

ΕΞΟΥΡΥΞΗ ΓΝΩΣΗΣ ΣΕ ΒΙΟΪΑΤΡΙΚΑ ΚΕΙΜΕΝΑ

Αριστείδης Βαγγελάτος¹, Άννα Ιορδανίδου², Γιάννης Σταματόπουλος¹,
Γιώργος Ορφανός³, Βασίλης Παπαπαναγιώτου⁴

1. EAITY
2. Πανεπιστήμιο Πατρών
3. Neurosoft A.E.
4. Γενική Κλινική «ΙΑΣΩ General»

Εισαγωγή

Στην παρουσίαση αυτή γίνεται αναφορά στο έργο «Ιατρολέξη» (ερευνητικό έργο χρηματοδοτούμενο από τη ΓΓΕΤ, στο πλαίσιο του Γ' ΚΠΣ - www.iatrolexi.gr).

Σκοπός

Σκοπός του έργου είναι η δημιουργία της κρίσιμης υποδομής για την ελληνική γλώσσα, η οποία θα αποτελέσει τη βάση για εξελιγμένες εφαρμογές επεξεργασίας φυσικής γλώσσας (ΕΦΓ) στο θεματικό πεδίο της βιοϊατρικής, όπως ευρετηρίαση κειμένων (text indexing), εξαγωγή και ανάκτηση πληροφορίας (information extraction and retrieval), εξόρυξη γνώσης, συστήματα ερωταποκρίσεων κτλ.

Υλικό και Μέθοδος

Για να επιτευχθεί ο παραπάνω στόχος, δημιουργούνται γλωσσικοί πόροι και εργαλεία κατάλληλα προσαρμοσμένα στα χαρακτηριστικά της ελληνικής γλώσσας και της βιοϊατρικής υπογλώσσας, που επιτρέπουν τη διαχείριση και την επεξεργασία της ψηφιοποιημένης πληροφορίας στο συγκεκριμένο γνωστικό πεδίο.

Αποτελέσματα

Τα αποτελέσματα του έργου είναι εργαλεία που απευθύνονται τόσο στον τελικό χρήστη, όπως είναι ο *ορθογραφικός διορθωτής ελληνικών βιοϊατρικών όρων*, αλλά και εργαλεία που θα ενισχύσουν την επεξεργασία βιοϊατρικών κειμένων στα ελληνικά και θα βελτιώσουν την αναζήτηση και την ανάκτηση βιοϊατρικών δεδομένων, όπως *μορφοσυντακτικός σχολιαστής* (morphosyntactic tagger) κατάλληλα προσαρμοσμένος στις ιδιαιτερότητες της βιοϊατρικής υπογλώσσας, και *οντολογία* της βιοϊατρικής ορολογίας, που θα χρησιμοποιηθεί ως τελικό προϊόν για αναζήτηση ορολογικών και εννοιολογικών πληροφοριών αλλά και ως ενδιάμεσος γλωσσικός πόρος, ως **βάση γνώσης** (knowledge base), για την τροφοδότηση ενός *σημασιολογικού σχολιαστή βιοϊατρικών κειμένων*.

Συμπεράσματα

Στις μέρες μας όπου η παραγωγή δεδομένων στο χώρο της υγείας έχει φτάσει σε δυσθεώρητα ύψη, είναι σημαντικό ο ερευνητής αλλά και ο απλός χρήστης να έχει τα κατάλληλα εργαλεία για να μπορεί να «ανακαλύψει» τη γνώση που χρειάζεται. Για να γίνει δυνατό αυτό, είναι απαραίτητη η υλοποίηση μιας βασικής υποδομής, μέρος της οποίας αναπτύσσεται στο πλαίσιο του έργου «Ιατρολέξη».

Λέξεις κλειδιά: Ιατρική Πληροφορική, Εξόρυξη κειμένων, οντολογίες.