

24 Κατασκευή βάσης δεδομένων ελληνικών ακρωνυμίων σε ελληνικά νομικά κείμενα

Τσιμπούρης Χαράλαμπος, Κυριάκος Σγάρμπας

ΠΕΡΙΛΗΨΗ

Στην παρούσα ανακοίνωση παρουσιάζεται η κατασκευή βάσης δεδομένων ελληνικών ακρωνυμίων και η περαιτέρω μελέτη τους, η οποία εκτελέστηκε υπό την αιγίδα της Επιτροπής Ερευνών του Πανεπιστημίου Πατρών, στα πλαίσια της βασικής έρευνας «Καραθεοδωρής». Για την ολοκλήρωση της εργασίας, χρησιμοποιήθηκε το σώμα κειμένων eThemis (<http://www.e-themis.gov.gr>) η οποία βασίστηκε στο έντυπο υλικό, τη θεματολογία και τη μεθοδολογία άτυπης κωδικοποίησης του έργου "Διαρκής Κώδικας Νομοθεσίας - Ραππάρχης", όπως συγκεντρώθηκε από το Τμήμα Εκδόσεων και Γραμματειακής Λογιστικής Εξυπηρέτησης της Διεύθυνσης Οικονομικού της Γενικής Γραμματείας Δημόσιας Διοίκησης και Ηλεκτρονικής Διακυβέρνησης του Υπουργείου Εσωτερικών.

Με την ολοκλήρωση της έρευνας, κατασκευάστηκε λίστα 503 διαφορετικών ακρωνυμίων, εκ των οποίων 339 δηλώθηκε η ανάλυσή τους μέσα στο κείμενο. Στην συνέχεια επισημειώθηκαν αντίστοιχα τα κατάλληλα αρχεία της βάσης eThemis με ειδικές λέξεις κλειδιά (xml tags). Σκοπός της μελέτης είναι η δημιουργία βάσης δεδομένων ελληνικών νομικών κειμένων για περαιτέρω μελέτη των ελληνικών ακρωνυμίων, αλλά και η απαραίτητη υποδομή για εκπαίδευση έμπειρων συστημάτων αναγνώρισης ακρωνυμίων, διαχωριστή προτάσεων (sentence tokenizer) ή λέξεων (word tokenizer).

Κατά τη πορεία της μελέτης, έγινε καταγραφή ποικίλων προβλημάτων που ανέκυψαν αναλύοντας τη συγκεκριμένη βάση δεδομένων κώδικα νομοθεσίας και σχετίζονται α) με τον τρόπο γραφής, β) αναφοράς μέσα στο κείμενο αλλά και γ) πλήρης ανάλυσής τους στο κείμενο όπου χρειάζεται. Τέλος, γίνεται σαφές πως παρόμοια προβλήματα εμφανίζονται ευρέως σε ελληνικά νομικά κείμενα, και πηγάζουν από το γεγονός πως δεν υπάρχει κεντρικός έλεγχος συντήρησης ακρωνυμίων σε συνδυασμό με την έλλειψη προτυποποίησης στο τρόπο κατασκευής τους.

Database construction of Greek acronyms, used in Greek law texts

Tsimpouris Charalampos, Kyriakos Sgarbas

ABSTRACT

This paper presents the construction of a dictionary of acronyms for legal texts in Modern Greek language, with a semi-automatic procedure, which was developed with the support of the Research Program Caratheodory, 2010-2013, of the Research Committee of the University of Patras, Greece.. A corpus of legal texts was gathered from online resources and a small portion of it was analysed manually by a legal and a language expert. The results of this paper are part of a greater project aiming the study of Greek legal texts for training acronym recognition systems and high precision sentence and word tokenizers.

During the semi-automatic procedure explanations of the target acronyms were located inside the texts (503 unique acronyms), most of which (339) were explained in-text, and the corpus was annotated accordingly with predefined xml tags for each acronym-explanation pair. Our goal was to create a database with Greek legal texts for further analysis of Greek acronyms, while also support future experiments of expert systems, sentence or word tokenizers.

