

Εργασία: **MITOS / grOOGLE 2008**

Γενική Περιγραφή

Έκδοση 0.4

18 Μαρτίου 2008

Γενικός στόχος αυτής της εργασίας είναι η εξοικείωση, βελτίωση και επέκταση της μηχανής αναζήτησης mitos/GRoogle (<http://google.csd.uoc.gr:8080/google/>) που κατασκευάστηκε τα δύο προηγούμενα χρόνια από τους σπουδαστές που παρακολούθησαν αυτό το μάθημα. Κάποιες από τις λειτουργικές μονάδες πρέπει να αντικατασταθούν, κάποιες να διορθωθούν, άλλες να βελτιωθούν και σίγουρα μπορούν να προστεθούν και νέες.

Για να πάρετε μια γενική εικόνα ξεκινήστε διαβάζοντας το άρθρο στην διεύθυνση:

<http://google.csd.uoc.gr/bzr/corr?cmd=dump;path=GoogleCorr.pdf>

Περισσότερες πληροφορίες θα είναι διαθέσιμες στο wiki:

<http://google.csd.uoc.gr/apache2-default/>

Οργάνωση

Η εργασία αυτή θα εκπονηθεί από ομάδες των 2 ή 3 ατόμων (ανάλογα με την ανατεθείσα εργασία).

Φάσεις και Χρονοδιάγραμμα

Δήλωση Ομάδων (έως 18 Μαρτίου)

Ενημερώστε το αντίστοιχο τμήμα της σελίδας:

http://google.csd.uoc.gr/apache2-default/index.php/Συντονισμός_2008

Επίσης μπορείτε να εκφράσετε την προτίμηση σας σχετικά με τη λειτουργική ομάδα που θα θέλατε να αναλάβετε ή να περιγράψετε τι θα επιθυμούσατε πολύ να κάνετε. Η τελική ανάθεση εργασιών σε ομάδες θα καθοριστεί από τους βοηθούς του μαθήματος με βάση τις προτιμήσεις σας, την εμπειρία σας και την προτεραιότητα που έχουν οι εργασίες.

Φ1. Εξοικείωση με Google'07 και σχεδίαση εργασίας (18 Μαρτίου έως 1 Απριλίου)

1/ Χρησιμοποιείστε τη μηχανή και εντοπίστε προβλήματα και σημεία που επιδέχονται βελτιώσεων (καταγράψτε τα στην η-δνση <http://google.csd.uoc.gr/apache2-default/index.php/Bugs>). Σκεφτείτε/επεξεργαστείτε ιδέες και προτάσεις για βελτίωση ή επέκταση της λειτουργικότητάς της.

2/ Εξοικειωθείτε με τον κώδικα του mitos/Google 2007 και τις πρακτικές ανάπτυξης κώδικα που θα ακολουθηθούν. Δοκιμάστε να το εγκαταστήσετε σε δικό σας μηχάνημα και αναφέρετε τα προβλήματα/λύσεις στη σχετική σελίδα του wiki. Ενημερώστε τους βοηθούς για το αν τα καταφέρατε ή όχι.

3/ Συγκεντρώστε τις απορίες, τα προβλήματα και τις ιδέες σας και συζητήστε τα με τους βοηθούς στο σχετικό φροντιστήριο που θα γίνει.

4/ Μελετήστε και σκεφτείτε την εργασία που σας έχει ανατεθεί.. Συντάξτε μια αναλυτικότερη περιγραφή της εκφώνησης που να περιγράφει τι θα κάνετε, αν θα απαιτηθούν αλλαγές στην διεπαφή της μονάδας σας ή σε άλλες, άλλου είδους εξαρτήσεις, ανοικτά ζητήματα, κλπ.

Επίσης, μπορείτε να κάνετε προτάσεις για βελτίωση της υπάρχουσας αρχιτεκτονικής και πάνω σε άλλα θέματα σχεδίασης. Μπορείτε να χρησιμοποιήσετε και διαγράμματα UML. Class Diagrams και Sequence Diagrams θα βοηθήσουν πολύ.

Επίσης κάθε ομάδα καλό είναι να διαμερίσει την εργασία της σε φάσεις και για κάθε φάση να περιγράψει τη λειτουργικότητα που σκοπεύει να παραδώσει, την ημερομηνία παράδοσης καθώς τον τρόπο ελέγχου της ορθότητας της λειτουργικότητας (ιδανικά σχεδιασμός JUnit test αφού οριστικοποιηθούν οι διεπαφές των κλάσεων).

Φ2. 1^η φάση ανάπτυξης (25 Μαρτίου έως 14 Απριλίου)

Μετά το τέλος της σχεδίασης και κατόπιν συναντήσεως και συζήτησης με τους βοηθούς (σε ώρες του φροντιστηρίου) θα αρχίσει η φάση της ανάπτυξης.

Παραδοτέο: Πρώτη έκδοση του λογισμικού.

