

Open Large Language Models For Greek: Meltemi & Krikri



Vassilis Katsouros

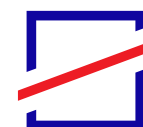
Director

Sokratis Sofianopoulos

Research Associate

Institute for Language and Speech Processing

Connect 2025 – Athens, June 5, 2025



ATHENA Research & Innovation
Information Technologies

What are Large Language Models (LLMs)

- / A language model is a type of artificial intelligence (AI) that can generate text
- / It is trained on a massive amount of text data
- / Can be used to generate new text that is like the text it was trained on
- / Training data can include books, articles, code, maths, and even social media conversations
- / By fine-tuning it you can have a model specialized for a specific task or domain



Our Motivation

- / **Underrepresentation:** Greek is not well served by major commercial models as are other languages
- / **Preservation of Language & Culture:** Creating AI that understands Modern, Ancient, and Polytonic Greek
- / **Local & Global Impact:** Supports Greek education, research, business, and culture
- / **Open-weight Benefits:** Transparency, adaptability, collaboration



Meltemi: Our first effort

On September 2024 we released **Meltemi**, the first open Greek LLM

/ 7B-parameters with 8K context length

/ Built on Mistral-7B

/ Demonstrated the feasibility of achieving fluent Greek text generation and comprehension by adapting a pre-existing model

/ Challenges

- Data curation for Greek
- Tokenizer adaptation
- Data mix (“catastrophic forgetting”)



Krikri: Taking the Next Step

- / Released in February 2025
- / Builds upon the foundation laid by Meltemi
- / It utilizes the more advanced Llama 3.1 architecture
- / Krikri scales up the model size (8B parameters) and context length (128k tokens) for enhanced capabilities
- / It incorporates more sophisticated alignment techniques to improve the model's helpfulness, harmlessness, and "honesty"



Key Improvements of Krikri over Meltemi

- / Enhanced natural language understanding and generation
- / Improved code-related task performance
- / Larger scale and greater capability
- / Deeper language comprehension and more sophisticated responses due to expanded context window





Building Meltemi & Krikri

Training strategy

Vocabulary Extension | mistral-meltemi

/ The original Mistral tokenizer did not contain meaningful Greek subwords

- It breaks up Greek into letters
- Extended it for Greek to (a) better capture context and semantics, (b) increase inference speed, and (c) lower training costs

Text:

Τα μεγάλα γλωσσικά μοντέλα χρειάζονται καλούς tokenizers

Tokenized with mistralai/Mistral-7B-v0.1:

['_', 'Τ', 'α', ' ', 'μ', 'ε', 'γ', 'ά', 'λ', 'α', ' ', 'γ', 'λ', 'ω', 'σ', 'σ', 'ι', 'κ', 'ά', ' ', 'μ', 'ο', 'ν', 'τ', 'έ', 'λ', 'α', ' ', 'χ', 'ρ', 'ε', 'ί', 'α', 'ζ', 'ο', 'ν', 'τ', 'α', 'ι', ' ', 'κ', 'α', 'λ', 'ο', 'ύ', 'ς', ' ', '_token', 'izers']

Tokenized with ilsp/Meltemi-7B-v1:

['_Τα', '_μεγάλα', '_γλωσσ', 'ικά', '_μοντέλα', '_χρειάζονται', '_καλούς', '_token', 'izers']



Vocabulary Extension | llama-krikri

/ Llama3: 21 tokens

<|begin_of_text|> Τα μεγάλα γλωσσικά μοντέλα χρειάζονται κα-
λοῦς token izers <|end_of_text|>

/ gpt 4o: 17 tokens

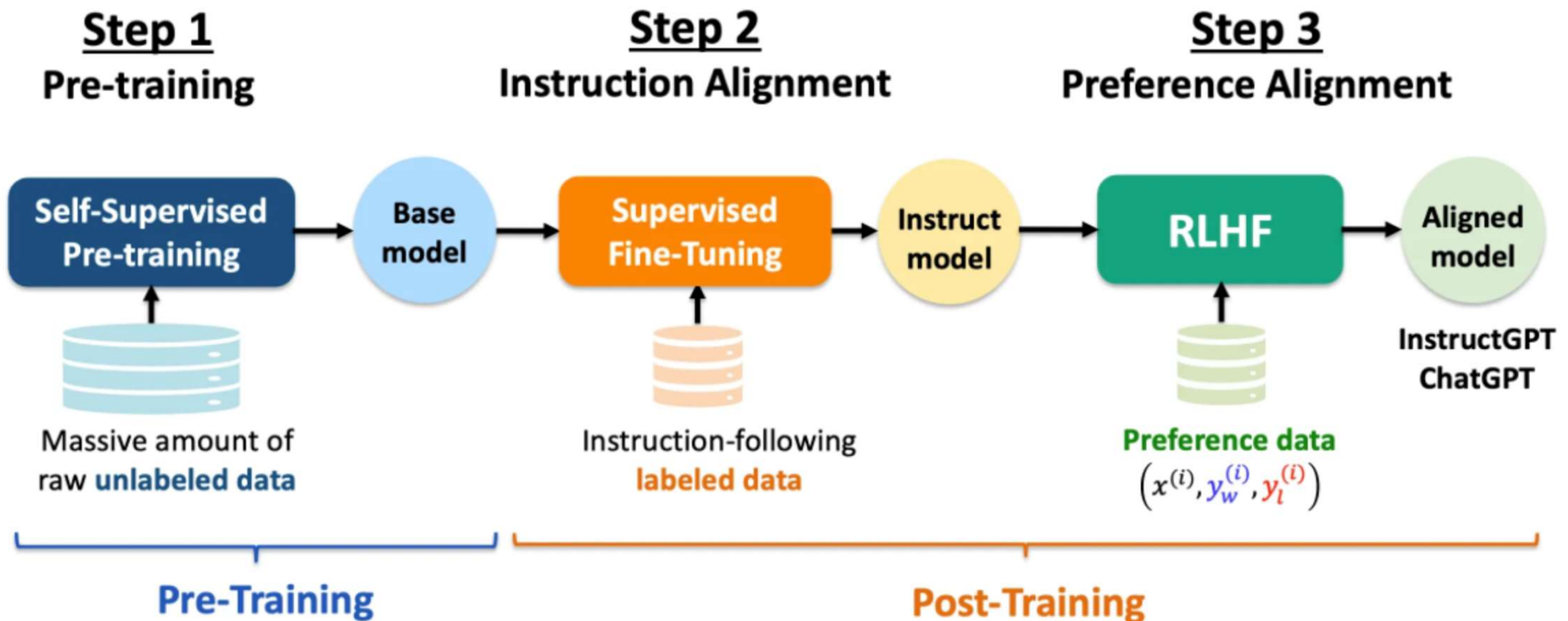
Τα μεγάλα γλωσσικά μοντέλα χρειάζονται καλούς tokenizers