0 Εισαγωγή

Οι συγγραφείς εγγράφων τείνουν να χρησιμοποιούν συχνά ακρωνύμια. Χρησιμοποιώντας ακρωνύμια, αποφεύγουν την επανάληψη μακρόσυρτων και δύσχρηστων τίτλων και διευκολύνουν τον αναγνώστη στην κατανόηση του νοήματος. Σαφώς, για να αποφεύγεται η σύγχυση, τα ακρωνύμια που χρησιμοποιούνται στα κείμενα θα πρέπει να επεξηγούνται ρητώς. Σε επίσημα κείμενα, (π.χ. ακαδημαϊκού χαρακτήρα) δύο τεχνικές ενστερνίζονται οι συγγραφείς, οι οποίες είναι συνάρτηση τόσο του μεγέθους όσο και του είδους του εκάστοτε εγγράφου. Πιο συγκεκριμένα, για τις μεγάλες σε μήκος και αυτόνομες εργασίες, όπως ένα εγχειρίδιο ή μια διατριβή, μια λίστα ορισμών επεξηγεί με σαφήνεια τα ακρωνύμια και αρκτικόλεξα που βρίσκονται διασκορπισμένα σε όλο το κείμενο. Η εν λόγω λίστα, βρίσκεται συνήθως στην αρχή του εγγράφου μετά τον πίνακα περιεχομένων. Από την άλλη μεριά, στα πιο μικρά κείμενα όπως τεχνικά εγχειρίδια (RFC), ιατρικά ή νομικά τα αρκτικόλεξα επεξηγούνται στην ίδια πρόταση, μία φορά κατά την πρώτη χρήση τους ενώ ο ορισμός βρίσκεται μέσα σε παρενθέσεις.

Δυστυχώς, δεν ακολουθούν όλοι οι συγγραφείς τις παραπάνω δύο τεχνικές. Είναι πολύ συχνό το φαινόμενο να χρησιμοποιούνται ακρώνυμα τα οποία δεν επεξηγούνται γιατί θεωρούνται γνωστά για το κοινό που απευθύνονται, με αποτέλεσμα η ερμηνεία να γίνεται διαφορετική ή αδύνατη. Σε τέτοιες περιπτώσεις μέθοδοι αποσαφήνισης μπορούν να εφαρμοστούν για να βοηθήσουν τους αναγνώστες [1, 2]. Πιο συγκεκριμένα, οι μηχανές αναζήτησης, είναι ένα από τα πολλά εργαλεία, που μπορούν να βοηθήσουν στη διαδικασία προσδιορισμού σχετικά με τους ορισμούς αρκτικόλεξων, όπως για παράδειγμα <http://www.asas.gr/>. Αλλά για να λειτουργήσουν κατάλληλα τα συγκεκριμένα εργαλεία χρειάζονται αντίστοιχη βάση δεδομένων η οποία να περιλαμβάνει τον σαφή ορισμό κάθε αρκτικόλεξου. Η ύπαρξη μιας τέτοιας βάσης δεδομένων δεν είναι πάντα διαθέσιμη ούτε για

κάθε αντικείμενο μελέτης ούτε για κάθε γλώσσα. Στην παρούσα εργασία εξετάζουμε τα νομικά κείμενα στα σύγχρονα ελληνικά. Τέλος, να σημειώσουμε πως χρησιμοποιείται ο όρος αρκτικόλεξο και ακρωνύμιο ισοδύναμα.