Φ3. 2^η φάση ανάπτυξης (20 Απριλίου έως 22 Μάιου)

Παραδοτέο: Δεύτερη έκδοση του λογισμικού (με τεκμηρίωση javadoc και JUnit tests).

Φ4. 1^{ος} Γύρος Δοκιμών – Διορθώσεων – Μετρήσεων - Επεκτάσεων (22 Μαΐου έως 12 Ιουνίου)

Παραδοτέο: Τελική έκδοση λογισμικού και γραπτή αναφορά μετρήσεων

Πρακτικές

- Θα πρέπει να ακολουθηθούν οι οδηγίες που περιγράφονται στο <http://google.csd.uoc.gr/apache2-default/index.php/Instructions>
- Ο κώδικας να συνοδεύεται από τεκμηρίωση (javadocs) και junit tests.
- Η χρήση του Wiki είναι επιβεβλημένη (για λόγους καλής επικοινωνίας, συντονισμού και διαμοίρασης χρήσιμων πληροφοριών)
- Συναντήσεις θα γίνονται κάθε Παρασκευή (στην ώρα του φροντιστηρίου) για επίλυση προβλημάτων, συζήτηση και λήψη αποφάσεων.
 - Τα μέλη των ομάδων μπορούν να ορίσουν κάποιες κοινές ώρες εργασίας ώστε να επιλύονται τα προβλήματα που προκύπτουν.
- Κάθε ομάδα μετά την τελική φάση θα πρέπει να ενημερώσει τις αντίστοιχες σελίδες του wiki. Εκεί μπορούν επίσης να καταγραφούν οι μετρήσεις και τα αποτελέσματα της αξιολόγησης. Τα παραπάνω θα ληφθούν υπόψη κατά τη βαθμολόγηση των ομάδων.

Αξιολόγηση Συστήματος

Το τελικό αποτέλεσμα θα αξιολογηθεί (εμπειρικά) σε πραγματικά δεδομένα. Η «ορθότητα» της λειτουργίας του θα δοκιμαστεί σε μικρές συλλογές σελίδων ενώ για την αξιολόγηση των επιδόσεων θα χρησιμοποιηθούν μεγαλύτερες συλλογές, για παράδειγμα όλες οι σελίδες του uoc.gr και του www.forth.gr (ιδανικά όλες οι σελίδες του .gr). Όσον αφορά τον σχεδιασμό του συστήματος κάντε την υπόθεση ότι το σύστημα θα πρέπει να δουλεύει ικανοποιητικά για τουλάχιστον 1.000.000 σελίδες.

Βαθμολόγηση

Θα υπάρξει ξεχωριστή βαθμολογία για κάθε ομάδα και για κάθε φάση. Το τελικό αποτέλεσμα (ήτοι το mitos2008 ver 1.0) θα λειτουργήσει ως καταλύτης (θετικός ή αρνητικός) στη βαθμολογία όλων των ομάδων. Αυτό για να ενθαρρύνουμε τη συνεργασία και το συντονισμό μεταξύ των ομάδων προσδοκώντας ένα καλό τελικό αποτέλεσμα.

Περισσότερες πληροφορίες θα αναρτώνται στο wiki και θα δίδονται από τους βοηθούς του μαθήματος .

Σχετική βιβλιογραφία και αναφορές μπορούν να καταγράφονται στη σελίδα http://google.csd.uoc.gr/apache2-default/index.php?title=Living_Bibliography

ΚΑΤΑΛΟΓΟΣ ΕΡΓΑΣΙΩΝ

Κωδικός	Εργασία	Περιγραφή
E1	Ερπυστής	Σχεδίαση και υλοποίηση ενός ερπυστή (κατά προτίμηση σε Java). Αναλυτική εκφώνηση και έτοιμα τμήματα λογισμικού θα δοθούν.
E2	Αξιολόγηση και Βελτίωση του Ελληνικού Στελεχωτή	Κατάρτιση συλλογής αξιολόγησης, υλοποίηση μέτρων αξιολόγησης και συγκριτική αξιολόγηση του στελεχωτή της ελληνικής γλώσσας του mitos. Αναλυτικές οδηγίες θα δοθούν σε ξεχωριστό κείμενο.
E3	Ανεστραμμένο Ευρετήριο	Σχεδίαση και υλοποίηση μιας μονάδας η οποία θα διαβάσει το DBMS-based ευρετήριο του mitos και θα κατασκευάζει ένα inverted file. Ενημέρωση του ευρετηριαστή και του αποτιμητή επερωτήσεων ώστε να χρησιμοποιούν το νέο ευρετήριο. Αναλυτικές οδηγίες θα δοθούν σε ξεχωριστό κείμενο
E4	Προσαρμογή γνωστού ερπυστή.	Προσαρμογή κάποιου έτοιμου ερπυστή (π.χ. Labrador, Java Crawler) για την χρήση του στο mitos.

Καλή εργασία & διασκέδαση