



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

**ΕΡΕΥΝΗΤΙΚΟ ΚΕΝΤΡΟ ΚΑΙΝΟΤΟΜΙΑΣ ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΤΩΝ
ΕΠΙΚΟΙΝΩΝΙΩΝ & ΤΗΣ ΓΝΩΣΗΣ «ΑΘΗΝΑ» /
ΙΝΣΤΙΤΟΥΤΟ ΠΛΗΡΟΦΟΡΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**



Τίτλος Πράξης

ΘΑΛΗΣ–Ερευνητικό Κέντρο Καινοτομίας στις Τεχνολογίες της Πληροφορίας, των Επικοινωνιών και της Γνώσης «Αθηνά» – ΕΙΚΟΣ : ΘΕΩΡΗΤΙΚΗ ΚΑΙ ΑΛΓΟΡΙΘΜΙΚΗ ΘΕΜΕΛΙΩΣΗ ΓΙΑ ΠΡΟΣΩΠΟΚΕΝΤΡΙΚΑ ΣΥΝΕΡΓΑΤΙΚΑ ΠΛΗΡΟΦΟΡΙΑΚΑ ΣΥΣΤΗΜΑΤΑ

THALIS – “ATHENA” RESEARCH AND INNOVATION CENTER IN INFORMATION, COMMUNICATION AND KNOWLEDGE TECHNOLOGIES - EICOS: foundations for perSONalized Cooperative Information Ecosystems

A/A Υποέργου: 1

**Ετήσια Έκθεση Προόδου Υποέργου
Από έναρξη υποέργου έως και 31/12/2013**



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

Πίνακας Περιεχομένων

1. Εισαγωγή.....	3
2. Σύντομη Περιγραφή του Υποέργου.....	5
3. Πορεία Υλοποίησης του Υποέργου	8
4. Προβλήματα κατά τη διάρκεια Υλοποίησης του Υποέργου	14
5. Ανακεφαλαίωση και Μελλοντικές Εργασίες	15



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

1. Εισαγωγή

Η παρούσα ετήσια έκθεση αφορά το υποέργο με τίτλο «ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα» που υλοποιείται στο Ινστιτούτο Πληροφοριακών Συστημάτων του Ερευνητικού Κέντρου «ΑΘΗΝΑ». Η έκθεση αφορά το χρονικό διάστημα από την έναρξη του υποέργου μέχρι και την 31/12/2013.

Ο **βασικός στόχος** του έργου είναι να προσφέρει τη μεθοδολογία, τη θεωρητική θεμελίωση, τις αλγοριθμικές τεχνικές και την αρχιτεκτονική του λογισμικού που απαιτείται ώστε τα πληροφοριακά συστήματα να μπορούν να προσφέρουν στους χρήστες αφενός την δυνατότητα εξατομίκευσης της παρεχόμενης πληροφορίας και αφετέρου τη δυνατότητα χρήσης ενσωματωμένων ετερογενών δεδομένων, ενδεχομένως διαφορετικής προέλευσης, με διαφανή τρόπο.

Η **θεμελιώδης ιδέα** πάνω στην οποία στηρίζεται το έργο για να επιτύχει τον προαναφερθέντα στόχο είναι η έννοια του *υπερχώρου δεδομένων (dataspace)* ο οποίος είναι μια οργάνωση της πληροφορίας με σημασιολογικά πλούσιο σε μεταπληροφορία τρόπο, που επιτρέπει τη διαχείριση δεδομένων άσχετα από τη διαμόρφωση (format), προέλευση, και δομή τους. Τα δεδομένα περιγράφονται σε σχέση με το μοντέλο (μετα)πληροφορίας του υπερχώρου και οι πράξεις των χρηστών ή οι επεμβάσεις του συστήματος στο να εξατομικεύει το περιεχόμενο και την παρουσίαση δεδομένων με βάση τον συγκεκριμένο κάθε φορά χρήστη μετατρέπονται, από πράξεις επί του υπερχώρου, σε πράξεις επί των υποκειμένων δεδομένων. Ο υπερχώρος επιτρέπει την ολιστική αντιμετώπιση και των δεδομένων και των εφαρμογών που τα διαχειρίζονται, ως ένα *οικοσύστημα πληροφορίας*.

Το **αντικείμενο του έργου** μπορεί να θεωρηθεί η πρόταση λύσεων για τα ακόλουθα ερευνητικά θέματα:

- Ικανότητα να δοθεί στον (κυρίαρχο, πλέον) τελικό χρήστη η δυνατότητα να συνθέτει πληροφορίες από πολλές πηγές με τρόπο εξατομικευμένο στις προτιμήσεις του.
- Έμφαση στην ευχρηστία και την εξατομίκευση της πληροφορίας (στο περιεχόμενο, τη δομή και την παρουσίαση).
- Ικανότητα να προσαρμόζουμε τα πληροφοριακά οικοσυστήματα σε μη παραδοσιακά δεδομένα (επιστημονικά, βιβλιογραφικά, προερχόμενα από τον Παγκόσμιο Ιστό, κ.ο.κ.).
- Κατανομή της υπολογιστικής υποδομής και αλγοριθμικής λειτουργίας ενός πληροφοριακού οικοσυστήματος σε πολλά, ομοτίμως συνεργαζόμενα συστήματα που είναι εύκολα στην εγκατάσταση, επερώτηση και διαχείριση.
- Ευελιξία και προσαρμοστικότητα στην εξέλιξη του περιβάλλοντος ενός πληροφοριακού οικοσυστήματος.

Οι επιμέρους **στόχοι του έργου** μπορούν να συνοψισθούν στα παρακάτω προσδοκώμενα αποτελέσματα, κάθε ένα από τα οποία οργανώνεται σε μια ξεχωριστή δράση (ή Πακέτο Εργασίας) ως εξής:

- Θεωρητική θεμελίωση, αλγοριθμικές τεχνικές και αρχιτεκτονικές λογισμικού για την ένταξη μη παραδοσιακών δεδομένων σε υπερχώρους. Η βασική προσέγγιση θα γίνει μέσω τεχνικών ανάκτησης μεταπληροφορίας και κατάλληλης δεικτοδότησης των πηγαίων δεδομένων για μη παραδοσιακά δεδομένα, με σκοπό να επιτρέψουμε στους χρήστες να εντάσσουν εύκολα νέα δεδομένα στο σύστημα.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
πρόγραμμα για την ανάπτυξη