2 Σχετικές εργασίες

Τα σύγχρονα ηλεκτρονικά λεξικά καταγράφουν μια πολύ μικρή μερίδα των λημμάτων τους ως αρκτικόλεξα ή συντμήσεις: περίπου 1% (857 από τα 81000 λήμματα) στο λεξικό Κοραΐς [3, 4], ακόμα λιγότερα στο Neurolingo MRD [5]. Από την μέχρι στιγμής έρευνα μας, δεν έχουν δημοσιευτεί, επίσημα, στατιστικά στοιχεία για το ποσοστό των ακρωνυμίων που χρησιμοποιούνται στην ελληνική γραμματεία. Παρόλα η δική μας έρευνα, στο σώμα κειμένου που συγκεντρώσαμε, δείχνει ότι αυτό το ποσοστό ανέρχεται στο 0,6% ενώ όσον αφορά τα ελληνικά νομικά κείμενα αυτό το ποσοστό ανεβαίνει σε 1,5%. Σύμφωνα με τον Ελληνικό Οργανισμό Τυποποίησης (<http://www.elot.gr>, τελευταία προσπέλαση Ιούνιος 2013), δεν υφίσταται κάποια τυποποίηση στον τρόπο δημιουργίας των αρκτικόλεξων. Κατά συνέπεια, η διαδικασία της δημιουργίας αρκτικόλεξων δεν είναι σαφώς προσδιορισμένη και ως εκ τούτου η κατανόηση του ορισμού τους υπόκειται σε μεγάλο βαθμό ασάφειας. Τέλος, δεν κατέστη δυνατό να βρεθεί κατάλληλη μελέτη στη βιβλιογραφία που να εξετάζει το συγκεκριμένο πρόβλημα σε νομικά κείμενα.

Από την έρευνα μας διαπιστώσαμε ότι οι αποδοτικότερες μεθοδολογίες τόσο για την ανίχνευση των αρκτικόλεξων όσο και τον προσδιορισμό της έννοιας / της εξήγησής τους βασίζονται σε κανονικές εκφράσεις (regular expressions) [6, 7], και αυτόματα πεπερασμένων καταστάσεων [8], όπως επίσης κάποιες εφαρμογές των ταξινομητών Naive Bayes [9].

3 Κατασκευή βάσης δεδομένων

Τα στοιχεία που χρησιμοποιήσαμε για την αυτόματη επεξεργασία του κειμένου αντλήθηκαν από την ελληνική βάση δεδομένων e-Themis (<http://www.e-themis.gov.gr>, τελευταία προσπέλαση Ιούνιος 2012). Η εν λόγω ιστοσελίδα υποστηρίζεται από το τμήμα λογιστικής και γραμματειακής υποστήριξης του Υπουργείου Εσωτερικών και παρέχει δωρεάν πρόσβαση σε 40 θεματικές ενότητες κώδικα νομοθεσίας (π.χ. Σύνταγματικό, Διοικητικό, Νομαρχίες, Πυροσβεστική Υπηρεσία, Αστυνομία κλπ). Η ιστοσελίδα e-Themis ήταν ενεργή μέχρι και τα τέλη του 2012. Από τότε, η πρόσβαση με τον κεντρικό υπολογιστή και κατ' επέκταση στις νομικές πληροφορίες δεν είναι πλέον διαθέσιμη. Διατηρήσαμε τη δομή της βάσης δεδομένων της ιστοσελίδας (40 κατηγορίες) και είναι αποθηκευμένη σε 105 έγγραφα κωδικοποίησης UTF-8, μεγέθους 410 Mbytes και 32.025.632 λέξεων (tokens).

4 Μεθοδολογία

Το σύνολο των νομικών κειμένων έχει υποβληθεί σε επεξεργασία χρησιμοποιώντας μια ημιαυτόματη διαδικασία. Ο κύριος στόχος μας, κατά την επεξεργασία, ήταν να ελαχιστοποιήσουμε το χειρονακτική επεξεργασία, και γι' αυτό το λόγο πήραμε δειγματοληπτικά κάποιες λέξεις από το σύνολο των κειμένων και δημιουργήσαμε την κανονική έκφραση $R = ((([\text{verb}=\text{A}-\Omega \] \setminus \{1,4\} \text{verb}=\text{.}) ?) +) ([\text{verb}=\text{A}-\Omega \] [\text{verb}=\text{.}] [\text{verb}=\text{ } =])$. Με τη συγκεκριμένη πρακτική, καταφέραμε να μετρήσουμε κατά προσέγγιση τα ακρωνύμια που παρουσιάζονται και στα 105 αρχεία της βάσης, και καταλήξαμε σε σύνολο 17841 καταχωρήσεις.