/ Llama-Krikri-8b: 9 tokens

_Ta _μεγάλα _γλωσσικά _μοντέλα _χρειάζονται _καλούς
_token izers

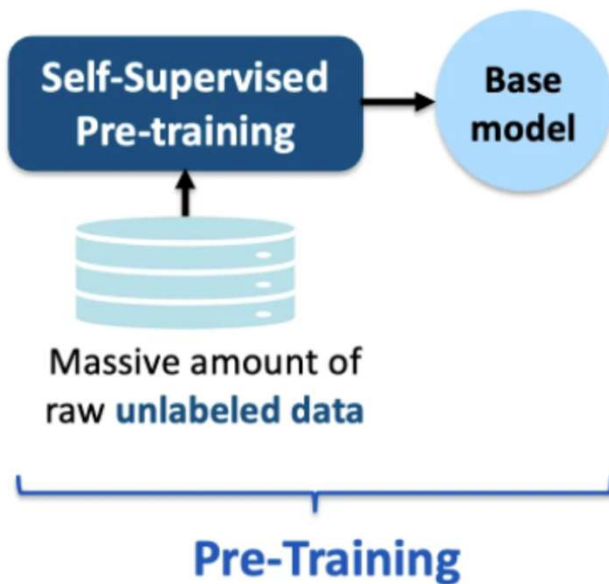


Steps for training Meltemi and Krikri

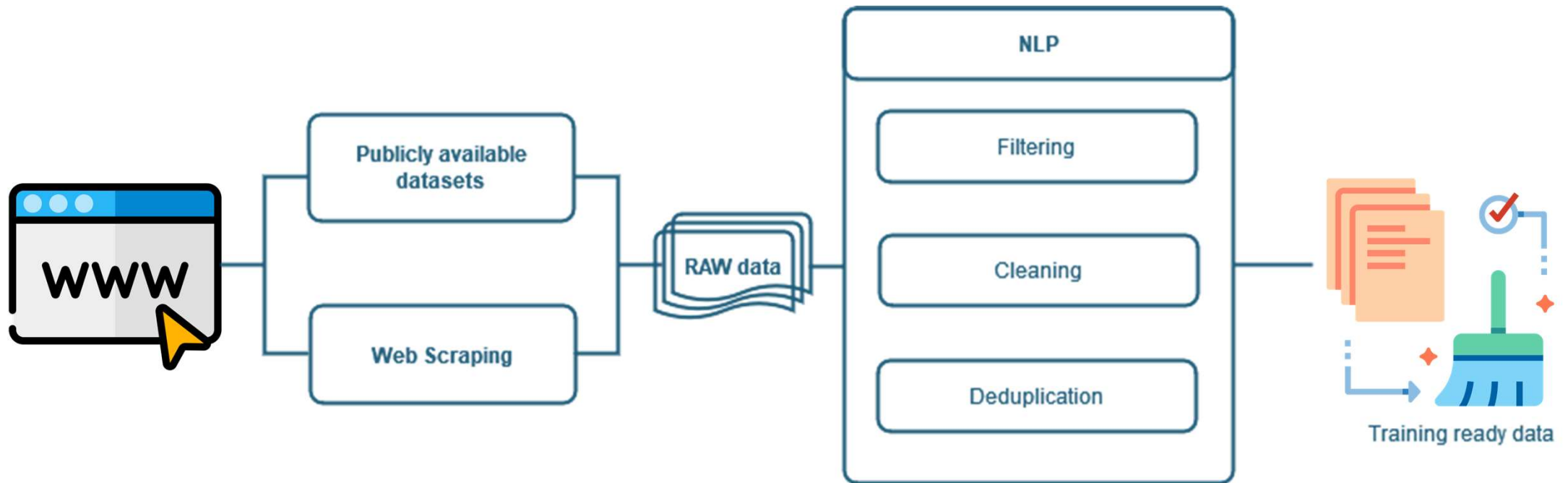


Step 1: Pre-training for the foundation LLM

Step 1 Pre-training



Data pre-processing



Data pre-processing | web example



Μαραθώνιος ληματογράφησης με αφορμή την Παγκόσμια Ημέρα των Ρομά
1 έως 15 Απριλίου 2025



[κλείσιμο]

Μεσολιθικός οικισμός Μαρουλάς

ΣΧ 6 γλώσσες ▾

Περιεχόμενα

απόκρυψη

Αρχή

Γενικά στοιχεία

▾ Αρχιτεκτονικές κατασκευές

Ταφικές κατασκευές

Λίθινα εργαλεία

Κατάλοιπα ζώων και φυτών

Παραπομπές

Βιβλιογραφικές πηγές

Λήμμα **Συζήτηση**

Ανάγνωση

Επεξεργασία

Επεξεργασία κώδικα

Προβολή ιστορικού

Εργαλεία ▾

Εμφάνιση

απόκρυψη

Κείμενο

☐ Μικρά

☒ Τυπικά

☐ Μεγάλα

Πλάτος

☒ Τυπικό

☐ Ευρύ

Χρώμα (beta)

☐ Αυτόματο

☒ Φωτεινό

☐ Σκοτεινό

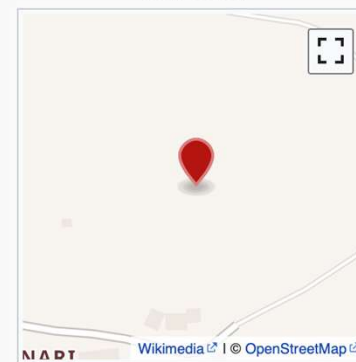
Ο **μεσολιθικός οικισμός Μαρουλάς** βρίσκεται στην **Κύθνο** κοντά στον σύγχρονο οικισμό **Λουτρά** και χρονολογείται στην 9η χιλιετία π.Χ, μεταξύ του 8.800 έως 8.600 π.Χ.^[1] Σύμφωνα με τον ανασκαφέα, η θέση χρησιμοποιήθηκε για λίγους αιώνες κατά την αρχή της 9ης π.Χ. χιλιετίας.^[1] Ο οικισμός διεκδικεί το πρωτείο του αρχαιότερου οικισμού στη νησιωτική περιοχή του Αιγαίου Πελάγους.^[2]

