

# Αξιολόγηση Μηχανικής Μετάφρασης: Ελληνικά προς Αγγλικά

Σφηναρολάκης Στέφανος inf2021218

Νικόλαος Τρυπάκης inf2021229

Κωνσταντίνος Καφτεράνης inf2021090

# Εισαγωγή

Η αυτόματη μετάφραση (MT) έχει σημειώσει σημαντική πρόοδο, διευκολύνοντας την επικοινωνία και γεφυρώνοντας γλωσσικά χάσματα. Ωστόσο, οι λιγότερο διαδεδομένες γλώσσες, όπως τα ελληνικά, παρουσιάζουν μοναδικές προκλήσεις

Η μελέτη μας εστιάζει σε δυσκολίες που αντιμετωπίζουν τα σύγχρονα μοντέλα MT όταν μεταφράζουν από ελληνικά σε αγγλικά.



# Προκλήσεις Ελληνικής Γλώσσας

Η ελληνική γλώσσα έχει πλούσια γραμματική και ευέλικτη σύνταξη, με:

- Ελεύθερη σειρά λέξεων.
- Πολλούς διαφορετικούς τύπους ρημάτων.
- Κλίσεις που δείχνουν το νόημα (όπως πτώσεις και χρόνοι).
- Ιδιαίτερες πολιτισμικές εκφράσεις που αντικατοπτρίζουν την ελληνική ιστορία και κουλτούρα.

Αυτά τα χαρακτηριστικά κάνουν τη μετάφραση σε γλώσσες όπως τα αγγλικά, που έχουν πιο αυστηρούς κανόνες σύνταξης, πιο δύσκολη.



# Επισκόπηση Βιβλιογραφίας 1



Η ανάπτυξη της Νευρωνικής Μετάφρασης (NMT) με βάση την αρχιτεκτονική Transformer έφερε σημαντική πρόοδο, αλλά παραμένουν προκλήσεις:

- **Πρόοδος:**
  - Τα “πάει” πολύ καλά σε μεταφράσεις ελληνικών κειμένων.
  - Υπερτερεί σε σύγκριση με παραδοσιακές μεθόδους για την ελληνική γλώσσα.
- **Περιορισμοί:**
  - Δυσκολία απόδοσης πολιτισμικών εκφράσεων και γλωσσικών διαφορών.
  - Αγγλοκεντρική εκπαίδευση των πολυγλωσσικών μοντέλων.
  - Έλλειψη δίγλωσσων δεδομένων, που οδηγεί σε μειωμένη ακρίβεια για γλώσσες με περιορισμένους πόρους.

# Επισκόπηση Βιβλιογραφίας 2

- **Ιστορικά Συστήματα:**
  - METIS και METIS-II: Στατιστικά μοντέλα χωρίς δίγλωσσους πόρους.
  - Uplug: Δημιουργία αξιόπιστου δίγλωσσου λεξικού.
  - KantanMT, Moses: Διαγλωσσική ανάκτηση πληροφοριών (υπερέχει το KantanMT).
- **Ανάγκες:**
  - Ανάπτυξη εξειδικευμένων dataset και προσαρμοσμένων μοντέλων.
  - Αντιμετώπιση γλωσσικών διαφορών και πολιτισμικών ιδιαιτεροτήτων.



# Μεθοδολογία

Στόχος μας ήταν να αξιολογήσουμε δύο σύγχρονα γλωσσικά μοντέλα:

- Claude: Ισχυρό στην διαχείριση σύνθετων φράσεων και προσαρμοσμένο σε προχωρημένες γλωσσικές κατασκευές.
- Microsoft Copilot: Εξαιρετικό σε βασικές προτάσεις και ενσωματωμένο σε πολλαπλές εφαρμογές.

Χρησιμοποιήσαμε dataset 1.500 προτάσεων με πέντε γλωσσικά φαινόμενα, εφαρμόζοντας την BLEU μέτρηση για αξιολόγηση.



# Γλωσσικά Φαινόμενα

Αναλύθηκαν πέντε φαινόμενα με βάση τις δυσκολίες που παρουσιάζουν στη μετάφραση:

1. **Άρνηση, 629 προτάσεις**
2. **Συντονισμός (Coordination), 401 προτάσεις**
3. **Υπόταξη (Subjunction), 138 προτάσεις**
4. **Στίξη, 332 προτάσεις**
5. **Μεταφορές και ειδικές φράσεις, 10 προτάσεις**



# Αποτελέσματα

**Claude:** Υπερέχει στη διαχείριση σύνθετων φράσεων και τεχνικών όρων.

**Copilot:** Καλύτερη απόδοση σε απλές γραμματικές δομές.

Αδυναμία και των δύο μοντέλων στη μετάφραση μεταφορών και ειδικών φράσεων.