Ο σκοπός αυτής της απλοποιημένης ανάλυσης ήταν να προσδιοριστούν τα αρχεία τα οποία εμφανίζουν τη μεγαλύτερη συχνότητα ακρωνυμίων. Τα πρώτα τρία αρχεία (στα οποία η συχνότητα των αρκτικόλεξων ήταν μεγαλύτερη) επιλέχθηκαν για περαιτέρω ανάλυση όπως εξηγείται στη συνέχεια.

Αυτά τα τρία αρχεία (τα αρχεία Α, Β και Γ που είναι 1^ο, 2^ο και 3^ο στη λίστα αντίστοιχα) εξετάστηκαν και προσδιορίστηκαν όλα τα ακρωνύμια τα οποία εμφανίζονται τουλάχιστον μια φορά, ενώ παράλληλα σημειώθηκαν και οι αντίστοιχες επεξηγήσεις, όπως δίνονταν μέσα στο κείμενο. Για το αρχείο Α, από τα 199 διαφορετικά αρκτικόλεξα που τακτοποιήθηκαν μόνο τα 161 είχαν επεξήγηση. Για το αρχείο Β, από τα 128 αρκτικόλεξα, τα 48 είχαν ορισμό, ενώ τέλος για το τρίτο αρχείο από τα 206 αρκτικόλεξα που βρέθηκαν μόνο τα 130 επεξηγούνταν. Σε αυτό το σημείο, πρέπει να τονίσουμε ότι δύο ακρωνύμια θεωρούνται διαφορετικά, όταν διαφέρουν οι ορισμοί τους παρόλα που ενδεχομένως απαρτίζονται από τα ίδια γράμματα.

Αρχικά, ολοκληρώθηκε η διαδικασία συλλογής όλων των ακρωνυμίων και των αντίστοιχων ορισμών τους από τα 3 αρχεία. Στην συνέχεια, έπρεπε να προσδιορίσουμε που βρέθηκαν αυτά τα ακρωνύμια μέσα στα 3 συγκεκριμένα κείμενα. Για τον λόγο αυτό, κατασκευάσαμε κώδικα (Python 2.6) το οποίο ημιαυτόματα αναγνώριζε όλα τα σημεία που εμφανιζόντουσαν τα ακρωνύμια, ενώ σε δύσκολες περιπτώσεις ο χρήστης έπαιρνε την αντίστοιχη απόφαση για να γίνει σαφές αν έχει δηλωθεί η επεξήγηση. Επίσης, επιλέξαμε παράθυρο καταγραφής 200 χαρακτήρων πριν και μετά τα σημεία μελέτης, αρκετό σε μέγεθος για οποιαδήποτε περαιτέρω ανάλυση.

5 Λεξικό

Το τελικό λεξικό και ο κατάλογος περιπτώσεων περιλαμβάνει όλα τα ελληνικά ακρωνύμια τα

οποία ανιχνεύτηκαν από τα κείμενα και αποθηκεύονται ως κείμενο χωρίς μορφοποίηση σε κωδικοποίηση UTF-8. Η καταγραφή των στοιχείων (features) του παραπάνω αρχείου οριστικοποιήθηκε με τον παρακάτω τρόπο:

1. Ακρωνύμιο (σε κανονική μορφή)
2. Ακρωνύμιο (όπως ανιχνεύτηκε)
3. Επεξήγηση (αν υπάρχει)
4. Θέση μέσα στο αρχείο
5. Παράθυρο προθέματος 200 χαρακτήρων
6. Παράθυρο επιθέματος 200 χαρακτήρων
7. «1» αν η επεξήγηση βρίσκεται εντός του παραθύρου χαρακτήρων πριν το ακρωνύμιο, και «0» σε διαφορετική περίπτωση.
8. «1» αν η επεξήγηση βρίσκεται εντός του παραθύρου χαρακτήρων μετά το ακρωνύμιο, και «0» σε διαφορετική περίπτωση.
9. «1» αν το ακρωνύμιο βρίσκεται μέσα σε παρενθέσεις, «0» σε διαφορετική περίπτωση