Η θέση αυτή είχε επισημανθεί από το 1975 ως προ-νεολιθικός οικισμός όμως η άποψη εκείνη αμφισβητήθηκε έντονα μέχρι την πρώτη έρευνα στον χώρο το 1995.^{[3][4]} Η ανασκαφική έρευνα συνεχίστηκε μεταξύ των ετών 2001 έως 2005 από το Πανεπιστήμιο του Αιγαίου και την ΚΑ' Εφορεία των Κυκλάδων καλύπτοντας την έκταση των 2.500 τμ.^[3] Η ανασκαφική

Συντεταγμένες:  37°26′47.900″N 24°25′43.453″E

Μεσολιθικός οικισμός Μαρουλάς Κύθνου

Μαρουλάς



Είδος

αρχαιολογική θέση



Data pre-processing | web example

```
1 <!DOCTYPE html>
2 <html class="client-nojs vector-feature-language-in-header-enabled vector-feature-language-in-main-page-header-disabled vector-feature-page-tools-pinned-disabled vector-
3 <head>
4 <meta charset="UTF-8">
5 <title>Μεσολιθικός οικισμός Μαρουλάς - Βικιπαίδεια</title>
6 <script>(function(){var className="client-js vector-feature-language-in-header-enabled vector-feature-language-in-main-page-header-disabled vector-feature-page-tools-pi
7 RLSTATE={"ext.globalCssJs.user.styles":"ready","site.styles":"ready","user.styles":"ready","ext.globalCssJs.user":"ready","user":"ready","user.options":"loading","ext.
8 <script>(RLQ=window.RLQ||[]).push(function(){mw.loader.impl(function(){return["user.options@12s5i",function($,jQuery,require,module){mw.user.tokens.set({"patrolToken":'
9 }})});});</script>
10 <link rel="stylesheet" href="/w/load.php?lang=el&modules=ext.cite.styles%7Cext.dismissableSiteNotice.styles%7Cext.kartographer.style%7Cext.uls.interlanguage%7Cext.
11 <script async="" src="/w/load.php?lang=el&modules=startup&only=scripts&raw=1&skin=vector-2022"></script>
12 <meta name="ResourceLoaderDynamicStyles" content="">
13 <link rel="stylesheet" href="/w/load.php?lang=el&modules=site.styles&only=styles&skin=vector-2022">
14 <meta name="generator" content="MediaWiki 1.44.0-wmf.23">
15 <meta name="referrer" content="origin">
16 <meta name="referrer" content="origin-when-cross-origin">
17 <meta name="robots" content="max-image-preview:standard">
18 <meta name="format-detection" content="telephone=no">
19 <meta name="viewport" content="width=1120">
20 <meta property="og:title" content="Μεσολιθικός οικισμός Μαρουλάς - Βικιπαίδεια">
21 <meta property="og:type" content="website">
22 <link rel="preconnect" href="//upload.wikimedia.org">
23 <link rel="alternate" media="only screen and (max-width: 640px)" href="//el.m.wikipedia.org/wiki/%CE%9C%CE%B5%CF%83%CE%BF%CE%BB%CE%B9%CE%B8%CE%B9%CE%BA%CF%8C%CF%82_%CE%
24 <link rel="alternate" type="application/x-wiki" title="Επεξεργασία" href="/w/index.php?title=%CE%9C%CE%B5%CF%83%CE%BF%CE%BB%CE%B9%CE%B8%CE%B9%CE%BA%CF%8C%CF%82_%CE%BF%
25 <link rel="apple-touch-icon" href="/static/apple-touch/wikipedia.png">
26 <link rel="icon" href="/static/favicon/wikipedia.ico">
27 <link rel="search" type="application/opensearchdescription+xml" href="/w/rest.php/v1/search" title="Βικιπαίδεια (el)">
28 <link rel="EditURI" type="application/rsd+xml" href="//el.wikipedia.org/w/api.php?action=rsd">
29 <link rel="canonical" href="https://el.wikipedia.org/wiki/%CE%9C%CE%B5%CF%83%CE%BF%CE%BB%CE%B9%CE%B8%CE%B9%CE%BA%CF%8C%CF%82_%CE%BF%CE%B9%CE%BA%CF%8C%CF%83%CE%BC%CF%8C%
30 <link rel="license" href="https://creativecommons.org/licenses/by-sa/4.0/deed.el">
31 <link rel="alternate" type="application/atom+xml" title="Βικιπαίδεια ροή Atom" href="/w/index.php?title=%CE%95%CE%B9%CE%B4%CE%B9%CE%BA%CF%8C:%CE%A0%CF%81%CF%8C%CF%83%CF%
32 <link rel="dns-prefetch" href="//meta.wikimedia.org" />
33 <link rel="dns-prefetch" href="auth.wikimedia.org">
34 </head>
35 <body class="skin--responsive skin-vector skin-vector-search-vue mediawiki ltr sitedir-ltr mw-hide-empty-elt ns-0 ns-subject mw-editable page-Μεσολιθικός_οικισμός_Μαροι
36 <div class="vector-header-container">
37 <header class="vector-header mw-header">
38 <div class="vector-header-start">
39 <nav class="vector-main-menu-landmark" aria-label="Ιστότοπος">
40
41 <div id="vector-main-menu-dropdown" class="vector-dropdown vector-main-menu-dropdown vector-button-flush-left vector-button-flush-right" title="Κύριο μενού" >
42 <input type="checkbox" id="vector-main-menu-dropdown-checkbox" role="button" aria-haspopup="true" data-event-name="ui.dropdown-vector-main-menu-dropdown" class="vec
43 <label id="vector-main-menu-dropdown-label" for="vector-main-menu-dropdown-checkbox" class="vector-dropdown-label cdx-button cdx-button--fake-button cdx-button--fal
44
45 <span class="vector-dropdown-label-text">Κύριο μενού</span>
46 </label>
```



Data pre-processing | web example

Ο μεσολιθικός οικισμός Μαρουλάς βρίσκεται στην Κύθνο κοντά στον σύγχρονο οικισμό Λουτρά και χρονολογείται στην 9η χιλιετία π.Χ., μεταξύ του 8.800 έως 8.600 π.Χ. Σύμφωνα με τον ανασκαφέα, η θέση χρησιμοποιήθηκε για λίγους αιώνες κατά την αρχή της 9ης π.Χ. χιλιετίας. Ο οικισμός διεκδικεί το πρωτείο του αρχαιότερου οικισμού στη νησιωτική περιοχή του Αιγαίου Πελάγους.