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

- Θεωρητική θεμελίωση, αλγοριθμικές τεχνικές και αρχιτεκτονικές λογισμικού για τον ορισμό και την αξιοποίηση υπερχώρων. Η βασική προσέγγιση θα γίνει μέσω τεχνικών διαχείρισης μοντέλων (model management) και οι σχετικές δράσεις θα επιτρέψει την ορθή και αποδοτική ανεύρεση και αξιοποίηση της μεταπληροφορίας για να επιτραπεί στον χρήστη να δουλεύει με τα συλλεγμένα δεδομένα.
- Θεωρητική θεμελίωση, αλγοριθμικές τεχνικές και αρχιτεκτονικές λογισμικού επί της προσαρμογής και εξατομίκευσης της πληροφορίας, με σκοπό να επιτρέψουμε στους χρήστες να επερωτούν το σύστημα με τρόπο διαισθητικό και εύκολο, να συνθέτουν επί μέρους απαντήσεις σε νέα αποτελέσματα και να εξειδικεύουν την κατάταξη των αποτελεσμάτων με τρόπο ανάλογο προς τις προτιμήσεις τους
- Αλγόριθμοι και μέθοδοι κατανομής της αποθήκευσης και του υπολογισμού, με σκοπό να εκμεταλλευθούμε την εξ αντικειμένου κατανεμημένη φύση των νέων υπολογιστικών συστημάτων και να εμπλουτίσουμε τον υπερχώρο με τη δυνατότητα να είναι κατανεμημένος, και να έχει στατιστική μεταπληροφορία, εξαγόμενη με κατανεμημένο τρόπο και για τα δεδομένα, αλλά για τους χρήστες
- Θεωρητική θεμελίωση, αλγοριθμικές τεχνικές και αρχιτεκτονικές λογισμικού για τη διαχείριση και σχεδίαση ενός υπερχώρου δεδομένων υπό το πρίσμα της εξέλιξής του στο χρόνο, με σκοπό να υποστηρίξουμε την ομαλή εξέλιξη της πληροφορίας στην πορεία του χρόνου από πλευράς σημασιολογίας, δομής, περιεχομένου και όγκου με την ελάχιστη δυνατή επίπτωση σε χρήστες και εφαρμογές

Ο σκοπός της έκθεσης αυτής είναι κατ' αρχήν η συγκεντρωτική παρουσίαση της προόδου του υποέργου στο διάστημα 1/2/2012 ως και 31/12/2013. Επιπλέον, περιγράφονται οι μελλοντικές προοπτικές και τα επόμενα βήματα για το 2014.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

2. Σύντομη Περιγραφή του Υποέργου

Ο βασικός μηχανισμός που θα χρησιμοποιηθεί είναι ο μηχανισμός υπερχώρων δεδομένων. Ένας **υπερχώρος δεδομένων** (dataspace) είναι μια πλατφόρμα η οποία έχει την ικανότητα να οργανώνει πληροφορίες από πολλές ετερογενείς πηγές δεδομένων (data sources). Για να λειτουργήσει ένας υπερχώρος δεδομένων χρειάζεται:

1. να μπορεί να δημιουργήσει ένα οργανωμένο χώρο πληροφοριών ο οποίος να μπορεί στη συνέχεια να επερωτηθεί,
2. να μπορεί να εξάγει πληροφορία από τις ετερογενείς πηγές δεδομένων την οποία να «τακτοποιεί» στην παραπάνω δομή,
3. να μπορεί να ανανεώνει την πληροφορία όταν αυτή αλλάζει, καθώς και να διαχειρίζεται «πολλαπλές εκδοχές της αλήθειας» (για δεδομένα με ασάφεια),
4. να μπορεί να (αυτο-)αναδιοργανώνει την εσωτερική του δομή,
5. να μπορεί να αφήνει το χρήστη να θέτει ερωτήσεις και να συνθέτει τις επί μέρους απαντήσεις,
6. να μπορεί να παρέχει προσεγγιστικές απαντήσεις στις ερωτήσεις του χρήστη, αν δεν έχει τις ακριβείς,
7. να μπορεί να εξειδικεύσει και εξατομικεύσει τις απαντήσεις ανά χρήστη.

Ένας υπερχώρος δεδομένων επιτρέπει τη διαχείριση δεδομένων άσχετα από τη διαμόρφωση (format), προέλευση, και δομή τους. Από πλευράς λογισμικού, ο πιο σημαντικός ρόλος παίζεται από το ενδιάμεσο λογισμικό -- DataSpace Support Platform (DSSP) -- που επιτρέπει την διαχείριση των δεδομένων μέσω συγκεκριμένων υπηρεσιών που παρέχει στον προγραμματιστή, τον τελικό χρήστη και το διαχειριστή του. Ένας DSSP εκμεταλλεύεται την ολοκλήρωση των δεδομένων σε χαμηλό επίπεδο για να παρέχει υπηρεσίες επερώτησης και διαχείρισης. Παρά τις σημαντικές ελπίδες που τρέφει η ερευνητική κοινότητα για τους υπερχώρους δεδομένων, δεν φαίνεται να υπάρχει ακόμα κάποια σοβαρή υλοποίηση πλατφόρμας που να υποστηρίζει τα αρχικά θεωρητικά αποτελέσματα. Ένα βασικό εμπόδιο είναι η εξαγωγή της κατάλληλης μεταπληροφορίας που να επιτρέπει την ομαλή ενσωμάτωση των δεδομένων υπό ολοκλήρωση στον υπερχώρο. Η εν λόγω λειτουργία πρέπει να μπορεί να επιτευχθεί για πολλές κατηγορίες δεδομένων (όχι μόνο σχεσιακά δεδομένα, αλλά και για ημι-δομημένα, πολυμεσικά και γεωγραφικά δεδομένα, δεδομένα αισθητήρων, κ.ά.). Έτσι, το θεμελιώδες ερώτημα του πώς θα επιτευχθεί η εξυπηρέτηση των υπηρεσιών που υπόσχεται ο υπερχώρος πάνω από ολοκληρωμένα δεδομένα είναι ακόμα ανοικτό.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

