

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

**Κοκολάκη Άννα
Μεταπτυχιακή Φοιτήτρια**

**Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης
Επόπτης Μεταπτ. Εργασίας: Αναπλ. Καθηγητής, Ι. Τζιτζικας**

Δευτέρα, 17/09/2018, 10:00

Αίθουσα “Α. Παγιατάκης”, Ινστιτούτο Πληροφορικής, ΙΤΕ

**“Facetize: Ένα Διαλογικό Εργαλείο για τον Καθαρισμό και το Μετασχηματισμό
Συνόλων Δεδομένων για τη Διευκόλυνση της Εξερεύνησής τους”**

ΠΕΡΙΛΗΨΗ

Υπάρχει πληθώρα συνόλων δεδομένων σε διάφορες μορφές τα οποία συνήθως αποθηκεύονται σε αρχεία, φιλοξενούνται σε καταλόγους δεδομένων ή προσπελούνται μέσω των τελικών σημείων SPARQL. Στις περισσότερες περιπτώσεις, αυτά τα σύνολα δεδομένων δεν μπορούν να εξερευνηθούν απ' ευθείας από τους τελικούς χρήστες. Για να συμπληρώσουμε αυτό το κενό, σε αυτή τη διατριβή παρουσιάζουμε το σχεδιασμό και την υλοποίηση του Facetize, ενός διαλογικού εργαλείου επεξεργασίας που επιτρέπει στους απλούς χρήστες, χωρίς τεχνικό υπόβαθρο, να μετασχηματίζουν σύνολα δεδομένων είτε στατικά (δηλ. αποθηκευμένα σε αρχεία) είτε δυναμικά (δηλ. απαντήσεις ερωτήσεων), σε σύνολα δεδομένων που μπορούν να εξερευνηθούν άμεσα από τους ίδιους ή άλλους χρήστες. Το τελευταίο επιτυγχάνεται μέσω του γνωστού παραδείγματος εξερευνητικής αναζήτησης, που ονομάζεται Πολυεδρική Αναζήτηση (και της Πολυεδρικής Αναζήτησης με Προτιμήσεις). Στη διατριβή περιγράφονται οι απαιτήσεις, εισάγεται το απαιτούμενο σύνολο μετασχηματισμών και στη συνέχεια αναλύεται η λειτουργικότητα και η υλοποίηση του εργαλείου Facetize που πραγματώνει αυτούς τους μετασχηματισμούς. Οι υποστηριζόμενες λειτουργίες καλύπτουν ένα ευρύ φάσμα εργασιών (επιλογή, ορατότητα,

διαγραφές, τροποποιήσεις, ορισμός ιεραρχιών και διαστημάτων, παράγωγα γνωρίσματα και άλλα) και το Facetize επιτρέπει στο χρήστη να τις εκτελέσει με φιλικό τρόπο, χωρίς να προϋποθέεται οποιοδήποτε τεχνικό υπόβαθρο (όσον αφορά τις γλώσσες αναπαράστασης δεδομένων ή τις γλώσσες επερωτήσεων). Τέλος παρουσιάζονται τα αποτελέσματα μιας αξιολόγησης με τους χρήστες που ήταν θετική.

Kokolaki Anna
M.Sc. Thesis

Computer Science Department
University of Crete

Master's Thesis Supervisor: Associate Professor Ioannis Tzitzikas

Monday, 17/09/2018, 10:00
Room "A. Pagiatakis", FORTH

"Facetize: An Interactive Tool for Cleaning and Transforming Datasets for Facilitating their Exploration"

Abstract

There is a plethora of datasets in various formats which are usually stored in files, hosted in data catalogs or accessed through SPARQL endpoints. In most cases, these datasets cannot be straightforwardly explored by end users. To fill this gap, in this thesis we present the design and implementation of Facetize, an editor that allows plain users with no particular technical background, to transform datasets, either static (i.e. stored in files), or dynamic (i.e. being the results of SPARQL queries), to datasets that can be directly explored effectively by themselves or other users. The latter is achieved through the familiar interaction paradigm of Faceted Search (and Preference-enriched Faceted Search). We describe the requirements, we introduce the required set of transformations, and then we detail the functionality and the implementation of the editor Facetize that realizes these transformations. The supported operations cover a wide range of tasks (selection, visibility, deletions, edits, definition of hierarchies, intervals, derived attributes, and others). In this way Facetize enables the user to carry them out in a user-friendly and guided manner, without presupposing any technical background (regarding data representation or query languages). Finally we present the results of task-based evaluation with users which was quite positive.