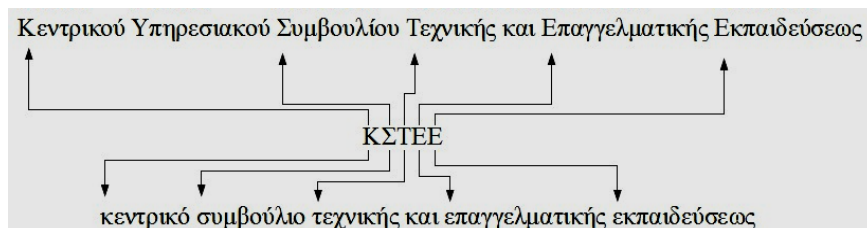
6 Προβλήματα αναγνώρισης

Χρησιμοποιώντας μια απλή κανονική έκφραση όπως " $[A-\Omega] \cdot []$ +" or " $[A-\Omega]$ +", είμαστε σε θέση να προσδιορίσουμε τα περισσότερα από τα ακρωνύμια που αναφέρονται στην βάση δεδομένων για τα νομικά κείμενα. Παρόλα αυτά, η παραπάνω έκφραση δεν αποτελεί βέλτιστη επιλογή γιατί δεν ακολουθούν όλα τα κείμενα το ίδιο μοτίβο. Για παράδειγμα, σε πολλές περιπτώσεις το ακρωνύμιο έχει ένα κενό μετά από κάθε τελεία ή να μην έχει τελεία, ή να μην έχει κενό. Η διαφοροποίηση αυτή δημιουργεί προβλήματα στη διαδικασία τεμαχισμού προτάσεων και λέξεων (sentence and word tokenizers). Για παράδειγμα, το Φύλλο Εφημερίδας της Κυβερνήσεως ως αρκτικόλεξο ανιχνεύεται ως "Φ.Ε.Κ." (χωρίς κενό), ή ως "Φ_Ε_Κ." (με κενό) ή και ως "ΦΕΚ" (χωρίς τελείες, χωρίς κενό). Επεκτείνοντας την παραπάνω έκφραση έτσι ώστε να περιλαμβάνει όλες τις παραπάνω περιπτώσεις δεν λύνει το πρόβλημα δεδομένου ότι εισάγει εσφαλμένα αποδεκτές λύσεις, όπως για παράδειγμα ολόκληρες λέξεις με κεφαλαία γράμματα (τίτλοι, υπότιτλοι, κλπ.). Στην συνέχεια ακολουθεί λίστα με προβλήματα τα οποία αναγνωρίστηκαν κατά τη διαδικασία της ανάλυσης.

1. Διαφορετική επεξήγηση σε ίδιο Ακρωνύμιο
Ακρωνύμιο: ΚΣΤΕΕ (Εικόνα 1)

Επεξήγηση, γραφή 1: *Κεντρικού Υπηρεσιακού Συμβουλίου Τεχνικής και Επαγγελματικής Εκπαίδευσης*

Επεξήγηση, γραφή 2: *κεντρικό συμβούλιο τεχνικής και επαγγελματικής εκπαίδευσης*



Εικόνα 1: Ανάλυση ακρωνυμίου ΚΣΤΕΕ

2. Δυσκολία συσχέτισης μεταξύ γραμμάτων ακρωνυμίου και λέξεων επεξήγησης

Ακρωνύμιο: ΚΥΣΕΠ

Επεξήγηση: *Κεντρικών Υπηρεσιακών και Πειθαρχικών Συμβούλιον Εποπτικού Προσωπικού Μέσης Τεχνικής και Επαγγελματικής Εκπαίδευσης*

3. Επαναλαμβανόμενη χρήση λέξης

Ακρωνύμιο: Ε.Ε.Κ.