Η θέση αυτή είχε επισημανθεί από το 1975 ως προ-νεολιθικός οικισμός όμως η άποψη εκείνη αμφισβητήθηκε έντονα μέχρι την πρώτη έρευνα στον χώρο το 1995. Η ανασκαφική έρευνα συνεχίστηκε μεταξύ των ετών 2001 έως 2005 από το Πανεπιστήμιο του Αιγαίου και την ΚΑ' Εφορεία των Κυκλάδων καλύπτοντας την έκταση των 2.500 τμ. Η ανασκαφική δραστηριότητα αποκάλυψε αρχιτεκτονικές κατασκευές, ταφές, αποθέτες με καύσεις, λίθινα εργαλεία και διατροφικά κατάλοιπα ζώων και φυτών.

Γίνονται προσπάθειες για αποκατάσταση, προστασία και αξιοποίηση του οικισμού ως αρχαιολογικό χώρο.

Γενικά στοιχεία

Ο Μαρουλάς είναι ο μοναδικός μεσολιθικός οικισμός σε ανοιχτό επίπεδο χώρο του οποίου σώζονται αρχιτεκτονικά στοιχεία. Ενώ η επιφάνεια της θάλασσας βρισκόταν 40–50 μέτρα (ή 50–60 μέτρα) πιο χαμηλά στη μεσολιθική περίοδο, σήμερα ο οικισμός βρίσκεται δίπλα στο κύμα και έχει πλέον καταστραφεί σημαντικό μέρος του. Φαίνεται πως είναι ο μοναδικός μεσολιθικός οικισμός στο Αιγαίο που εν μέρει διασώθηκε από την άνοδο της στάθμης της θάλασσας καθώς τότε δεν ήταν παραθαλάσσιος.

Αρχιτεκτονικές κατασκευές

Με βάση τα ανασκαφικά δεδομένα, τα οικήματα του Μαρουλά είχαν κυκλικό σχήμα με διάμετρο 3 έως 4 μ και το δάπεδό τους επιστρωνόταν με πέτρες, ενώ κάποια από αυτά χρησιμοποιήθηκαν περισσότερες από μία φορές. Οι κατασκευές αλλού εντοπίζονται συγκεντρωμένες, όπως συμβαίνει στο κεντρικό τμήμα του οικισμού και αλλού κατανέμονται πιο αραιά. Αριθμούνται 31 (ή 15) τέτοιες κατασκευές. Οι περισσότερες και πιο κατεστραμμένες οικίες βρίσκονται στην ανατολική πλευρά του οικισμού, η οποία γειτνιάζει με τη θάλασσα.

Ο τύπος του κυκλικού οικήματος στον Μαρουλά, κατοικήθηκε από ομάδες θηρευτών–τροφοσυλλεκτών με αμφισβητούμενο, προς το παρόν βαθμό μονιμότητας αν και φαίνεται πως γίνονταν ανακατασκευές, κάτι που δηλώνει μακροχρόνια χρήση τους. Παρόμοιοι τύποι οικημάτων δεν συναντώνται στο πλαίσιο της Μεσολιθικής περιόδου της Νοτιανατολικής Ευρώπης.

Ταφικές κατασκευές

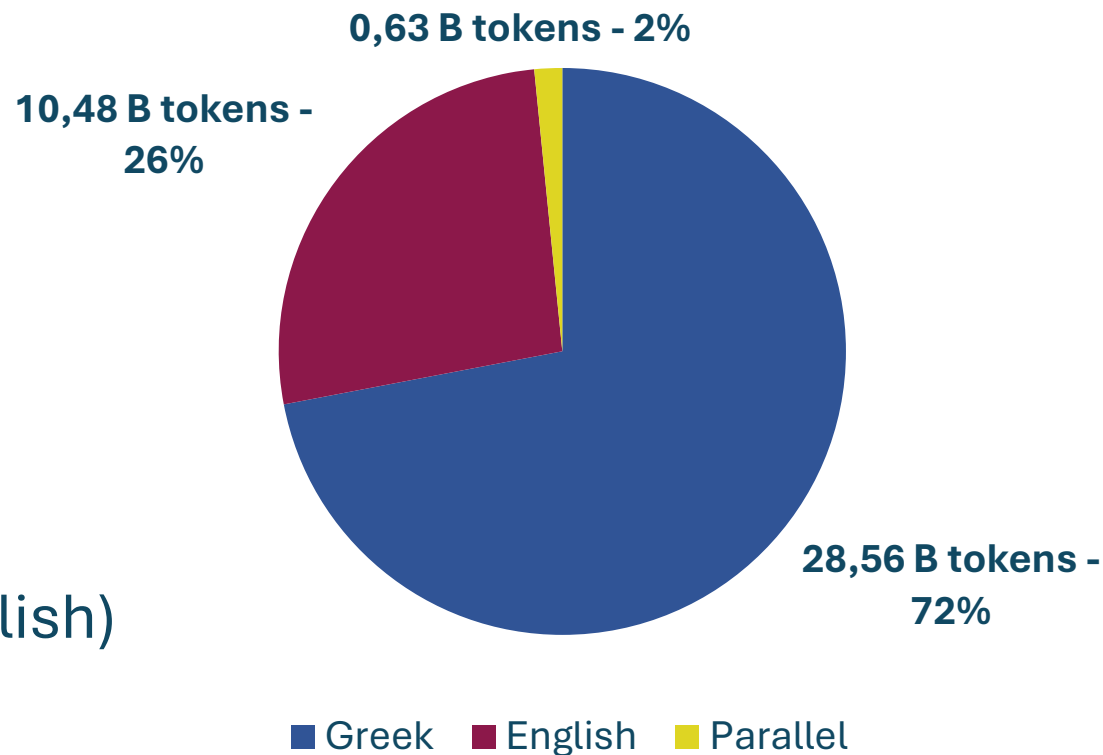
Στον Μαρουλά φαίνεται πως υπάρχει ποικιλομορφία ως προς τις κατασκευές χώρων ταφής αφού δύο τάφοι είχαν σκαφτεί στο μαλακό σχιστόλιθο, δύο άλλοι είχαν τη μορφή απλών αβαθών λάκκων ελλειψοειδούς σχήματος με μικρές πέτρες περιμετρικά και μία ταφή εντοπίστηκε σε αβαθή λάκκο στην περίμετρο του οποίου υπήρχαν πέτρες που στήριζαν μια μεγάλη πλάκα. Σε όλες τις περιπτώσεις ο νεκρός ήταν τοποθετημένος σε συνεσταλμένη στάση και δεν συνοδευόταν από κτερίσματα. Η παρουσία μακρών οστών και μιας σιαγόνας κάτω από τα δάπεδα δύο οικημάτων, όπως επίσης και της συγκέντρωσης οστών πάνω στο δάπεδο μιας κατασκευής, αλλά και η απουσία



