δεδομένα FCD του κεντρικού εξυπηρετητή.
- (ii) Δυνατότητες συλλογής FCD από τη φορητή συσκευή και αποστολή των δεδομένων στον κεντρικό εξυπηρετητή.

Ειδικότερα, στην διπλωματική εργασία προβλέπεται να καλυφθούν τα εξής ζητήματα:

- Σχεδιασμός και υλοποίηση ενός interface για τη συλλογή δεδομένων FCD από την κινητή συσκευή.
- Αποστολή των δεδομένων FCD που συλλέγονται στον κεντρικό εξυπηρετητή.
- Οπτικοποίηση των FCD του κεντρικού εξυπηρετητή στη φορητή συσκευή

Για την εργασία θα υπάρξει πρόσβαση στα ακόλουθα δεδομένα:

Δεδομένα κίνησης της Αθήνας μέσω του έργου TRACK&TRADE – [www.trackandtrade.org](http://www.trackandtrade.org).





## GEO-BLOGGING – ΔΗΜΙΟΥΡΓΙΑ ΕΝΟΣ ΣΗΜΑΣΙΟΛΟΓΙΚΟΥ ΠΑΓΚΟΣΜΙΟΥ ΧΑΡΤΗ

ΠΛΗΡΟΦΟΡΙΕΣ: Dieter Pfoser, [pfoser@imis.athena-innovation.gr](mailto:pfoser@imis.athena-innovation.gr)

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι η δημιουργία ενός διαδικτυακού portal που θα παρέχει τη δυνατότητα geo blogging, δηλαδή τη συλλογή και αποθήκευση πολυμεσικού υλικού, τοποθεσιών και διαδρομών από τους χρήστες του portal.

**ΑΤΟΜΑ:** 1++

**ΠΕΡΙΒΑΛΛΟΝ ΕΡΓΑΣΙΑΣ:** MapServer, OpenLayers, AJAX, PostgreSQL, PostGIS, Google Maps API.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Η διπλωματική θα δημιουργήσει ένα διαδικτυακό portal με δυνατότητες geo blogging, δηλαδή τη συλλογή και αποθήκευση πολυμεσικού υλικού, τοποθεσιών και διαδρομών από τους χρήστες του portal. Στόχος της διπλωματικής είναι η χρήση των δεδομένων (πολυμεσικού υλικού, τοποθεσιών και πορειών) που συλλέγονται από τους χρήστες, προκειμένου να δημιουργηθεί ένας παγκόσμιος χάρτης σημασιολογικά πλούσιος σε πολυμεσικό περιεχόμενο. Έτσι μια γεωγραφική τοποθεσία θα μπορεί να συνοδεύεται από φωτογραφίες, ήχους, video ή κείμενο που συλλέχθηκε από τους χρήστες του portal. Η Web εφαρμογή που θα αναπτυχθεί, θα πρέπει να επιτρέπει στους χρήστες να προσθέτουν (μεταφορτώνουν), μεταβάλλουν και να δημοσιεύουν τα δεδομένα τους με ευέλικτο τρόπο. Κεντρικός στόχος είναι η ανάπτυξη μιας πλατφόρμας που θα ενθαρρύνει τους χρήστες να διαμοιράζουν τα γεωγραφικά δεδομένα που συνέλεξαν, προκειμένου να δημιουργηθεί ένας παγκόσμιος χάρτης σημασιολογικά πλούσιος σε πολυμεσικό περιεχόμενο.

Αντικείμενο της διπλωματικής είναι η δημιουργία μιας Web εφαρμογής που θα παρέχει:

- (i) Μια απλή και κομψή διασύνδεση χρήστη που θα επιτρέπει στους χρήστες να μεταφορτώνουν κείμενο και πολυμεσικό υλικό που σχετίζεται με τη θέση τους, συνδέοντας έτσι τοποθεσίες του πραγματικού κόσμου με εικόνες, video, ήχους και κείμενο.
- (ii) Την απαραίτητη λειτουργικότητα, ώστε οι χρήστες να μπορούν να επεξεργαστούν τα δεδομένα που έχουν μεταφορτωθεί, δηλαδή:
  - a. Τη μεταβολή της τοποθεσίας με δυνατότητα drag and drop τοποθεσιών και διαδρομών
  - b. Προσθήκη, μεταβολή και διαγραφή πολυμεσικού περιεχομένου.

**Η δημιουργία ενός Γεω-ιστολογίου πρέπει να είναι το ίδιο εύκολη με τη σύνταξη ενός email!**

Ειδικότερα, στην διπλωματική εργασία προβλέπεται να καλυφθούν τα εξής ζητήματα:

- Δημιουργία και υλοποίηση μιας διαδικτυακής διασύνδεσης χαρτών – Google Maps Mashup (<http://code.google.com/apis/maps/>)
- Δυνατότητα μεταφόρτωσης δεδομένων
  - GPS δεδομένα σε διάφορες μορφές (gpx)
  - Πολυμεσικό υλικό (εικόνες, video, ήχοι και σημειώσεις κειμένου)
- Δυνατότητα επεξεργασίας των δεδομένων που έχουν μεταφορτωθεί
  - Drag and drop λειτουργικότητα για την επεξεργασία μιας τοποθεσίας ή μιας διαδρομής. Παρόμοια με τις δυνατότητες που παρέχει το Google Maps στη μεταβολή μιας διαδρομής.

- Drag and drop λειτουργικότητα για το πολυμεσικό περιεχόμενο, e.g., μετακίνηση της τοποθεσίας των εικόνων, ήχων και των σημειώσεων κειμένου
- Δυνατότητα συλλογικής δεικτοδότησης (tagging) δεδομένων, π.χ. με προσθήκη κοινών ετικετών, όπως τουρισμός, ποδηλατοδρόμιο στα δεδομένα
- Δυνατότητα επερωτήσεων στα δεδομένα
  - Γεωγραφικές επερωτήσεις (spatial queries)– ο χρήστης θα επιλέγει το παράθυρο επερωτήσεων
  - Επερωτήσεις κειμένου με τη χρήση λέξεων – κλειδιά (keyword search)
- Σχεδιασμός και υλοποίηση μια βάσης δεδομένων (PostgreSQL/PostGIS) για την αποθήκευση των δεδομένων του συγκεκριμένου Web portal που θα δημιουργήσουμε.

Ενδεικτικές πηγές στο Διαδίκτυο:

- Google Maps API: <http://code.google.com/apis/maps/>
- Collaborative mapping efforts:
  - <http://www.openstreetmap.org/>,
  - <http://www.openrouteservice.org/>,
  - <http://platial.com/>.