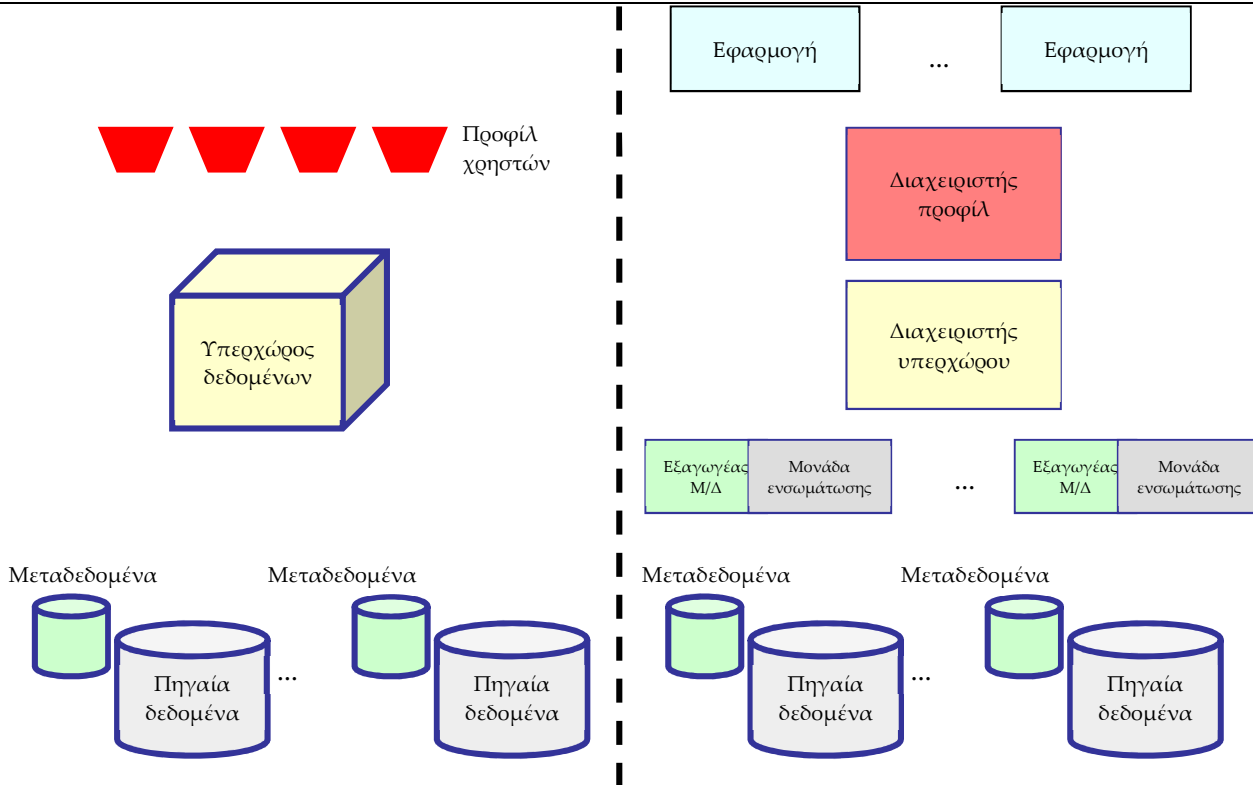
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα



Σχήμα 1. Αριστερά: βασικές δομές του υπερχώρου. Δεξιά: κομμάτια λογισμικού που απαρτίζουν ένα σύστημα διαχείρισης υπερχώρων δεδομένων

Το έργο θα στηριχθεί στη γενική δομή ενός υπερχώρου δεδομένων που φαίνεται στο σχήμα 1. Ένας υπερχώρος οργανώνει την πληροφορία σε ένα πολυδιάστατο μοντέλο. Κάθε κομμάτι πληροφορίας, κάθε δεδομένο (ή ομάδα δεδομένων) χαρακτηρίζεται από κάποιες «συντεταγμένες τιμών» πάνω στις διαφορετικές διαστάσεις του υπερχώρου. Ο χαρακτηρισμός αυτός αποτελεί την μεταπληροφορία ή μεταδεδομένα της πληροφορίας και επιτρέπει την αρμονική ενσωμάτωσή της στον υπερχώρο. Κάθε φορά που κάποιος χρήστης ενσωματώνει νέα δεδομένα, ένα τμήμα λογισμικού υπεύθυνο για την εξαγωγή μεταδεδομένων αναλαμβάνει να υποστηρίξει τη διαδικασία αυτή. Επιπλέον, κάθε χρήστης διαθέτει ένα προσωπικό προφίλ με προτιμήσεις, ιστορικό αναζητήσεων, προσωρινά αποθηκευμένες απαντήσεις και συνθέσεις επί μέρους αποτελεσμάτων. Η διαδικασία επερώτησης χρησιμοποιεί το εν λόγω προφίλ για να εξατομικεύσει την απάντηση, ενδεχομένως να ενημερώσει προηγουμένως αποθηκευμένα αποτελέσματα αν υπάρχουν αλλαγές στα πηγαία δεδομένα και να παρουσιάσει τα αποτελέσματα ανάλογα με τις προτιμήσεις του χρήστη. Ένας διαχειριστής προφίλ είναι υπεύθυνος για την χρήση και ενημέρωση των προφίλ των χρηστών. Στο κέντρο της αρχιτεκτονικής βρίσκεται ένας διαχειριστής του υπερχώρου που συντονίζει τις επί μέρους λειτουργίες. Το οικοσύστημα διαχείρισης της πληροφορίας συμπληρώνεται από εφαρμογές που εκτελούνται πάνω από τον υπερχώρο και εξαρτώνται από αυτόν.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης

ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

Για να μπορέσει ο διαχειριστής του υπερχώρου να λειτουργήσει, πρέπει να υποστηρίζεται από τη διαδικασία ενσωμάτωσης δεδομένων και από τη διαδικασία εξατομίκευσης της επερώτησης. Για πρακτικούς λόγους, η διαδικασία ολοκλήρωσης και επερώτησης παρουσιάζει προβλήματα κλιμάκωσης λόγω του όγκου των δεδομένων και των χρηστών και έχει ανάγκη από μη παραδοσιακές τοπολογίες εκτέλεσης σε συστημικό επίπεδο. Για λόγους προσαρμογής, το όλο οικοσύστημα εφαρμογών και υπερχώρου πρέπει στη διάρκεια του χρόνου να μπορεί να αλλάζει εσωτερική δομή, μεγέθη σε χρήστες, μεταπληροφορία και δεδομένα με τρόπο που επιφέρει την ελάχιστη δυνατή απαιτούμενη αναπροσαρμογή.

Τα ερευνητικά και πρακτικά ζητήματα που προκύπτουν, μπορούν να ομαδοποιηθούν, λοιπόν στα παρακάτω θεμελιώδη ερωτήματα:

1. Πώς μπορούμε να επιτρέψουμε στους χρήστες να εντάσσουν εύκολα νέα δεδομένα στο σύστημα – ενδεχομένως σε μεγάλους όγκους – και μάλιστα με έμφαση σε μη παραδοσιακά δεδομένα;
2. Πώς μπορούμε να συνθέσουμε επί μέρους όψεις της πληροφορίας και από πλευράς χρηστών και από πλευράς υπαρχόντων δεδομένων σε ένα ενιαίο υπερχώρο δεδομένων που να επιτρέπει την μετέπειτα επερώτηση της πληροφορίας;
3. Πώς μπορούμε να επιτρέψουμε στους χρήστες να επερωτούν το σύστημα με τρόπο διαισθητικό και εύκολο, να συνθέτουν επί μέρους απαντήσεις σε νέα αποτελέσματα και να εξειδικεύουν την κατάταξη των αποτελεσμάτων με τρόπο ανάλογο προς τις προτιμήσεις τους;
4. Πώς μπορούμε να κατανείμουμε την τοπολογία λειτουργίας ενός υπερχώρου / οικοσυστήματος δεδομένων ώστε να επιτρέψουμε την ομαλή κλιμάκωση από πλευράς χρηστών και δεδομένων; Επιπλέον, πώς μπορούμε να υποστηρίξουμε την προσωποπαγή διαχείριση της πληροφορίας, να αναβαθμίσουμε τους χρήστες σε πρωταγωνιστές και μέσω της συνεργατικότητάς τους να επιτρέψουμε τη σημασιολογική συνένωση διάσπαρτων οικοσυστημάτων δεδομένων; Πώς μπορούμε, δηλαδή, να αναδείξουμε τη «σοφία» της κοινότητας χρηστών (community wisdom) και να την εκμεταλλευτούμε κατάλληλα; Τί είδους νέα μοντέλα δεδομένων και μηχανές επεξεργασίας ερωτημάτων απαιτούνται;
5. Πώς μπορούμε να υποστηρίξουμε την ομαλή εξέλιξη της πληροφορίας στην πορεία του χρόνου από πλευράς σημασιολογίας, δομής, περιεχομένου και όγκου με την ελάχιστη δυνατή επίπτωση σε χρήστες και εφαρμογές;