Composition of pre-training data | Meltemi

Sources include:

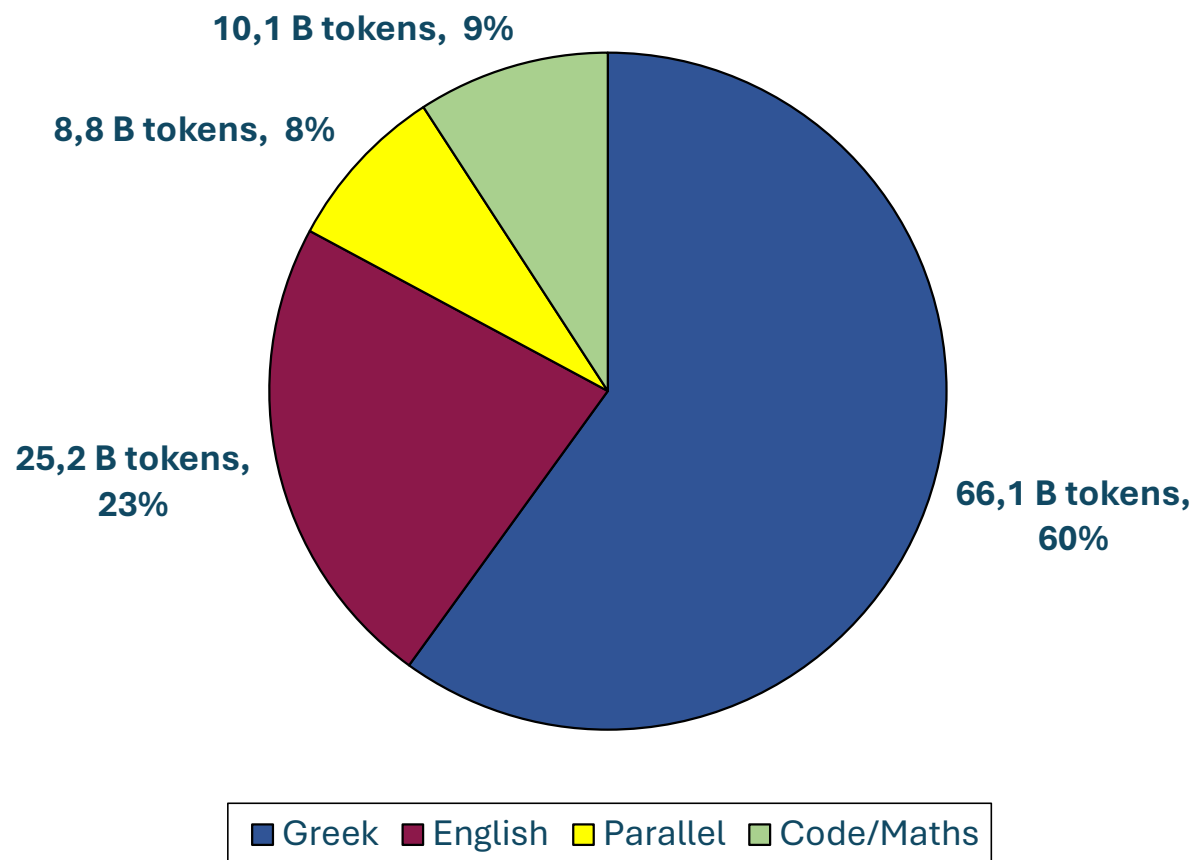
- / Wikipedia
- / Theses & Dissertations
- / Parliamentary Proceedings
- / Parallel data (Greek <-> English)



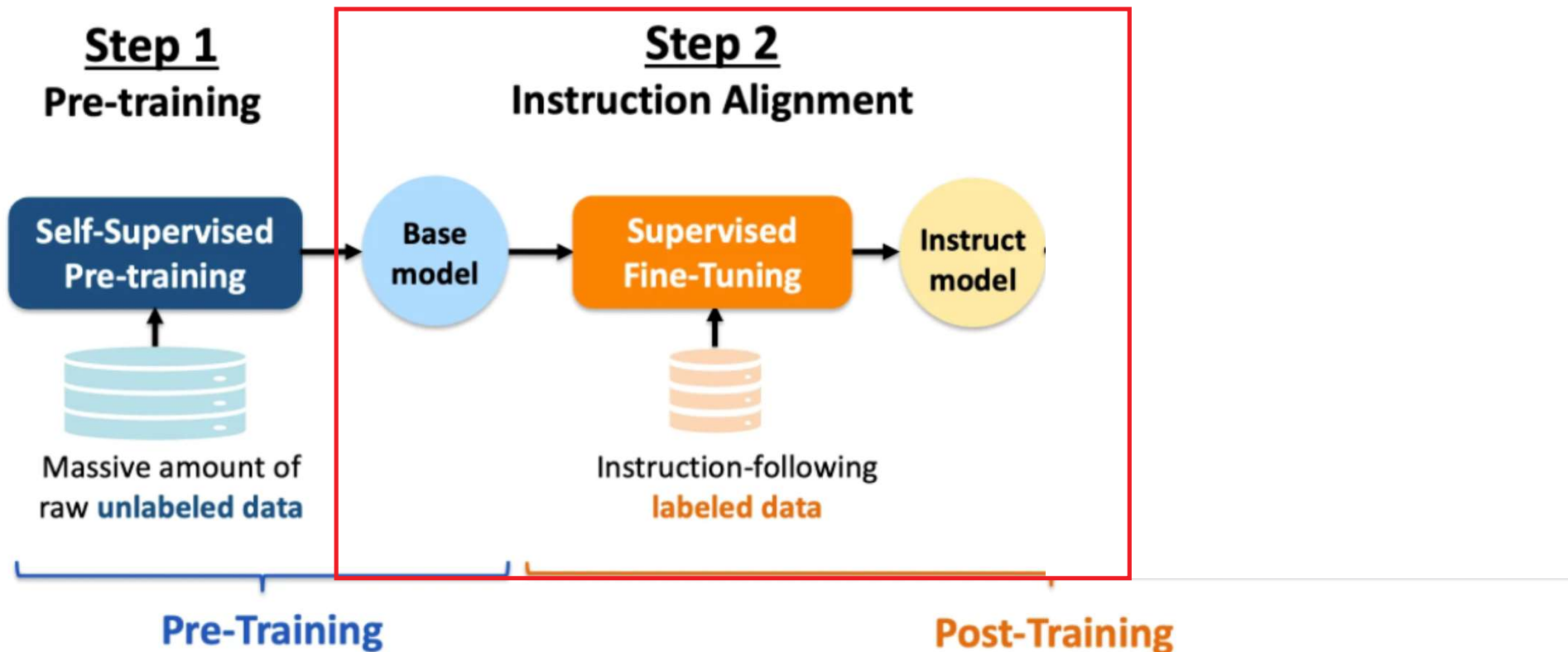
Composition of pre-training data | Krikri

