



**ΕΡΓΑΛΕΙΑ ΓΙΑ ΤΗ ΣΥΓΧΩΝΕΥΣΗ ΔΕΔΟΜΕΝΩΝ
ΜΕ ΣΗΜΑΣΙΟΛΟΓΙΚΑ ΚΑΙ ΓΕΩΧΩΡΙΚΑ ΚΡΙΤΗΡΙΑ**

ΠΛΗΡΟΦΟΡΙΕΣ: Δημήτρης Σκούτας, [dskoutas \[at\] imis.athena-innovation.gr](mailto:dskoutas@imis.athena-innovation.gr)
Γιώργος Γιαννόπουλος, [giann \[at\] dblab.ece.ntua.gr](mailto:giann [at] dblab.ece.ntua.gr)

ΠΕΡΙΛΗΨΗ: Στόχος της διπλωματικής είναι η σχεδίαση και υλοποίηση ενός εργαλείου που θα διευκολύνει την αντιστοίχιση και συγχώνευση γεωχωρικών οντοτήτων με τεχνικές που συνδυάζουν σημασιολογική και γεωχωρική πληροφορία.

ΑΤΟΜΑ: 1

ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ: Java, PostGIS, OpenStreetMap

ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ: Το ανοιχτό μοντέλο λειτουργίας του Παγκόσμιου Ιστού συνεπάγεται ότι συνήθως μπορεί κανείς να αντλήσει πληροφορία για μία οντότητα από μια πληθώρα διαφορετικών, ετερογενών και ανεξάρτητων μεταξύ τους πηγών. Οι πηγές αυτές συχνά χρησιμοποιούν διαφορετικά αναγνωριστικά για να αναφερθούν στην ίδια οντότητα, διαφορετικούς τρόπους αναπαράστασης, και παρέχουν διαφορετικές πληροφορίες που ενδέχεται να είναι συμπληρωματικές, επικαλυπτόμενες ή και αντιφατικές. Για το λόγο αυτό, ένα μεγάλο μέρος της έρευνας έχει επικεντρωθεί στην ανάπτυξη αλγορίθμων και τεχνικών για την αντιστοίχιση και συγχώνευση τέτοιων δεδομένων [1].

Η συνήθης πρακτική είναι η κατασκευή εργαλείων τα οποία με τη βοήθεια ευριστικών αλγορίθμων προσπαθούν να εντοπίσουν πιθανές αντιστοιχίσεις μεταξύ οντοτήτων, και κατόπιν μέσω μια γραφικής διεπαφής επιτρέπουν στον χρήστη να δει τις προτεινόμενες αντιστοιχίσεις και να τις επιβεβαιώσει, απορρίψει ή συμπληρώσει. Συνεπώς είναι σημαντικό το γραφικό περιβάλλον να είναι εύχρηστο, κατανοητό και σχεδιασμένο με τρόπο που να υποστηρίζει και να διευκολύνει τη διαδικασία, ώστε να μειώνεται ο απαιτούμενος χρόνος και προσπάθεια, να αποφεύγονται λάθη, και να βελτιώνεται η ποιότητα και πληρότητα των αποτελεσμάτων.

Ωστόσο τα εργαλεία αυτά συνήθως δεν λαμβάνουν κάποια ειδική μέριμνα για τη διαχείριση γεωχωρικών δεδομένων, όπως π.χ. τη δυνατότητα απεικόνισής τους σε χάρτη ή την αξιοποίηση της χωρικής διάστασης στην ανεύρεση αντιστοιχίσεων. Το κενό αυτό είναι σημαντικό, καθώς τα γεωχωρικά δεδομένα αποτελούν μια πολύ μεγάλη και σημαντική κατηγορία δεδομένων, με πληθώρα πρακτικών εφαρμογών, και προέρχονται συχνά από ετερογενείς πηγές οπότε είναι υπαρκτή η ανάγκη για αντιστοίχιση και συγχώνευσή τους.

Σκοπός λοιπόν της διπλωματικής είναι να μελετηθούν διάφορα υπάρχοντα εργαλεία για διασύνδεση δεδομένων (π.χ. [2][3]), προκειμένου να σχεδιαστεί και να υλοποιηθεί ένα εργαλείο (είτε ως επέκταση υπάρχοντος είτε ανεξάρτητο) που θα είναι προσανατολισμένο και προσαρμοσμένο σε γεωχωρικά δεδομένα. Αυτό θα αφορά τόσο τις τεχνικές που θα ενσωματώνει όσο και τη δομή, σχεδίαση και λειτουργικότητα που θα προσφέρει το γραφικό του περιβάλλον.

Η διπλωματική θα πραγματοποιηθεί στα πλαίσια του ερευνητικού έργου GeoKnow το οποίο είναι ένα τριετές, χρηματοδοτούμενο από την ΕΕ ερευνητικό έργο, που αφορά στο γεωχωρικό Σηματολογικό Ιστό. Συνοπτικά, το GeoKnow καταπιάνεται με τη διασύνδεση, διαχείριση, ποιότητα, συνάθροιση, οπτικοποίηση και δημιουργία γεωχωρικών διαδικτυακών δεδομένων. Τα ερευνητικά μας αποτελέσματα θα εφαρμοστούν στις περιοχές των εφοδιαστικών αλυσίδων και των ταξιδιωτικών εταιρειών.

ΣΧΕΤΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ:

- [1] Bernstein et al. Generic Schema Matching, Ten Years Later, VLDB'11
http://www.sigmod.org/publications/sigmod-record/0906/publications/1003/p41_survey.drosou.pdf
- [2] Interlinking tools: <http://stack.lod2.eu/>
- [3] Fusion tools: <http://sieve.wbsg.de/>
- [4] Datasets, <http://linkedgeodata.org>