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

3. Πορεία Υλοποίησης του Υποέργου

Σε ότι αφορά το χρονικό διάστημα που καλύπτεται από την παρούσα έκθεση, οι κύριες δραστηριότητες του υποέργου εστιάζονται και στις 5 ερευνητικές Δράσεις καθώς επίσης και στην 6η Δράση, η οποία αφορά στη διαχείριση του έργου. Οι δραστηριότητες αυτές περιγράφονται πιο αναλυτικά στη συνέχεια.

Δράση 1: ΑΝΑΚΤΗΣΗ ΚΑΙ ΑΝΑΠΑΡΑΣΤΑΣΗ ΜΗ ΠΑΡΑΔΟΣΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ

Η Δράση 1 «Αναπαράσταση και ενσωμάτωση μη-παραδοσιακών δεδομένων» έχει σκοπό να παρέχει την διεπαφή του συστήματος με τα μη παραδοσιακά δεδομένα. Σε αυτό το πλαίσιο αναγνωρίζονται τρεις βασικές επιδιώξεις οι οποίες αποτελούν τις 3 υποδράσεις της συγκεκριμένης δράσης: α) η δημιουργία ενός επεκτάσιμου και γενικού μοντέλου για μη παραδοσιακά δεδομένα, β) η σχεδίαση τεχνικών ανάκτησης πληροφορίας που θα έχουν την δυνατότητα να εκμαιεύουν την απαραίτητη μεταπληροφορία για την ένταξη των δεδομένων στο υπερχώρο δεδομένων και γ) η ανάπτυξη μηχανισμών δεικτοδότησης των μη παραδοσιακών δεδομένων και της μεταπληροφορίας που τα συνοδεύει, οι οποίοι θα έχουν την δυνατότητα να υποστηρίξουν γενικού τύπου ερωτήσεις που θα παρέχονται στον υπερχώρο των δεδομένων. Η Δράση 1 έχει διάρκεια από 1/2/12 έως το τέλος του έργου, 31/7/15.

Στη διάρκεια του 2012, πραγματοποιήθηκαν εργασίες στα πλαίσια της Δράσης 1 που αφορούσαν στον προσδιορισμό διαθέσιμων πηγών μη παραδοσιακών δεδομένων και ειδικότερα προκρίθηκαν τα διασυνδεδεμένα δεδομένα στον Ιστό, τα επιστημονικά δεδομένα και τα γεωχωρικά δεδομένα. Συγκεκριμένα, έγιναν 2 δημοσιεύσεις: η πρώτη περιλαμβάνεται στο Π1.3 και αφορά στην κατασκευή μηχανισμών δεικτοδότησης μη παραδοσιακών δεδομένων με αλφαριθμητική αναπαράσταση (1). Η δεύτερη περιλαμβάνεται στο Π1.1 και αφορά σε δημοσίευση μη-παραδοσιακών δεδομένων συνόλων (όπως ιατρικών, ιστορικών, ημερολογίων πιστωτικών καρτών και επίσκεψης ιστοτόπων) με τρόπο που θα διασφαλίζει την ανωνυμία τους (2).

Στη διάρκεια του 2013, δημοσιεύτηκε μια εργασία στα πλαίσια της Δράσης 1 η οποία περιλαμβάνεται στο Π1.2. Η εργασία προτείνει μία πρωτότυπη οντολογία για τον εμπλουτισμό χωρο-χρονικών δεδομένων τα οποία αναπαριστούν τροχιές κινούμενων αντικειμένων, με επισημειώσεις από ελεύθερα διασυνδεδεμένα δεδομένα (3). Η οντολογία αυτή αποτελεί μέρος ενός συστήματος που ανακτά χωρο-χρονικά δεδομένα πλούσια σε μετα-δεδομένα, δρώντας ουσιαστικά ως ο ενδιάμεσος κρίκος μεταξύ γλωσσών επερωτήσεων για διασυνδεδεμένα και σημασιολογικά επαυξημένα δεδομένα κίνησης.

Δράση 2: ΟΛΟΚΛΗΡΩΣΗ ΠΑΡΑΔΟΣΙΑΚΩΝ ΚΑΙ ΜΗ ΔΕΔΟΜΕΝΩΝ, ΠΛΟΗΓΗΣΗ ΚΑΙ ΑΝΑΖΗΤΗΣΗ

Η Δράση 2 «Ολοκλήρωση παραδοσιακών και μη δεδομένων, πλοήγηση και αναζήτηση» έχει σκοπό να παρέχει θεωρητική θεμελίωση, αλγοριθμικές τεχνικές και αρχιτεκτονικές λογισμικού για τον ορισμό υπερχώρων δεδομένων και την αξιοποίηση παραδοσιακών και μη δεδομένων σε τέτοια περιβάλλοντα.

Η Δράση 2 έχει διάρκεια από 1/6/12 έως 31/7/15 και οργανώνεται σε τρεις θεμελιώδεις υποδράσεις, εκ των οποίων η πρώτη αφορά στον ορισμό του εννοιολογικού μοντέλου αναπαράστασης υπερχώρων, η



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
πρόγραμμα για την ανάπτυξη

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

δεύτερη την περιγραφή του μηχανισμού ενσωμάτωσης νέων πηγών σε έναν υπερχώρο και την εύρεση αντιστοιχίσεων μεταξύ ετερογενών πηγών δεδομένων και τέλος η τρίτη αφορά στην αρχιτεκτονική και τους μηχανισμούς που θα πρέπει να διαθέτει ένα σύστημα υποστηρίξης υπερχώρων δεδομένων για την εφαρμογή επερωτήσεων και την ανάκτηση πληροφορίας από αυτό.

Στη διάρκεια του 2012, οι εργασίες που έγιναν στα πλαίσια της Δράσης 2 αφορούσαν στον ορισμό ενός γραφοθεωρητικού μοντέλου υπερχώρων δεδομένων. Δημιουργήθηκε μια τεχνική αναφορά που θα αποτελέσει μέρος του Π2.1 με το reference example που θα χρησιμοποιηθεί από όλες τις δράσεις.