Sources include:

- / Wikipedia
- / Theses & Dissertations
- / Parliamentary Proceedings
- / Greek & EU Laws
- / Maths & Code
- / Parallel data (Greek <-> English)



Supervised Fine-tuning of the foundation LLM



Instruction following data



Below is an instruction that describes a task, paired with an input that provides further context. Write a response that appropriately completes the request.

```
### Instruction:
{instruction}
```

```
### Input:

```

Response:

```
"instruction": "What are the three primary colors?",  
"input": "",  
"output": "The three primary colors are red, blue, and yellow."
```

```
"instruction": "Classify the following into animals, plants, and minerals",

"output": "Oak tree: Plant\n Copper ore: Mineral\n Elephant: Animal"
```



Fine-tuning Data (1)

USER:

Μπορείς να μου εξηγήσεις την τριεπίπεδη εξελικτική δομή του εγκεφάλου των θηλαστικών, ξεκινώντας από το κατώτερο επίπεδο;

ASSISTANT:

Βέβαια, η τριεπίπεδη εξελικτική δομή (ή ιεραρχική οργάνωση) του εγκεφάλου των θηλαστικών (και κατ' επέκταση των ανώτερων σπονδυλωτών) έχει ως εξής:

****1. Φυλογενετική Προέλευση (Κατώτερο Επίπεδο) – Ριζωματική Οντογένεση:****

* Ξεκινάμε από τις ****αρχαιότερες δομές**** που διατηρούνται σε όλα τα σπονδυλωτά και αποτελούν τη βάση της εγκεφαλικής εξέλιξης.

– ****Εγκεφαλικό Στέλεχος (Brain Stem):**** Περιέχει βασικές λειτουργίες όπως αναπνοή, καρδιακή λειτουργία, αυτόνομη ρύθμιση. Περιλαμβάνει τον μεσεγκέφαλο, μυελό, εγκεφαλικά νεύρα.

– ****Υποθάλαμος:**** Κεντρικό τμήμα του διημισφαιρικού εγκεφάλου, ρυθμίζει βασικές ομοιόστασεις (πείνα, δίψα, θερμοκρασία), ενδοκρινείς αδένες και κύκλους ύπνου-εγρήγορσης.

****2. Εγκεφαλικός Φλοιός (Μεσαίο Επίπεδο):****

* ****Εξελίχθηκε** πάνω από το εγκεφαλικό στέλεχος** και αποτελεί την "επιδερμίδα" του εγκεφάλου.

– ****Νεοφλοιός (Neocortex):**** Χαρακτηριστικό των θηλαστικών, ανώτερη γνωστική λειτουργία, αισθητηριακή επεξεργασία, γλώσσα, μνήμη, λογική σκέψη.

– ****Αρχαιότερος Φλοιός:**** Περιλαμβάνει παλαιότερες δομές όπως ιππόκαμπος (μνήμη), αμυγδαλή (συναισθήματα).

****3. Ανώτερο Επίπεδο – Νευρωνικά Δίκτυα και Σύνδεση:****

* ****Συνειρμική Επεξεργασία:**** Οι κατώτερες δομές (1ο επίπεδο) τροφοδοτούν με πληροφορίες τον φλοιό (2ο επίπεδο) ο οποίος με τη σειρά του δημιουργεί πολύπλοκα ****νευρωνικά κυκλώματα**** και ****δίκτυα****.

* ****Λειτουργική Ολοκλήρωση:**** Η τελική λειτουργία προκύπτει από την ****ολοκληρωμένη δραστηριότητα**** όλων των επιπέδων, με το καθένα να έχει εξειδικευμένο ρόλο στην επεξεργασία πληροφοριών και την παραγωγή συμπεριφοράς.

Αυτή η τριεπίπεδη οργάνωση επιτρέπει στα θηλαστικά να εξελίσσονται από βασικές αισθήσεις και αυτόματες αντιδράσεις σε σύνθετες γνωστικές λειτουργίες και προσαρμοστική συμπεριφορά.

Η κατανόηση αυτής της δομής είναι θεμελιώδης για την νευροεπιστήμη και την κατανόηση της ανθρώπινης νοημοσύνης και συμπεριφοράς.

Copyright 2020. All rights reserved.



Fine-tuning Data (2)

USER:

Αν υποθέσουμε ότι ένα πεύκο στο Στρεβλό Δάσος έχει ύψος 15 μέτρα και η καμπύλωσή του ξεκινά 0.5 μέτρα από το έδαφος με εκτροπή 2.5 μέτρα προς τα βόρεια, υπολόγισε τη γωνία κλίσης του κορμού στο σημείο καμπύλωσης.

