**ΠΡΟΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ (ΜΔΕ)**

Θεσσαλονίκη …. /… / 201...

**ΕΙΣΗΓΗΤΗΣ:** Σιδηρόπουλος Αντώνης

**ΤΙΤΛΟΣ ΕΡΓΑΣΙΑΣ**: Αναγνώριση Ονομάτων Ερευνητών-Συγγραφέων.

ΣΤΟΧΟΣ - ΑΝΤΙΚΕΙΜΕΝΟ (περίπου 50 λέξεις)

|  |
| --- |
| Στόχος της εργασίας είναι η εφαρμογή μεθόδων αναγνώρισης συγγραφέων και συνωνυμιών σε μια βάση βιβλιογραφικών δεδομένων. Χρησιμοποιείται η βάση “Open Academic Graph”. Σε αυτήν την βάση, η οποία παρέχεται από την microsoft, για την πληροφορία του συγγραφέα, δεν καταχωρείται κάποιο ID, αλλά το ονοματεπώνυμο και το Ίδρυμα συνεργασίας. Δεδομένου, ότι υπάρχουν συνωνυμίες, ή μερικές φορές κακή καταγραφή του ονόματος (πχ “Antonios Sidiropoulos” αντί του “Antonis Sidiropoulos” ή “Sidiropoulos Antonis” αντί “Antonis Sidiropoulos” κα.) θα πρέπει να γίνει αναγνώριση αυτών των περιπτώσεων αξιοποιώντας άλλες πληροφορίες όπως: λέξεις κλειδιά, ερευνητικός τομέας, συν-συγγραφείς και οποιαδήποτε άλλη πληροφορία είναι διαθέσιμη. Επιπλέον μπορούν (πρέπει) να χρησιμοποιηθούν εξωτερικές υπηρεσίες για την συμπλήρωση των δεδομένων. Για παράδειγμα μπορεί να χρησιμοποιηθεί το API του orcID. Το orcID είναι ένα μοναδικό αναγνωριστικό ερευνητών, το οποίο όμως δεν περιλαμβάνεται μέσα στην βιβλιογραφική βάση που έχουμε. Επιπλέον η υπηρεσία doi.org η οποία παρέχει πληροφορίες για έγγραφα, μπορεί να χρησιμοποιηθεί για την συμπλήρωση των δεδομένων. Για παράδειγμα η εντολή:  curl -LH "Accept: application/json; style=json" <https://doi.org/10.1016/j.neucom.2017.08.076> θα κατεβάσει σε json μορφή μετα-πληροφορίες για το έγγραφο με doi: 10.1016/j.neucom.2017.08.076. Σε αυτές τις πληροφορίες, για τον έναν εκ των τριών συγγραφέων, υπάρχει η πληροφορία orcID:  "author" : [  {  "family" : "Ougiaroglou",  "sequence" : "first",  "ORCID" : "http://orcid.org/0000-0003-1094-2520",  "given" : "Stefanos"  },  {  "given" : "Konstantinos I.",  "family" : "Diamantaras",  "sequence" : "additional",  },  {  "family" : "Evangelidis",  "sequence" : "additional",  "given" : "Georgios"  }  ].... |

ΜΕΘΟΔΟΛΟΓΙΑ - ΕΡΕΥΝΗΤΙΚΑ ΕΡΓΑΛEΙΑ (περίπου 50 λέξεις)

|  |
| --- |
| Η εργασία χωρίζεται σε 4 φάσεις:   1. Διερεύνηση και καταγραφή της σχετικής βιβλιογραφίας. 2. Κατανόηση και οργάνωση των δεδομένων. 3. Συμπλήρωση των δεδομένων όπως περιγράφηκε 4. Συγχρόνως μελέτη, υλοποίηση και βελτίωση υπαρχόντων αλγορίθμων αναγνώρισης συγγραφέων. |

ΑΝΑΜΕΝΟΜΕΝΑ ΑΠΟΤΕΛΕΣΜΑΤΑ (50-100 λέξεις)

|  |
| --- |
| Στόχος είναι η υλοποίηση ενός συστήματος το οποίο λειτουργεί σε πραγματικά δεδομένα, το Academic Graph το οποίο παρέχεται δωρεάν από την microsoft. |

ΠΡΟΑΠΑΙΤΟΥΜΕΝΕΣ ΓΝΩΣΕΙΣ

|  |
| --- |
| Προγραμματισμός, **Δομές δεδομένων**, **Βάσεις δεδομένων**. |

ΠΡΟΤΕΙΝΟΜΕΝΗ ΕΝΔΕΙΚΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

ΑΛΛΕΣ ΠΗΓΕΣ

……………………………………………………………

……………………………………………………………

……………………………………………………………

ΤΥΠΙΚΟΙ ΠΕΡΙΟΡΙΣΜΟΙ

* Γνώση αγγλικών

ΧΡΟΝΟΣ ΚΑΙ ΤΡΟΠΟΙ ΕΠΙΚΟΙΝΩΝΙΑΣ ΜΕ ΕΙΣΗΓΗΤΗ

Τόπος: Γρ. 106

Χρόνος (τουλάχιστον 1 ώρα/εβδ.)

Τηλ.: 2310 013(984)

E-mail: asidirop@gmail.com

Υπογραφή Εισηγητή