

ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ
Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

ΕΞΕΡΕΥΝΗΣΗ, ΟΠΤΙΚΟΠΟΙΗΣΗ ΚΑΙ ΑΝΑΛΥΣΗ ΜΕΓΑΛΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΛΗΡΟΦΟΡΙΕΣ: Νίκος Μπικάκης, 210 7721402, bikakis[a]dmlab.ece.ntua.gr

Σταύρος Μαρούλης, 210 6875403, stavmars[a]imis.athena-innovation.gr

Γιώργος Παπαστεφανάτος, 210 6875403, gparas[a]imis.athena-innovation.gr

ΠΕΡΙΛΗΨΗ: Σκοπός της εργασίας είναι η ανάπτυξη καινοτόμων εφαρμογών και τεχνικών για την οπτικοποίηση και ανάλυση (π.χ., ανάλυση συσχετίσεων, εντοπισμός τάσεων, στατιστική ανάλυση, κτλ.) μεγάλων δεδομένων (*Big Data*). Οι εφαρμογές και οι τεχνικές θα αναπτυχθούν στο πλαίσιο εξερεύνηση (*exploration*), οπτικοποίηση (*visualization*) και ανάλυση διασυνδεδεμένων δεδομένων (*Linked Data*) καθώς και οικονομικών δεδομένων (*Finance and Market Data*).

ΑΤΟΜΑ: 1-2

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Θα καθοριστεί σε συνεννόηση με τον ενδιαφερόμενο

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Ο τεράστιος όγκος δεδομένων (*Big Data*) που δημοσιεύονται σήμερα στο Web καθιστά τη χρήση και την αξιοποίησή τους πρακτικά αδύνατη από τον άνθρωπο και ειδικότερα από την πλειοψηφία των χρηστών στο Web που δε διαθέτουν εξειδικευμένες γνώσεις προγραμματισμού. Η οπτική αναπαράσταση των δεδομένων καθώς και η παροχή τεχνικών εξερεύνησης και ανάλυσης, αποτελεί βασική προϋπόθεση για την αξιοποίησή τους.

Στόχος της διπλωματικής είναι η ανάπτυξη καινοτόμων εφαρμογών και τεχνικών για την οπτικοποίηση και ανάλυση (π.χ., ανάλυση συσχετίσεων, εντοπισμός τάσεων, στατιστική ανάλυση, κτλ.) μεγάλων δεδομένων τα οποία είναι διαθέσιμα στο Web σήμερα. Οι εφαρμογές και οι τεχνικές θα αναπτυχθούν στο πλαίσιο οπτικοποίηση και ανάλυση διασυνδεδεμένων δεδομένων (*Linked Data*) καθώς και οικονομικών δεδομένων (*Finance and Market Data*). Ιδιαίτερη έμφαση θα δοθεί στην ανάπτυξη εφαρμογών και τεχνικών οι οποίες θα μπορούν να διαχειρίζονται αποδοτικά μεγάλους όγκους δεδομένων.

Πιο συγκεκριμένα, σε συνεννόηση με τους ενδιαφερόμενους θα μελετηθούν θέματα σχετιζόμενα με: (α) την σταδιακή δημιουργία οπτικοποιήσεων, βασιζόμενες στην πλοήγηση του χρήστη (*incremental visualization*), (β) εξερεύνηση η οποία προσαρμόζεται στις προτιμήσεις και στις ανάγκες των χρηστών (*personalized exploration*), (γ) εξερεύνηση και πλοήγηση μέσω όψεων (*facets browsing*) (δ) δομές δεδομένων και ευρετηρίων (*data structures & indexes*) για την αποδοτική εξερεύνηση και ανάλυση δεδομένων, (ε) δημιουργία περιλήψεων (*data summarization/abstraction*), κτλ.

Η συγκεκριμένη εργασία θα πραγματοποιηθεί σε συνεργασία με το Ινστιτούτο Πληροφοριακών Συστημάτων του Ερευνητικού Κέντρου "Αθηνά".

ΣΧΕΤΙΚΟ ΥΛΙΚΟ

- *Big Data*: <http://www.ibm.com/big-data>
- *Linked Data*: <http://linkeddata.org/>
- *Business Analytics*: https://en.wikipedia.org/wiki/Business_analytics
- *Market Analysis*: https://en.wikipedia.org/wiki/Market_analysis
- *Faceted Search*: http://en.wikipedia.org/wiki/Faceted_search