ASSISTANT:

Για να υπολογίσουμε τη γωνία κλίσης (γωνία ως προς τον ορίζοντα) του κορμού του πεύκου στο σημείο καμπύλωσης, θα χρησιμοποιήσουμε τριγωνομετρία.

Ας ορίσουμε:

- $h_{\text{κορύφη}} = 15$ μέτρα (συνολικό ύψος)
- $h_{\text{εδάφους}} = 0.5$ μέτρα (απόσταση από έδαφος μέχρι το σημείο έναρξης καμπύλωσης)
- $d_{\text{εκτροπή}} = 2.5$ μέτρα (εκτροπή προς τα βόρεια)
- Θέλουμε τη γωνία θ (γωνία κλίσης)

Από το ορθογώνιο τρίγωνο που σχηματίζεται:

- Η υποτείνουσα είναι $L = h_{\text{κορύφη}} - h_{\text{εδάφους}} = 15 - 0.5 = 14.5$ μέτρα
- Η κάθετη πλευρά είναι $y = d_{\text{εκτροπή}} = 2.5$ μέτρα
- Η οριζόντια πλευρά είναι $x = h_{\text{εδάφους}} = 0.5$ μέτρα

Χρησιμοποιώντας την εφαπτομένη: $\tan(\theta) = \frac{y}{x} = \frac{2.5}{0.5} = 5$

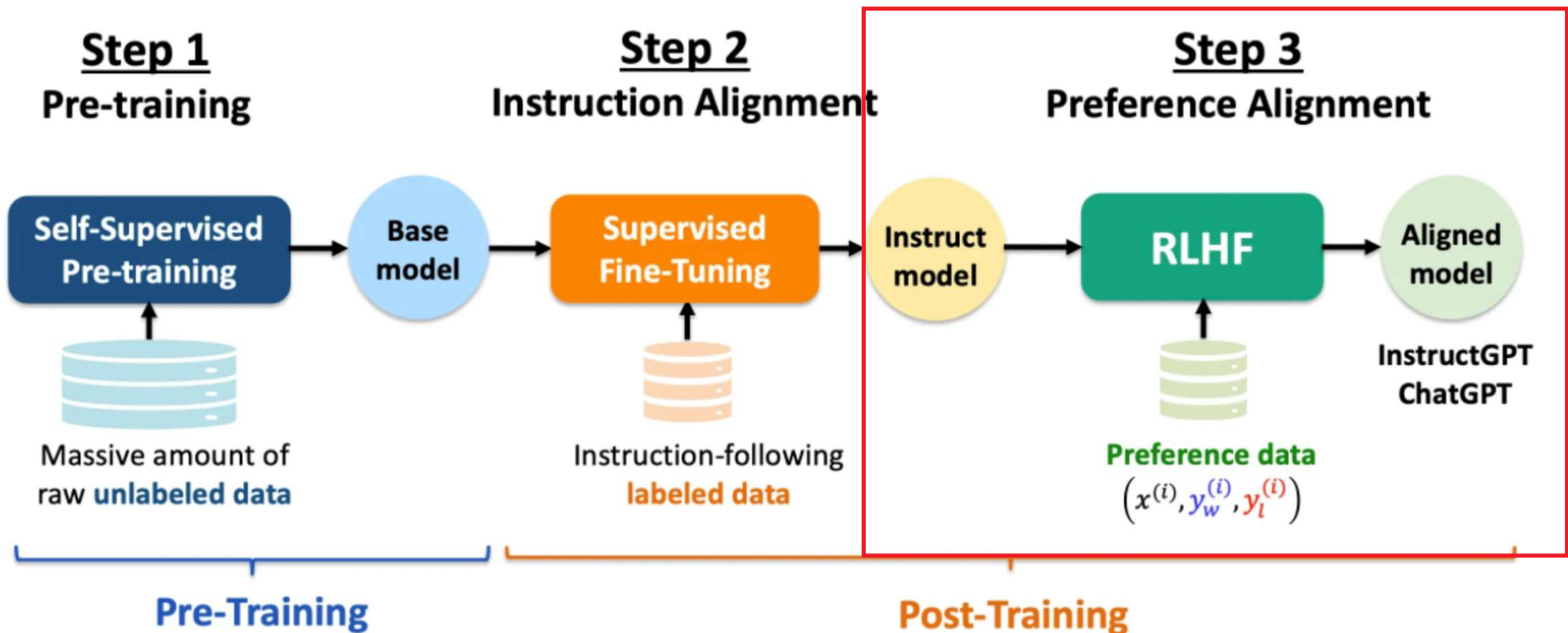
Άρα: $\theta = \arctan(5) \approx 78.69$ μοίρες

Η γωνία κλίσης του κορμού στο σημείο καμπύλωσης είναι περίπου 78.69° (με ακρίβεια δύο δεκαδικών ψηφίων).