Στη διάρκεια του 2013, έγιναν οι εξής δραστηριότητες στα πλαίσια της Δράσης 2:

1. Μια δημοσίευση που περιλαμβάνεται στο Π2.1 και παρουσιάζει μια μεθοδολογία για την ενσωμάτωση πολυδιάστατων στατιστικών δεδομένων σε υπερχώρους δεδομένων (4).
2. Μια δημοσίευση που περιλαμβάνεται στο Π2.3 και προτείνει μηχανισμούς επεξεργασίας επερωτήσεων για δομημένα δεδομένα έτσι ώστε να επιτυγχάνεται διαφοροποίηση των αποτελεσμάτων αναζήτησης (5).
3. Μια δημοσίευση που περιλαμβάνεται στο Π2.2 και παρουσιάζει ένα σύνολο αλγορίθμων για την επιτάχυνση της διαδικασίας ενσωμάτωσης μεγάλων συνόλων δεδομένων σε έναν υπερχώρο (6).
4. Μια συμμετοχή σε ένα θερινό σχολείο με αντικείμενο την μελέτη των αλγορίθμων για τη διαχείριση και επερώτηση σύνθετων δομών δεδομένων. Οι τεχνικές που εξετάστηκαν είναι χρήσιμες για τη σχεδίαση και την ανάπτυξη αλγορίθμων για την αποδοτική ενσωμάτωση μεγάλων συνόλων δεδομένων σε υπερχώρους.
5. Μια σειρά συναντήσεων με καθηγητές από το Πανεπιστήμιο του Τρέντο με θέμα τον σχεδιασμό ενός μεταμοντέλου για την ευέλικτη περιγραφή δεδομένων σε υπερχώρους, τη δημιουργία τελεστών για τη διαχείριση και τον μετασχηματισμό στιγμιότυπων ενός υπερχώρου, και την ανάπτυξη αλγορίθμων για εξαγωγή αντιστοιχίσεων από στιγμιότυπα του μεταμοντέλου που περιγράφουν διάφορα δεδομένα.

Δράση 3: CONTEXTUALIZATION ΚΑΙ ΕΞΑΤΟΜΙΚΕΥΣΗ ΠΕΡΙΕΧΟΜΕΝΟΥ ΚΑΙ ΠΑΡΟΥΣΙΑΣΗΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

Η Δράση 3 «Contextualization και εξατομίκευση περιεχομένου και παρουσίασης της πληροφορίας» αποσκοπεί στον ορισμό γενικευμένων κι εκφραστικών μοντέλων και στην εξαγωγή αλγοριθμικών αποτελεσμάτων για την παροχή εξατομικευμένων και προσαρμοσμένων υπηρεσιών αναζήτησης σε υπερχώρους δεδομένων, καθώς και στο γενικότερο σχεδιασμό ενός ισχυρού συστήματος εξατομίκευσης, ικανού να ανταποκριθεί και να ικανοποιήσει τις διαρκώς αυξανόμενες και μεταβαλλόμενες απαιτήσεις πληθώρας εφαρμογών.

Η Δράση 3 έχει διάρκεια από 1/2/12 έως 31/7/15 και οργανώνεται σε τρεις θεμελιώδεις υποδράσεις, εκ των οποίων η πρώτη αφορά τη διαχείριση και τη συντακτική και σημασιολογική ολοκλήρωση των δεδομένων για τη συμπεριφορά και το περιβάλλον του χρήστη, ενώ η δεύτερη αφορά τη μοντελοποίηση του χρήστη και τη διαχείριση των προφίλ χρηστών, τα οποία περιέχουν πληροφορία για της προτιμήσεις και τα χαρακτηριστικά τους. Τέλος, η τρίτη αφορά τον ευρύτερο σχεδιασμό ενός γενικευμένου



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα συστήματος εξατομίκευσης, έχοντας υπόψη τα διαθέσιμα μοντέλα και τις ορισμένες τεχνικές από τις δύο προηγούμενες δράσεις .

Στη διάρκεια του 2012, οι ενέργειες που έγιναν στα πλαίσια της Δράσης 3 αφορούσαν στη μελέτη της βιβλιογραφίας και στον προσδιορισμό των μεθόδων και των τεχνικών εξατομίκευσης, ανάλογα με το είδος μη παραδοσιακών δεδομένων.

Στη διάρκεια του 2013 έγιναν οι εξής ενέργειες στα πλαίσια της Δράσης 3:

1. Μια δημοσίευση που περιλαμβάνεται στο Π3.1 και προτείνει ένα πρωτότυπο τρόπο διεπαφής του χρήστη με ένα πληροφοριακό σύστημα έτσι ώστε να διευκολύνεται η πλοήγηση του στη βάση δεδομένων (7). Η διεπαφή ενσωματώνει ένα αλγόριθμο που υπολογίζει και προτείνει στους χρήστες δεδομένα που δεν ανήκουν στο αποτέλεσμα του αρχικού τους ερωτήματος, αλλά σχετίζονται στενά με αυτό ("You May Also Like" αποτελέσματα). Έτσι, ο χρήστης ξεκινά από ένα βασικό ερώτημα και με τη βοήθεια του συστήματος συστάσεων μπορεί να διατυπώνει ερωτήματα με βάση τις προσωπικές του προτιμήσεις και τις ατομικές πληροφοριακές του ανάγκες.
2. Μια δημοσίευση που περιλαμβάνεται στο Π3.2 και παρέχει μια αρχική θεμελίωση για πρότυπα πρότασης συμπερασμάτων προς το χρήστη, τα οποία στο μέλλον μπορεί να βασίζονται σε πρότυπα συμπεριφοράς ή προτιμήσεων (8). Η μεθοδολογία αυτή συνεισφέρει στην υλοποίηση εργαλείων διαχείρισης της εξέλιξης καθώς τα αποτελέσματα εντάχθηκαν στο εργαλείο Cinecubes.
3. Στα πλαίσια του Π3.3, ξεκίνησε η ανάπτυξη ενός πρωτότυπου συστήματος εξατομίκευσης, το οποίο υποστηρίζει ένα ενιαίο τρόπο μοντελοποίησης του προφίλ των χρηστών και επιτρέπει την ενσωμάτωση όλων των πληροφοριών και κινήσεων που έχει για ένα χρήστη από όλες τις διαθέσιμες εφαρμογές. Επιπροσθέτως, διαθέτει ένα σύνολο τελεστών που επιτρέπει τη δημιουργία αλγορίθμων που εξυπηρετούν διαφορετικές ανάγκες, αλλά μπορούν να συνδυαστούν αποδοτικά κάτω από ένα ενιαίο πλαίσιο ροής επεξεργασίας δεδομένων.

