

Εφαρμογές Τεχνολογιών Γλωσσικής Επεξεργασίας στα Συστήματα Αναζήτησης των Ελληνικών Ακαδημαϊκών Βιβλιοθηκών

Άννα Μάστορα¹, Μανόλης Πεπονάκης² και Σαράντος Καπιδάκης¹

¹Ίονιο Πανεπιστήμιο, Εργαστήριο Ψηφιακών Βιβλιοθηκών και Ηλεκτρονικής Δημοσίευσης, Τμήμα Αρχιτεκτονικής και Βιβλιοθηκονομίας

²Εθνικό Κέντρο Τεκμηρίωσης, Εθνικό Ίδρυμα Ερευνών

[mastora\(a\).ionio.sr](mailto:mastora(a).ionio.sr), [eueuo\(a\):ekt.sr](mailto:eueuo(a):ekt.sr), saraMosiaionio.gr

Περίληψη

Η αναντιστοιχία μεταξύ των όρων της ερώτησης που υποβάλλει ο χρήστης και των όρων που έχουν ευρετηριαστεί είναι ένα σημαντικό πρόβλημα, το οποίο επηρεάζει την ανάκτηση σχετικών τεκμηρίων κατά την αναζήτηση πληροφοριών. Σε γλώσσες με έντονη μορφολογία, όπως είναι η Ελληνική γλώσσα, η λέξη παίρνει διαφορετικές μορφές για να εκφράσει αριθμούς, πτώσεις, γένη, χρόνους κτλ. δυσχεραίνεται, συνεπώς, το έργο της αναζήτησης πληροφοριών καθώς πρέπει να είναι εκ των προτέρων γνωστή η μορφή της λέξης που έχει ευρετηριαστεί ώστε να υποβληθεί ομοίως στο σύστημα κατά την αναζήτηση. Ένας τρόπος αντιμετώπισης του προβλήματος της έντονης μορφολογίας των γλωσσών, στο πλαίσιο της αναντιστοιχίας μεταξύ των υποβαλλόμενων όρων σε μια ερώτηση και εκείνων που περιλαμβάνονται στο αντίστοιχο ευρετήριο, σχετίζεται με τις Εφαρμογές Τεχνολογιών Γλωσσικής Επεξεργασίας (ΕΤΓΕ).

Στόχος της παρούσας μελέτης είναι να καταδείξει τα πλεονεκτήματα των Εφαρμογών Τεχνολογιών Γλωσσικής Επεξεργασίας στα συστήματα αναζήτησης των Ελληνικών Ακαδημαϊκών Βιβλιοθηκών και ταυτόχρονα να καταδείξει τα προβλήματα που προκύπτουν όταν υλοποιούνται με διαφορετικό τρόπο οι παραπάνω τεχνολογίες. Για το λόγο αυτό, εξετάσαμε ως προς την εφαρμογή αυτών των τεχνολογιών Ελληνικές Ακαδημαϊκές Βιβλιοθήκες που διαθέτουν και OPAC και Ίδρυματικό Αποθετήριο καθώς και συγκρίναμε τη συνέπεια εφαρμογής αυτών τόσο μεταξύ των συστημάτων του ίδιου ιδρύματος όσο και οριζόντια, δηλαδή των ιδρυμάτων μεταξύ τους για να διαπιστώσουμε το βαθμό διαλειτουργικότητας που επιτυγχάνεται.

Λέξεις κλειδιά: Εφαρμογές Τεχνολογιών Γλωσσικής Επεξεργασίας, Πληροφοριακά Συστήματα, Δημόσιοι Κατάλογοι Βιβλιοθηκών, Ακοθετίψια, Αναζήτησι/ Πληροφοριών

Implementation of Language Processing Techniques in Greek Academic Libraries' Information Systems

Abstract

A way of dealing with the problem of highly inflectional languages as well as the query - document terms mismatch problem is by implementing Language Processing Techniques. The aim of this study is to report, through presenting empirical data, on the Language Processing Techniques and their advantages if implemented by the information retrieval systems of Greek Academic Libraries.

The objectives of this study are twofold. First goal is to acknowledge these techniques and then try to designate the interoperability issues deriving from the varying implementations. For this purpose, we examined Greek Academic Libraries which host both an OPAC and an Institutional Repository towards whether they implement any kind of language technology. Additionally, we report whether these techniques are implemented consistently in terms of the OPAC and the Institutional Repository of the same institution, as well as among different institutions.

Keywords: Language Processing Techniques, Information Systems, Libraries' OPACs, Repositories, Information Search