

## Εργασία Bonus (0.5/10 μονάδες)

**Προθεσμία Ολοκλήρωσης: 4 Φεβρουαρίου 2024 στις 23:59**

Τα τελευταία χρόνια η περιοχή της Επεξεργασίας Φυσικής Γλώσσας (υποπεριοχή της Τεχνητής Νοημοσύνης) έχει σημειώσει αλματώδη πρόοδο. Μεταξύ των επιμέρους τομέων της Επεξεργασίας Φυσικής Γλώσσας, η Αναγνωστική Κατανόηση (Reading Comprehension) έχει σημειώσει αντίστοιχη σημαντική πρόοδο τα τελευταία χρόνια. Ωστόσο, η πλειονότητα των επιστημονικών εργασιών αυτού του τομέα αφορούν την Αγγλική γλώσσα, καθώς τα διαθέσιμα δεδομένα άλλων γλωσσών, όπως τα Ελληνικά, είναι περιορισμένα.

Το **GreekQA Dataset (Greek Question Answering Dataset)** στοχεύει να γίνει το πρώτο Ελληνικό σύνολο δεδομένων Αναγνωστικής Κατανόησης. Θα περιλαμβάνει ερωτήσεις και απαντήσεις σε ένα σύνολο άρθρων της ελληνικής Wikipedia και θα αναπτυχθεί με τη βοήθεια εθελοντών όπως εσείς.

Η συγκεκριμένη πρωτοβουλία γίνεται στα πλαίσια σχετικής έρευνας της Ομάδα Τεχνητής Νοημοσύνης του Τμήματος Πληροφορικής και Τηλεπικοινωνιών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών (<https://ai.di.uoa.gr/>). Η ανάπτυξη του GreekQA ακολουθεί την μεθοδολογία του Stanford Question Answering Dataset (SQuAD) για το οποίο μπορείτε να βρείτε περισσότερες πληροφορίες στην ιστοσελίδα <https://rajpurkar.github.io/SQuAD-explorer/>.

**Στην εργασία αυτή σας ζητάμε να διαβάσετε παραγράφους της Ελληνικής Wikipedia και να απαντήσετε σε τουλάχιστον 50 ερωτήσεις που αναφέρονται σε αυτές τις παραγράφους.** Για να το κάνετε αυτό, θα χρειαστεί να γραφτείτε στην πλατφόρμα <https://greekqa.di.uoa.gr/>. Εκεί θα βρείτε οδηγίες πως να ξεκινήσετε. Οι απάντησεις πρέπει να είναι το **μικρότερο δυνατό μέρος του κειμένου** (φράση, λέξη ή υποπρόταση) που απαντά στην εκάστοτε ερώτηση και όχι ολόκληρη πρόταση.

Μπορείτε να κάνετε αυτή την εργασία ακόμα κι αν την έχετε ήδη κάνει σε άλλο δικό μου μάθημα μια που οι ερωτήσεις που θα σας ανατεθούν για απάντηση από την πλατφόρμα θα είναι διαφορετικές.

Αφού ολοκληρώσετε τη συνεισφορά σας, παρακαλούμε να εισάγετε τα στοιχεία σας στη φόρμα <https://forms.gle/BjoeN4gWZwVrbZZF9> ώστε να επαληθεύσουμε ότι ολοκληρώσατε την εργασία και να σας βαθμολογήσουμε.

Η πλατφόρμα έχει αναπτυχθεί από τον πρώην φοιτητή του Τμήματος Στάθη Σιάτρα στα πλαίσια της πτυχιακής του.

Ευχαριστούμε για την συνεισφορά σας!