Δράση 4: ΚΑΤΑΝΕΜΗΜΕΝΕΣ ΥΠΟΔΟΜΕΣ ΑΠΟΘΗΚΕΥΣΗΣ, ΠΡΟΣΠΕΛΑΣΗΣ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗΣ ΔΕΔΟΜΕΝΩΝ

Ο υπερχώρος δεδομένων στον οποίο εστιάζει η πρότασή μας, εγγενώς περιλαμβάνει ένα κατανεμημένο υπολογιστικό περιβάλλον, όπου οικοσυστήματα δεδομένων είναι γεωγραφικά κατανεμημένα, όπως άλλωστε είναι και οι χρήστες του υπερχώρου. Η Δράση 4 «Κατανεμημένες Υποδομές Αποθήκευσης, Προσπέλασης και Διαχείρισης Δεδομένων» σκοπό έχει να παράσχει αρχιτεκτονικές και αλγορίθμους οι οποίοι είτε οι ίδιες θα παρέχουν τρόπους για την κατανεμημένη οργάνωση χώρων δεδομένων στα χαμηλότερα επίπεδα του συστήματος, είτε θα παρέχουν κατάλληλα στατιστικά στοιχεία (όπως συνόψεις δεδομένων που περιγράφουν το φόρτο του συστήματος, ή το ποια δεδομένα ζητούνται από ποιούς χρήστες, τι ρόλο παίζουν διάφοροι χρήστες, κλπ) που να υποβοηθούν την οργάνωση αυτή. Επιπλέον, στόχος είναι οι ίδιοι οι χρήστες να αναβαθμιστούν από απλοί καταναλωτές πληροφορίας σε πρωταγωνιστές που μέσω της συνεργατικότητάς τους να βοηθούν στην κατανόηση των δεδομένων, της σχέσης τους και στην εξατομίκευση των αποτελεσμάτων με βάση το ποιός ερωτά και την «κοινωνική σοφία» αναφορικά με τα περιεχόμενα του οικοσυστήματος. Η δυνατότητα να συλλάβουμε, πέραν των



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

παραπάνω χαρακτηριστικών, την αβεβαιότητα των αποθηκευμένων δεδομένων στις παραπάνω συνόψεις, μας προσφέρει ένα επιπλέον εργαλείο για τη δυνατότητα σύνοψης «πολλαπλών εκδοχών της αλήθειας» (για δεδομένα με ασάφεια), και να μπορεί να παρέχει προσεγγιστικές απαντήσεις στις ερωτήσεις του χρήστη, αν δεν έχει τις ακριβείς, ή αν ο ρυθμός άφιξης των δεδομένων (π.χ. ροές δεδομένων) είναι εξαιρετικά υψηλός. Επιπλέον, οι απαντήσεις δυνητικά θα εξαρτώνται από τα χαρακτηριστικά του ερωτώντα χρήστη, τις παρεμβάσεις άλλων (με διαφορετικές διαβαθμίσεις) χρηστών στο χαρακτηρισμό των δεδομένων, κλπ., που συνεισφέρουν άλλη μια διάσταση στις «πολλαπλές εκδοχές της αλήθειας». Η Δράση 4 έχει διάρκεια από 1/6/12 έως τις 31/7/15.

Στη διάρκεια του 2012, δημοσιεύτηκαν οι εξής εργασίες στα πλαίσια της Δράσης 4:

1. Μια εργασία που περιλαμβάνεται στο Π4.1 και προτείνει ένα πλαίσιο για τη συλλογή και την επεξεργασία πληροφορίας από τους χρήστες (9). Συγκεκριμένα, ορίζει τη διεπαφή που οι τελευταίοι έχουν με το σύστημα και προτείνει έναν αλγόριθμο για τη δημιουργία μιας δομής δεικτοδότησης. Η δομή αυτή είναι αποτέλεσμα συλλογικής εργασίας, στα πλαίσια της κοινότητας των χρηστών και αποτυπώνει τη γενική αντίληψη που έχει η κοινότητα ως προς το φυσικό περιβάλλον γύρω της. Με τον τρόπο αυτό η εργασία αυτή εξυπηρετεί τους στόχους της υποδράσης 1 που είναι η διαχείριση δεδομένων που προέρχονται από μια ομάδα χρηστών καταναμημένων σε διαφορετικές γεωγραφικά περιοχές και συγκεκριμένα η επεξεργασία των δεδομένων αυτών μέσω ανθρωποκεντρικών τεχνικών μέσα σε περιβάλλοντα συνεργασίας.
2. Μια εργασία που περιλαμβάνεται στο Π4.2 και προτείνει μια μεθοδολογία για την απάντηση πολλαπλών επερωτήσεων κορυφογραμμής σε μεγάλο όγκο δεδομένα, δεδομένων κάποιων προτιμήσεων χρηστών (10). Ο συγκεκριμένος αλγόριθμος είναι πολύ πιο αποδοτικός από προηγούμενες τεχνικές για μεγάλης κλίμακας δεδομένα και έχει κατασκευαστεί ώστε να μπορεί να εκτελεστεί παράλληλα και καταναμημένα.
3. Μια εργασία που περιλαμβάνεται στο Π4.3 και παρουσιάζει μεθόδους εύρεσης ομοιότητας χρηστών υπερχώρων δεδομένων (11). Για την ακρίβεια, προτείνει ένα καινοτόμο τρόπο εκτίμησης της ομοιότητας χρηστών σε μια αποθήκη δεδομένων χρησιμοποιώντας μια καθιερωμένη τεχνική παραγωγής συνόψεων δεδομένων. Συνδυάζεται μάλιστα με μια τεχνική κατακερματισμού δεδομένων έτσι ώστε να είναι δυνατή μια αρκετά ακριβής και σε μικρό χρόνο εκτίμηση της ομοιότητας χρηστών σε ένα εξαιρετικά μεγάλο υπερχώρο δεδομένων. Με τον τρόπο αυτό, δημιουργείται ένα σχήμα ευρετηρίασης των δεδομένων σε δύο επίπεδα και επιτυγχάνεται ο εντοπισμός ακραίων τιμών-έκτοπων χρηστών (outliers).

Στη διάρκεια του 2013, εκπονήθηκαν δυο διπλωματικές εργασίες στα πλαίσια της Δράσης 4, οι οποίες περιλαμβάνονται στο Π4.2 σαν εκτεταμένες περιλήψεις. Η πρώτη εργασία εξετάζει την τεχνική Locality Sensitive Hashing (LSH), που παρέχει ένα αποδοτικό μηχανισμό για τη σύνοψη δεδομένων, επιτυγχάνοντας μια πιθανοτική μείωση της διαστασιμότητας τους. Για την ακρίβεια, προτείνει μια παράλληλη και καταναμημένη υλοποίηση της παρασκευής συνόψεων με LSH, κάνοντας χρήση του συστήματος απεικόνισης-μείωσης (Map-Reduce) Apache Hadoop. Η δεύτερη εργασία εξετάζει τα ιστογράμματα, τα οποία αποτελούν τον πιο ευρέως χρησιμοποιούμενο και αποτελεσματικό μηχανισμό σύνοψης δεδομένων. Συγκεκριμένα, η εργασία μελέτησε τους αλγόριθμους κατασκευής ιστογραμμάτων για ντερμινιστικά δεδομένα ώστε να καταστεί δυνατή η προσαρμογή τους για πιθανοτικά δεδομένα.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΣΠΑ
2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