BLEU scores δείχνουν γενική ικανοποιητική απόδοση, με προκλήσεις σε συγκεκριμένα γλωσσικά φαινόμενα.





Grammatical Phenomenon	Number of Sentences	Score (Copilot)	Score (Claude)
Negation	629	57.14	60.55
Coordination	401	21.21	47.05
Subjunctions	138	26.54	49.38
Punctuation	332	59.97	59.79
Metaphor/Special Phrase	10	27.69	40.21

Table 1: Linguistic phenomena analysis: number of sentences and BLUE scores by Microsoft Copilot and Claude.

BLEU:

59.79

52.97

Precision x brevity: 60.63 × 98.61

52.97 × 100.00

Type 1-gram 2-gram 3-gram 4-gram

1-gram 2-gram 3-gram 4-gram

Individual 82.47 66.67 55.18 44.54

72.26 57.29 48.04 39.58

Cumulative 81.33 73.12 66.26 59.79

72.26 64.34 58.37 52.97

Export data [CSV](#)

Sentence 494	BLEU	Length ratio	Text
Human	100.00	1.00	Don ' t look a gift horse in the mouth .
Machine	5.06	1.55	Someone was given a donkey as a gift , and they were looking at its teeth .
Machine	18.80	1.27	Someone was being given a donkey and was looking it in the mouth .

Sentence 200	BLEU	Length ratio	
Source	-	-	Do they know you're here?
Human	100.00	1.00	Do they know you ' re here ?
Machine	23.26	0.75	They know you are here .
Machine	42.38	0.88	Do they know you are here ?

Sentence 381	BLEU	Length ratio	
Source	-	-	Tom is naked from the waist up.
Human	100.00	1.00	Tom is naked from the waist up .
Machine	100.00	1.00	Tom is naked from the waist up .
Machine	13.01	0.50	Tom is shirtless .

# Περιορισμοί Μελέτης

Περιορισμένο dataset με μόλις 1.500 προτάσεις.

Μικρό δείγμα δεδομένων για τις μεταφορές και ειδικές φράσεις.

Περιορισμοί BLEU metric:

- Δυσκολία στην αξιολόγηση σωστών παραφράσεων.
- Θεωρεί λάθος μικρές διαφορές, όπως "can't" vs. "cannot".



# Συμπεράσματα και Προτάσεις

- Απαιτείται ανάπτυξη ειδικών μοντέλων προσαρμοσμένων στη μετάφραση ελληνικών προς αγγλικά.
- Χρειάζεται η χρήση μεγαλύτερων και πιο εξειδικευμένων dataset.
- Πρόταση για νέο πείραμα με μεγαλύτερο dataset ειδικών εκφράσεων και χρήση μοντέλου εξειδικευμένο στην μεταγραφή σε Ελληνικά-Αγγλικά



# Βιβλιογραφία

**Omkar Khade**, *Challenges in Adapting Multilingual LLMs to Low-Resource Languages using LoRA PEFT Tuning*

<https://arxiv.org/pdf/2411.18571>

**Ourania Kolovou**, *Machine Translation from Ancient Greek to English: Experiments with OpenNMT :*

[https://gupea.ub.gu.se/bitstream/handle/2077/81765/Machine%20Translation%20from%20Ancient%20Greek%20to%20English\\_Experiments%20with%20OpenNMT.pdf?sequence=1&isAllowed=y](https://gupea.ub.gu.se/bitstream/handle/2077/81765/Machine%20Translation%20from%20Ancient%20Greek%20to%20English_Experiments%20with%20OpenNMT.pdf?sequence=1&isAllowed=y)

**Gabriel Nicholas Aliya Bhatia**, *Lost in Translation: Large Language Models in Non-English Content Analysis*

<https://arxiv.org/pdf/2306.07377>

**Dimitris Papadopoulos**, *PENELOPIE: Enabling Open Information Extraction for the Greek Language through Machine Translation:*

<https://arxiv.org/abs/2103.15075>

**K. Skianis, J. Pavlopoulos, A. S. Dogruoz**, *Severity Prediction in Mental Health: LLMbased Creation, Analysis, Evaluation of a Novel Multilingual Dataset:* <https://arxiv.org/pdf/2409.17397>



