



**ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ**

**ΕΡΓΑΣΙΑ ΣΤΟ ΜΑΘΗΜΑ «ΓΛΩΣΣΙΚΗ ΤΕΧΝΟΛΟΓΙΑ»
ΑΚΑΔ. ΕΤΟΣ 2024-2025**

ΘΕΜΑ: ΑΞΙΟΛΟΓΗΣΗ ΜΗΧΑΝΙΚΗΣ ΜΕΤΑΦΡΑΣΗΣ

Στόχος της παρούσας εργασίας είναι η αξιολόγηση της μετάφρασης που παράγεται από σύγχρονα συστήματα μεγάλων μοντέλων γλώσσας (GPT, ChatGPT, GEMINI κλπ.). Τα ζεύγη γλωσσών που ενδιαφέρουν είναι 1. το Ελληνικά-Αγγλικά και 2. το Αγγλικά-Ελληνικά.

A. ΕΠΙΣΚΟΠΗΣΗ ΤΗΣ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Στο openourses παρατίθενται κάποιες εργασίες που έχουν πραγματοποιήσει αξιολόγηση μετάφρασης σε διάφορα ζεύγη γλωσσών.

ΒΗΜΑ 1 (ΜΕΧΡΙ 4/11): Αναζητήστε περισσότερη βιβλιογραφία σχετικά με την εργασία αυτή στο scholar.google.com

Χρησιμοποιείτε για την αναζήτησή σας λέξεις-κλειδιά όπως “machine translation”, “evaluation”, “GPT” κλπ.

B. ΕΠΙΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

ΒΗΜΑ 2 (ΜΕΧΡΙ 7/11): Θα σας δοθεί σεν δεδομένων με ζεύγη Ελληνικών-Αγγλικών προτάσεων. Διαλέξτε 500 ζεύγη προτάσεων, έτσι ώστε να καλύπτουν γλωσσολογικά φαινόμενα στα οποία σας ενδιαφέρει να εστιάσετε την αξιολόγησή σας. Διαλέξτε τουλάχιστον 5 φαινόμενα. Τέτοια φαινόμενα μπορεί να είναι

Φαινόμενο	Παράδειγμα
Στίξη	Ο Κώστας πήγε τον γιό του, Γιάννη, για ποδόσφαιρο. Ο Κώστας πήγε τον γιό του Γιάννη για ποδόσφαιρο.
Συνθετικότητα	Hairless, doll house, fire wood, skin deep
Δομική Μορφολογία	This not is untieable. Αυτός ο κόμπος δεν λύνεται. Αυτός ο κόμπος μπορεί να λυθεί.
Μορφολογία	Happy? ---> Χαρούμενος; Χαρούμενη; You can call me. Μπορείς να με καλέσεις. Μπορείτε να με καλέσετε.
Λειτουργικές λέξεις	Περιμένω την απόφαση. I am waiting for the decision.
Άρνηση	Δεν τον χωνεύω καθόλου. Φοβάμαι μη δεν έρθει.
Κύρια Ονόματα	Ένωση Ξενοδοχοϋπαλλήλων Κέρκυρας, Joe Biden
Χρόνος ρήματος	Διδάσκω εδώ και δυο χρόνια.--> I have been teaching for two years. Had he not decided to come, this wouldn't have happened.
Mediopassive voice	This book reads well.

Ορίσματα (Αντικείμενα) του ρήματος (verb valency)	I gave the dog a bone. Έδωσα του σκύλου ένα κόκκαλο.
False friends	costume (στολή), κουστούμι
Phrasal verbs	The student picked up a new language. John picked up the children.
Modal verbs	We need this --> We don't need this. We can go --> We cannot go.
Multi-word expressions	More often than not such problems are difficult to solve.
Αμφισημία	Πλησιάζω τον καλόγερο.
Σημασιολογικοί ρόλοι	The bike accident broke Sara's arm. Η Σάρα έσπασε το χέρι της στο ατύχημα με το ποδήλατο.
Functional shift	You can whatsapp me on this number. Μπορείς να επικοινωνήσεις μαζί μου στο whatsapp σε αυτό το νούμερο.
Αναφορά	The cat climbed up a tree. It was afraid.
Ερωτηματικές προτάσεις	You do believe in happy endings, don't you?
Inversion	Here comes the bus!
Topicalization	That pizza, I won't eat.
Substitution	I'll get the books in the hall and you get the ones in the classroom.
Προτάσεις μεγάλου μήκους	He was also seen wearing harem-style trousers as he tapped his feet along with his new track.
Εξαρτήσεις μεγάλης απόστασης	The professor that Alexis was confident the students ignored continued with his lecture.
Ορολογία	μετατροπέας πεπερασμένων καταστάσεων, μελάνωμα
Ιδιαίτερο λεξιλόγιο, με ιδιαίτερη πολιτισμική χροιά	onomatopoeia
Παράταξη (Coordination)	Θέλω ή το ένα ή το άλλο. I want either the one or the other.
Υπόταξη (Subordination)	Αναρωτιέμαι αν είναι αλήθεια. I wonder whether it is true. Αν βρέξει θα πάμε μέσα. If it rains we will go inside.
Έλλειψη	Η Ελένη δέχτηκε την πρόσκληση, και η Μαίρη το ίδιο.
Ιδιωματισμοί	He reacted at the drop of a hat.
Resultative predicates	The man wiped the table clean.