Επεξήγηση, γραφή 1: *επαγγελματική εκπαίδευση και κατάρτιση*

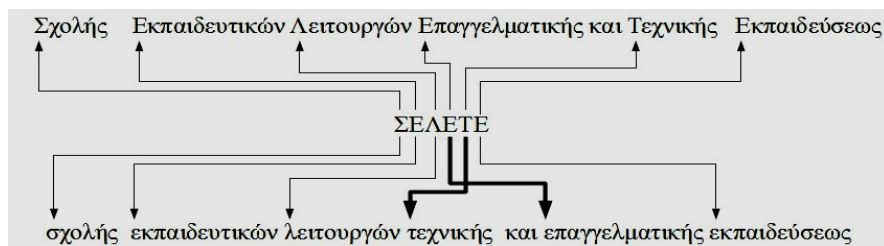
Επεξήγηση, γραφή 2: *επαγγελματική εκπαίδευση και επαγγελματική κατάρτιση*

4. Αλλαγή σειράς λέξεων

Ακρωνύμιο: ΣΕΛΕΤΕ

Επεξήγηση, γραφή 1: *Σχολής Εκπαιδευτικών Λειτουργών Επαγγελματικής και Τεχνικής Εκπαίδευσης*

Επεξήγηση, γραφή 2: *σχολής εκπαιδευτικών λειτουργών τεχνικής και επαγγελματικής εκπαίδευσης*



Εικόνα 2: Ανάλυση ακρωνυμίου ΣΕΛΕΤΕ

5. Διαφορετική κλίση ή τρόπος γραφής

Ακρωνύμιο: Α.Ε.

Επεξήγηση, γραφή 1: *Ανώνυμος Εταιρεία*

Επεξήγηση, γραφή 2: *Ανωνύμων Εταιρειών*

Επεξήγηση, γραφή 3: *Ανωνύμων Εταιριών*

6. Η επεξήγηση βασίζεται σε άλλες συντομεύσεις

Ακρωνύμιο: ΥΠΕΠΘ

Επεξήγηση, γραφή 1: *Υπουργείο Εθνικής Παιδείας και Θρησκευμάτων*

Επεξήγηση, γραφή 2: *υπ. εθν. παιδείας και θρησκευμάτων*

7. Ασάφεια στον επίσημο τίτλο δημόσιας υπηρεσίας

Ακρωνύμιο: Ε.Σ.Α.Π.

Επεξήγηση: Εθνικό Συμβούλιο Ανώτατης Παιδείας

Ακρωνύμιο 2: Σ.Α.Π.

Επεξήγηση: Συμβούλιο Ανώτατης Παιδείας

Ακρωνύμιο 3: Ε.Σ.Π.

Επεξήγηση: Εθνικό Συμβούλιο Παιδείας

7 Συμπεράσματα

Η ανάλυση κατέστησε σαφές ότι υπάρχει σημαντική έλλειψη τυποποίησης και μεγάλη διαφοροποίηση στον τρόπο με τον οποίο τα αρκτικόλεξα κατασκευάζονται ανά περίπτωση, ενώ θα περίμενε κανείς ότι αυτό δεν θα συμβεί για τα νομικά κείμενα (δεδομένου ότι συνήθως ελέγχονται πολλαπλώς). Κατά συνέπεια, το πρόβλημα αναγνώρισης επεξηγήσεων έχει αποδειχθεί αρκετά δύσκολο. Η εργασία μας στόχευσε να δημιουργήσει ένα χειροποίητο σώμα ελληνικών ακρωνυμίων από νομικά κείμενα χρησιμοποιώντας με τη λιγότερη δυνατή ανθρώπινη προσπάθεια.