Δράση 5: ΥΠΟΣΤΗΡΙΞΗ ΕΞΕΛΙΞΗΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΑΥΤΟ-ΡΥΘΜΙΣΗΣ ΣΥΣΤΗΜΑΤΩΝ

Η Δράση 5 «Υποστήριξη εξέλιξης της πληροφορίας και αυτορύθμισης συστημάτων» έχει σκοπό να παρέχει αλγοριθμικά αποτελέσματα για τη σχεδίαση, προσαρμογή ενός υπερχώρου δεδομένων και του περιρρέοντος οικοσυστήματος δεδομένων του σε ότι αφορά εξελικτικές μεταβολές που αφορούν στη σημασιολογία, δομή και περιεχόμενο της πληροφορίας. Η Δράση 5 έχει διάρκεια από 1/6/12 έως τις 31/7/15 και οργανώνεται σε τρεις θεμελιώδεις υποδράσεις, εκ των οποίων η πρώτη αφορά τη μοντελοποίηση των γεγονότων εξέλιξης και των αλληλεξαρτήσεών τους, η δεύτερη αφορά την ρύθμιση και αυτοματοποίηση της εξέλιξης ενός οικοσυστήματος και η τρίτη την α-πριορι σχεδιάσή του, έχοντας υπόψη τις διαθέσιμες τεχνικές προσαρμογής από την προηγούμενη υποδράση.

Στη διάρκεια του 2012, δημοσιεύθηκαν οι εξής εργασίες στα πλαίσια της Δράσης 5:

1. Μια εργασία που περιλαμβάνεται στο Π5.1 και ασχολείται με τη διαχείριση γεγονότων εξέλιξης σε συστήματα διαχείρισης δεδομένων (12). Συγκεκριμένα, προτείνει μια μεθοδολογία που βασίζεται στο γράφημα αρχιτεκτονικής για να καθορίσει πώς θα διαδοθούν οι αλλαγές σε ένα οικοσύστημα δεδομένων.
2. Μια εργασία που περιλαμβάνεται στο Π5.3 και εστιάζει στον ορισμό μετρικών ποιότητας για την αποτίμηση της εξέλιξης πολυδιάστατων υπερχώρων και οικοσυστημάτων δεδομένων όπως είναι οι αποθήκες και οι ροές δεδομένων (ETL) (13). Επιπλέον, αξιολογεί τις μετρικές αυτές σε πραγματικά συστήματα αποθηκών δεδομένων και προτείνει ένα σύνολο από «καλές πρακτικές» για την αποτίμηση και σχεδίαση οικοσυστημάτων δεδομένων με στόχο την διαχείριση της εξέλιξής τους στο χρόνο.

Στη διάρκεια του 2013, δημοσιεύθηκαν οι εξής εργασίες στα πλαίσια της Δράσης 5:

1. Μια εργασία που περιλαμβάνεται στο Π5.1 και προτείνει μια μεθοδολογία για την καταγραφή της εξέλιξης οικοσυστημάτων δεδομένων με τη βοήθεια μιας γραφικής αναπαράστασης που επιτρέπει την αναζήτηση της προέλευσης των δεδομένων και της ανάκτησης προηγούμενων εκδόσεων τους (14). Η εργασία αναλύει πώς η συγκεκριμένη γραφική αναπαράσταση δημιουργείται και ενσωματώνει κάθε φορά τις αλλαγές εξέλιξης που συμβαίνουν στην τρέχουσα έκδοση δεδομένων.
2. Μια εργασία που περιλαμβάνεται στο Π5.2 και προτείνει μια μέθοδο για την επανεγγραφή οικοσυστημάτων δεδομένων με αλγοριθμικό, πλήρως αυτοματοποιημένο τρόπο, όταν τα οικοσυστήματα είναι στη θεμελιώδη μορφή που περιλαμβάνει βάσεις δεδομένων και ερωτήσεις (αλλά όχι κείμενα ή RDF γράφηματα) (15).
3. Μια εργασία που περιλαμβάνεται στο Π5.3 και εξετάζει την εξέλιξη και ενημέρωση υπερχώρων βασισμένη σε ροές δεδομένων. Για την ακρίβεια, προτείνει μεθόδους για την επιτυχή αναδιάταξη του υπερχώρου με αυτοματοποιημένο τρόπο μέσα από την εφαρμογή αποδοτικών αλγορίθμων βελτιστοποίησης έτσι ώστε να μειωθεί ικανοποιητικά ο φόρτος εργασίας μέσα σε αυτόν (16).



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
πρόγραμμα για την ανάπτυξη

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

Δράση 6: ΣΥΝΤΟΝΙΣΜΟΣ – ΠΑΡΑΚΟΛΟΥΘΗΣΗ ΦΥΣΙΚΟΥ ΚΑΙ ΟΙΚΟΝΟΜΙΚΟΥ ΑΝΤΙΚΕΙΜΕΝΟΥ

Στα πλαίσια της Δράσης 6 πραγματοποιήθηκε η εναρκτήρια συνάντηση του έργου στην Αθήνα το Φεβρουάριο του 2012, η 2^η συνάντηση όλων των μελών της ΚΕΟ στα Χανιά τον Ιούνιο του 2012, συνάντηση των μελών της Δράσης 1 και 2 στην Αθήνα τον Ιούλιο 2012 και η 3^η συνάντηση όλων των μελών της ΚΕΟ το Δεκέμβριο στην Αθήνα. Επίσης δημιουργήθηκαν 2 sites για το έργο: το site διάχυσης (<http://web.imis.athena-innovation.gr/projects/eicos/en/index.html>) και το collaboration site του έργου. Τον Μάρτιο του 2013 έγινε η 4^η συνάντηση στην Αθήνα, στα πλαίσια της οποίας πραγματοποιήθηκε ανοιχτό σεμινάριο (με τη συμμετοχή φοιτητών) γύρω από το ερευνητικά αντικείμενα του έργου. Για κάθε συνάντηση υπάρχουν αναλυτικά πρακτικά.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
πρόγραμμα για την ανάπτυξη

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

4. Προβλήματα κατά τη διάρκεια Υλοποίησης του Υποέργου

Η πρόοδος του φυσικού αντικείμενου εντοπίστηκε κυρίως στην προετοιμασία των εργασιών που θα πραγματοποιήσει η κάθε ερευνητική ομάδα, στον προσδιορισμό των βασικών ερευνητικών κατευθύνσεων. Υπήρξε μια καθυστέρηση της συγκρότησης της ΟΕΣ λόγω της αναμονής διευκρινήσεων από πλευράς της Διαχειριστικής Αρχής σχετικά με την επιλεξιμότητα των δαπανών για υποτροφίες σε υπ. Διδάκτορες – μέλη της ΟΕΣ.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

5. Ανακεφαλαίωση και Μελλοντικές Εργασίες

Ανακεφαλαιώνοντας, μπορούμε να πούμε ότι μέχρι στιγμής *όλα τα Πακέτα Εργασίας είναι εντός των χρονικών ορίων που είχαν προβλεφθεί και σε ότι αφορά την πρόοδο της εργασίας του και σε ότι αφορά τα παραδοτέα τους.*