Γ. ΜΕΤΑΦΡΑΣΗ

ΒΗΜΑ 3 (MEXPI 28/11): Επιλέξτε δυο σύγχρονα εργαλεία μετάφρασης (GPT3.5, GPT4, ChatGPT, GEMINI, MELTEMI και μεταφράστε τις προτάσεις ελέγχου της γλώσσας-πηγή. Αν επιλέξετε το ζεύγος Ελληνικά-Αγγλικά, η γλώσσα-πηγή είναι τα Ελληνικά και η γλώσσα στόχος είναι τα Αγγλικά.. Αν επιλέξετε το ζεύγος Αγγλικά-Ελληνικά, η γλώσσα-πηγή είναι τα Αγγλικά και η γλώσσα-στόχος τα Ελληνικά.

Δ. ΑΞΙΟΛΟΓΗΣΗ

ΒΗΜΑ 4 (MEXPI 9/12): Χρησιμοποιήστε κάποιο εργαλείο υπολογισμού αυτόματων μέτρων αξιολόγησης (πχ BLEU), όπως το <https://translate.tilde.ai/bleu/> και υπολογίστε αυτόματα μέτρα αξιολόγησης για τις προτάσεις σας και για τα δυο συστήματα μετάφρασης που επιλέξατε.

Συγκρίνετε τα αποτελέσματα. Τις προτάσεις στην γλώσσα στόχο που έχετε από τα δεδομένα σας τις χρησιμοποιείτε σαν αναφορά (reference).

ΒΗΜΑ 5 (ΜΕΧΡΙ 16/1): Μελετήστε μια-μια τις μεταφράσεις ως προς τα φαινόμενα που έχετε επιλέξει. Συγκρίνετε τα αυτόματα μεταφράσματα με τις μεταφράσεις-αναφορά στα δεδομένα σας. Καταγράψτε αναλυτικά τα σφάλματα που γίνονται, ανά φαινόμενο και ανά σύστημα μετάφρασης. Κάντε ποιοτική αξιολόγηση των αποτελεσμάτων και σύγκριση ανάμεσα στα συστήματα.

Ε. ΣΥΝΤΑΞΗ ΕΡΓΑΣΙΑΣ

ΒΗΜΑ 6 (ΜΕΧΡΙ ΤΗΝ ΗΜΕΡΟΜΗΝΙΑ ΤΩΝ ΕΞΕΤΑΣΕΩΝ ΤΟΥ ΜΑΘΗΜΑΤΟΣ): Συντάξτε επιστημονική εργασία στα Αγγλικά (σαν τις εργασίες της βιβλιογραφίας) μέχρι 10 σελίδες το πολύ, στην οποία θα παρουσιάζετε

- το θεωρητικό υπόβαθρο της δουλειάς σας
- το σετ δεδομένων σας
- την πειραματική σας διαδικασία και τα αποτελέσματά σας
- την αξιολόγηση των αποτελεσμάτων σας
- την σύγκριση των αποτελεσμάτων σας με αυτά άλλων προσεγγίσεων στη βιβλιογραφία
- τα συμπεράσματα στα οποία καταλήξατε και μελλοντικές βελτιώσεις που προτείνετε για την προσέγγισή σας
- την διαφορετικότητα στην προσέγγισή σας (τι διαφορές/καινοτομίες/συνεισφορά έχει η δική σας δουλειά σε σχέση με τις παραπλήσιες της βιβλιογραφίας)
- την λίστα των βιβλιογραφικών σας παραπομπών

ΑΛΛΑ

Η εργασία θα πραγματοποιηθεί σε ομάδες των τεσσάρων τουλάχιστον ατόμων.

Η εργασία είναι υποχρεωτική και πιάνει το 50% στον τελικό βαθμό.

Η ομάδα που θα πετύχει το μεγαλύτερο σκορ στα αποτελέσματα θα έχει

- επιπλέον βαθμολογικό bonus
- ευκαιρία να προωθήσει την εργασία της για δημοσίευση σε επιστημονικό συνέδριο