8 Μελλοντική κατεύθυνση

Μια διαδικτυακή έκδοση ενός συστήματος προσδιορισμού αρκτικόλεξων είναι υπό ανάπτυξη, η οποία θα μπορεί να δέχεται τα έγγραφα χρηστών και να αναγνωρίζει τα αρκτικόλεξα που περιλαμβάνονται, καθώς επίσης και να προσπαθεί να εξάγει τυχόν επεξηγήσεις. Αυτό το σύστημα θα δημιουργήσει τη δυνατότητα για την περαιτέρω αξιολόγηση της συζητημένης μεθοδολογίας, και τη δυνατότητα να μετρηθεί η ακρίβεια αλλά και να αξιολογηθεί σε πραγματικές απαιτήσεις χρηστών. Επίσης, αποτελεί σκοπό της μελέτης μας και η βελτίωση του αλγορίθμου σε άλλου είδους νομικά κείμενα όπως αποφάσεις δικαστηρίων, υλικό το οποίο είναι αρκετά πιο δύσκολο να συγκεντρωθεί.

9 Υποστήριξη

Η συγκεκριμένη εργασία εκτελέστηκε υπό την αιγίδα της Επιτροπής Ερευνών του Πανεπιστημίου Πατρών, στα πλαίσια της βασικής έρευνας «Καραθεοδωρής».

10 Βιβλιογραφία

- [1] Sumita, E. and Sugaya, F. (2006). *Using the web to disambiguate Ακρωνύμιοι*, *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, Association for Computational Linguistics, USA.
- [2] Joshi, M., Pakhomov, S., Pedersen, T., and Chute, G. C. (2006). *A Comparative Study of Supervised Learning as Applied to Ακρωνύμιο Expansion in Clinical Reports*, AMIA Annu Symp Proc. 2006; 2006: 399–403.
- [3] Kokkinakis, D. (2006). *Collection, Encoding and Linguistic Processing of a Swedish Medical Corpus: The MEDLEX Experience*, *Proceedings of the 5th LREC*. Genoa, Italy.
- [4] Lyras, D., Kokkinakis, G., Lazaridis, A., Sgarbas, K., and Fakotakis, N. (2009). *A Large Greek-English Dictionary with Incorporated Speech and Language Processing Tools*, *Interspeech 2009, 10th Annual Conference of the International Speech Communication Association*, pp.1891-1894, 6-10 September 2009, Brighton, UK.
- [5] Gakis, P., Panagiotakopoulos, Ch., Sgarbas, K., and Tsalidis, Ch. (2012). *Design and Implementation of an Electronic Lexicon for Modern Greek*, *Literary and Linguistic Computing*, Vol.27, No.2, pp.155-169.
- [6] Dannells, D. (2006). *Automatic Ακρωνύμιο Recognition*, *Annual Meeting of The European Chapter of The Association of Computational Linguistics*, Sydney.
- [7] Xiaonan, J., Gu, X., James, B. and Hang, L. (2008). *Mining, Ranking, and Using Ακρωνύμιο Patterns*, *Proceedings of the 10th Asia-Pacific web conference on Progress in WWW research and development*, Shenyang, China, pp. 371-382.
- [8] Kempe, A. (2006). *Ακρωνύμιο-meaning extraction from corpora using multi-tape weighted finite-state machines*, *CoRR*, abs/cs/0612033.
- [9] Yoon, Y., C., Park, S., Y., Song, Y., I. Rim, H., C., and Rhee, D., W. (2008) *Automatic Ακρωνύμιο Dictionary Construction Based on Ακρωνύμιο Generation Types*, *IEICE Transactions on Information and Systems E91-D*, pp. 1584-1587.

Χαράλαμπος Τσιμπούρης

Υπ. Διδάκτορας, Τμήμα Ηλεκτρολόγων Μηχανικών, Πανεπιστήμιο Πατρών
Τμ. Ηλ.Μηχ & Τεχν. Υπολ, Εργαστήριο Ενσύρματης Τηλεπικοινωνίας
Ηλ-ταχ.: xtsimpouris@upatras.gr

Δρ. Μηχ. Κυριάκος Σγάρμπας

Επ. Καθηγητής, Τμήμα Ηλεκτρολόγων Μηχανικών, Πανεπιστήμιο Πατρών
Τμ. Ηλ.Μηχ & Τεχν. Υπολ, Εργαστήριο Ενσύρματης Τηλεπικοινωνίας
Ηλ-ταχ.: sgarbas@upatras.gr