Σε ότι αφορά τις μελλοντικές ενέργειες, το έργο θα συνεχιστεί όπως έχει προβλεφθεί. Συγκεκριμένα:

- Στα πλαίσια της Δράσης 1 θα συνεχιστούν οι εργασίες μελέτης τεχνικών ανάκτησης πληροφορίας και μεταπληροφορίας για διαφορετικά είδη δεδομένων.
- Στα πλαίσια της Δράσης 2 θα μελετηθούν τεχνικές ανίχνευσης αντιστοιχίσεων και ολοκλήρωσης δεδομένων, λαμβάνοντας υπόψη τόσο τη δομή του κάθε συνόλου δεδομένων που εντάσσεται σε ένα υπερχώρο όσο και τη σημασιολογία των δεδομένων. Ιδιαίτερη έμφαση θα δοθεί στις τεχνικές meta-blocking, που στοχεύουν να αυξήσουν τη απόδοση (efficiency) της διαδικασίας ενσωμάτωσης με μικρό κόστος στην αποτελεσματικότητα της (effectiveness), αλλά και στην ανάπτυξη μεθόδων για τη σταδιακή (incremental) ενσωμάτωση υπερχώρων.
- Στα πλαίσια της Δράσης 3, θα μελετηθούν μοντέλα εξατομίκευσης και ειδικότερα τεχνικές μοντελοποίησης της συμπεριφοράς χρήστη των χαρακτηριστικών του προφίλ και τεχνικές συλλογής και διαχείρισης των προτιμήσεων του.
- Στα πλαίσια της Δράσης 4 θα συνεχίσουν οι εργασίες για την διαχείριση πληροφορίας στο cloud καθώς και θα μελετηθούν τεχνικές παραγωγής στατιστικών και συνόψεων δεδομένων για την αποδοτικότερη ανάκτηση της πληροφορίας από ένα υπερχώρο δεδομένων.
- Στα πλαίσια της Δράσης 5, θα μελετηθούν τεχνικές μοντελοποίησης, διαχείρισης και μετάδοσης αλλαγών εξέλιξης σε υπερχώρους δεδομένων, καθώς και τεχνικές αποτίμησης της ποιότητας σχεδιασμού οικοσυστημάτων πληροφορίας με βάση τη δυνατότητα προσαρμογής τους σε αλλαγές εξέλιξης.
- Στα πλαίσια της Δράσης 6 θα πραγματοποιηθεί μια συνάντηση όλων των μελών της ΚΕΟ, ενέργειες διάχυσης των αποτελεσμάτων του έργου σε συνέδρια που θα πραγματοποιηθούν το προσεχές εξάμηνο σχετικά με τη διαχείριση δεδομένων και τα πληροφοριακά συστήματα, ενώ τέλος θα διοργανωθεί ένα 2^ο ανοιχτό σεμινάριο με συμμετοχή φοιτητών - προπτυχιακών \ μεταπτυχιακών - με στόχο την παρουσίαση των μέχρι τώρα ερευνητικών αποτελεσμάτων του έργου.

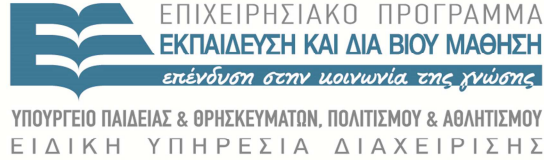
Με τιμή

Ο Επιστημονικός Υπεύθυνος του Υποέργου

Καθ. Ιωάννης Ιωαννίδης
Πρόεδρος «ΑΘΗΝΑ»



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
πρόγραμμα για την ανάπτυξη

ΕΙΚΟΣ : Θεωρητική Και Αλγοριθμική Θεμελίωση για Προσωποκεντρικά Συνεργατικά Πληροφοριακά Συστήματα

6. Βιβλιογραφία

1. *ERa efficient serial and parallel suffix tree construction for very long strings*. **E. Mansour, A. Allam, S. Skiadopoulos, and P. Kalnis**. 2011. PVLDB. σσ. 49–60.
2. *Privacy preservation by disassociation*. **M. Terrovitis, J. Liagouris, N. Mamoulis, and S. Skiadopoulos**. 2012.
3. *Baquara: A Holistic Ontological Framework for Movement Analysis with Linked Data*. **R. Filletto, M. Krueger, N. Pelekis, Y. Theodoridis, C. Renso**. s.l. : ER, 2013.
4. *Publishing Census as Linked Open Data. A Case Study*. **I. Petrou, G. Papastefanatos, T. Dalamagas**. s.l. : 1st Workshop on Open Data, 2013.
5. *RDivF: Diversifying Keyword Search on RDF Graphs*. **N. Bikakis, G. Giannopoulos, J. Liagouris, D. Skoutas, T. Dalamagas, T. Sellis**. s.l. : TPDFL, 2013.
6. **G. Papadakis, G. Koutrika, T. Palpanas, W. Nejdl**. Meta-Blocking: Taking Entity Resolution to the Next Level . *TKDE*.
7. **M. Drosou, E. Pitoura**. YmalDB: exploring relational databases via result-driven recommendations. *The VLDB Journal*. 2013.
8. *CineCubes: Cubes As Movie Stars with Little Effort*. **D. Gkesoulis, P. Vassiliadis**. s.l. : DOLAP, 2013.
9. *Crowdsourcing Taxonomies*. **D. Karampinas, P. Triantafillou**. s.l. : ESWC, 2012.
10. *Efficient Influence-Based Processing of Market Research Queries*. **A. Arvanitis, A. Deligiannakis, Y. Vassiliou**. s.l. : CIKM, 2012.
11. *Towards Enabling Outlier Detection in Large, High Dimensional Data Warehouses*. **K. Georgoulas, Y. Kotidis**. s.l. : SSDBM, 2012.
12. *Propagating Evolution Events in Data-Centric Software Artifacts*. **G. Papastefanatos, P. Vassiliadis, A. Simitis**. s.l. : 3rd Workshop on Hot Topics in Software Upgrades, 2011.
13. **G. Papastefanatos, P. Vassiliadis, A. Simitis, Y. Vassiliou**. Metrics for the Prediction of Evolution Impact in ETL Ecosystems: A Case Study. *Journal on Data Semantics*. 2012.
14. *Capturing the History and Change Structure of Evolving Data*. **G. Papastefanatos, Y. Stavrakas, T. Galani**. s.l. : 5th International Conference on Advances in Databases, Knowledge, and Data Applications , 2013.
15. *Automating the Adaptation of Evolving Data-Intensive Ecosystems*. **P. Manousis, P. Vassiliadis, G. Papastefanatos**. s.l. : ER, 2013.
16. *Declarative expression and optimization of data-intensive* . **G. Kougka, A. Gounaris**. s.l. : DaWaK, 2013.