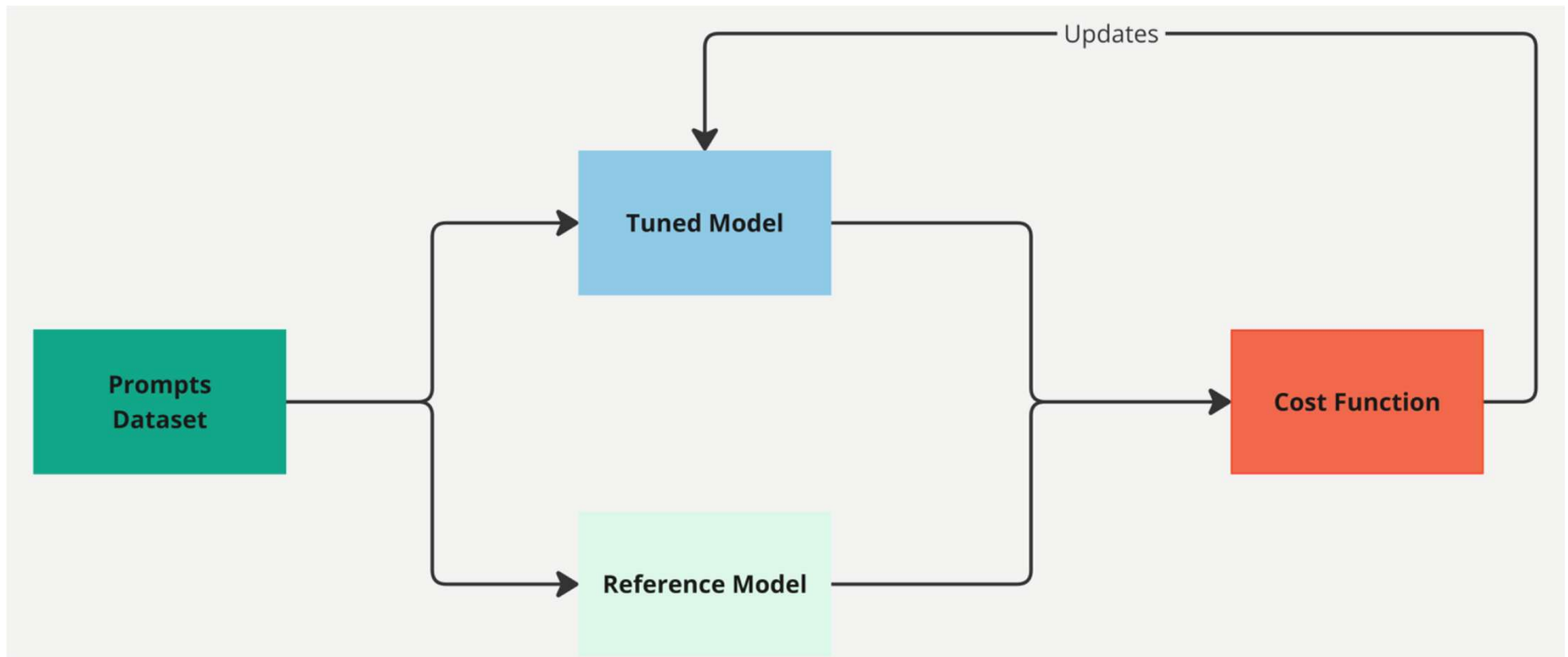
Η γωνία αυτή μας δίνει την κλίση ως προς τον ορίζοντα και είναι η ζητούμενη απάντηση.



Step 3: Preference alignment



Step 3: Preference alignment



Preference Alignment Data (1)

USER:

Έχω έναν προϋπολογισμό 322.683,61€ χωρίς ΦΠΑ για την προμήθεια καυσίμων. Με ΦΠΑ 24%, ποιο θα είναι το συνολικό ποσό και πόσα είναι τα χρήματα του ΦΠΑ;

CHOSEN:

Ας το λύσουμε βήμα-βήμα:

1. **Ποσό χωρίς ΦΠΑ:** 322.683,61€
2. **Ποσοστό ΦΠΑ:** 24% = 0,24
3. **Υπολογισμός ΦΠΑ:**
* $\text{ΦΠΑ} = 322.683,61\text{€} \times 0,24 = 77.444,07\text{€}$
4. **Συνολικό ποσό με ΦΠΑ:**
* $\text{Συνολικό} = \text{Ποσό χωρίς ΦΠΑ} + \text{ΦΠΑ}$
* $\text{Συνολικό} = 322.683,61\text{€} + 77.444,07\text{€} = \textbf{400.127,68\text{€}}$

Άρα:

- * Συνολικό ποσό με ΦΠΑ: 400.127,68€
- * Ποσό ΦΠΑ: 77.444,07€

REJECTED:

Ας λύσουμε το πρόβλημα βήμα-βήμα:

- 1) Έστω x το ποσό με ΦΠΑ 24%
- 2) Για το ποσό χωρίς ΦΠΑ (322.683,61€) ισχύει η σχέση:
* $x = \text{ποσό_με_ΦΠΑ} = (\text{ποσό_χωρίς_ΦΠΑ} \times 1.24)$
* $322.683,61 = 322.683,61 \times 1.24$ [διαιρέσαμε και τα δύο μέλη με 1.24 για να βρούμε το αρχικό ποσό]



Preference Alignment Data (2)

USER

Το άθροισμα των ψηφίων ενός τριψήφιου αριθμού είναι 12 και ο αριθμός διαιρείται με το 4. Ποιος είναι ο μεγαλύτερος τέτοιος ακέραιος;

ASSISTANT

<think>

Ας αναλύσω το πρόβλημα βήμα προς βήμα:

1) Πρώτα, χρειάζεται να βρω τους περιορισμούς:

- * Ο αριθμός είναι τριψήφιος (100–999)
- * Το άθροισμα των ψηφίων είναι 12
- * Ο αριθμός διαιρείται με το 4

2) Για να διαιρείται ένας αριθμός με το 4, πρέπει τα δύο τελευταία ψηφία του να σχηματίζουν αριθμό διαιρετό με το 4.

3) Έστω ότι τα ψηφία του αριθμού είναι a , b , c

- * $a + b + c = 12$
- * $100a + 10b + c$ πρέπει να διαιρείται με 4
- * $a \geq 1$ (αφού είναι τριψήφιος)

4) Για να είναι ο μεγαλύτερος δυνατός αριθμός:

- * Το πρώτο ψηφίο (a) πρέπει να είναι όσο μεγαλύτερο γίνεται
- * Δεδομένου ότι $a + b + c = 12$, το a δεν μπορεί να είναι μεγαλύτερο από 9

Ας δοκιμάσω ξεκινώντας με $a = 9$:

- * Αν $a = 9$, τότε $b + c = 3$
- * Τα b και c πρέπει να δίνουν διαιρετό με 4 στα δύο τελευταία ψηφία
- * Οι πιθανοί συνδυασμοί για b, c που αθροίζουν σε 3 είναι: $(3, 0)$, $(2, 1)$, $(1, 2)$, $(0, 3)$
- * Από αυτούς, πρέπει να ελέγξω ποιοι δίνουν διαιρετό με 4

Ελέγχω:

- * $90 + 3 = 93$ (δεν διαιρείται με 4)
- * $92 + 1 = 93$ (δεν διαιρείται με 4)
- * $91 + 2 = 92$ (διαιρείται με 4)
- * $90 + 3 = 93$ (δεν διαιρείται με 4)

Άρα ο αριθμός 912 είναι ο μεγαλύτερος που ικανοποιεί όλες τις συνθήκες.



Evaluating Meltemi & Krikri

Introduced an LLM evaluation suite for Greek

- **Language Understanding and Reasoning**
- **General Question Answering** (also Medical)
- **Truthfulness**
- **Instruction Following** (e.g. write in 3 paragraphs)
- **Code & Maths**
- **General Chat Capabilities**
- **Translation** (even for Ancient Greek)

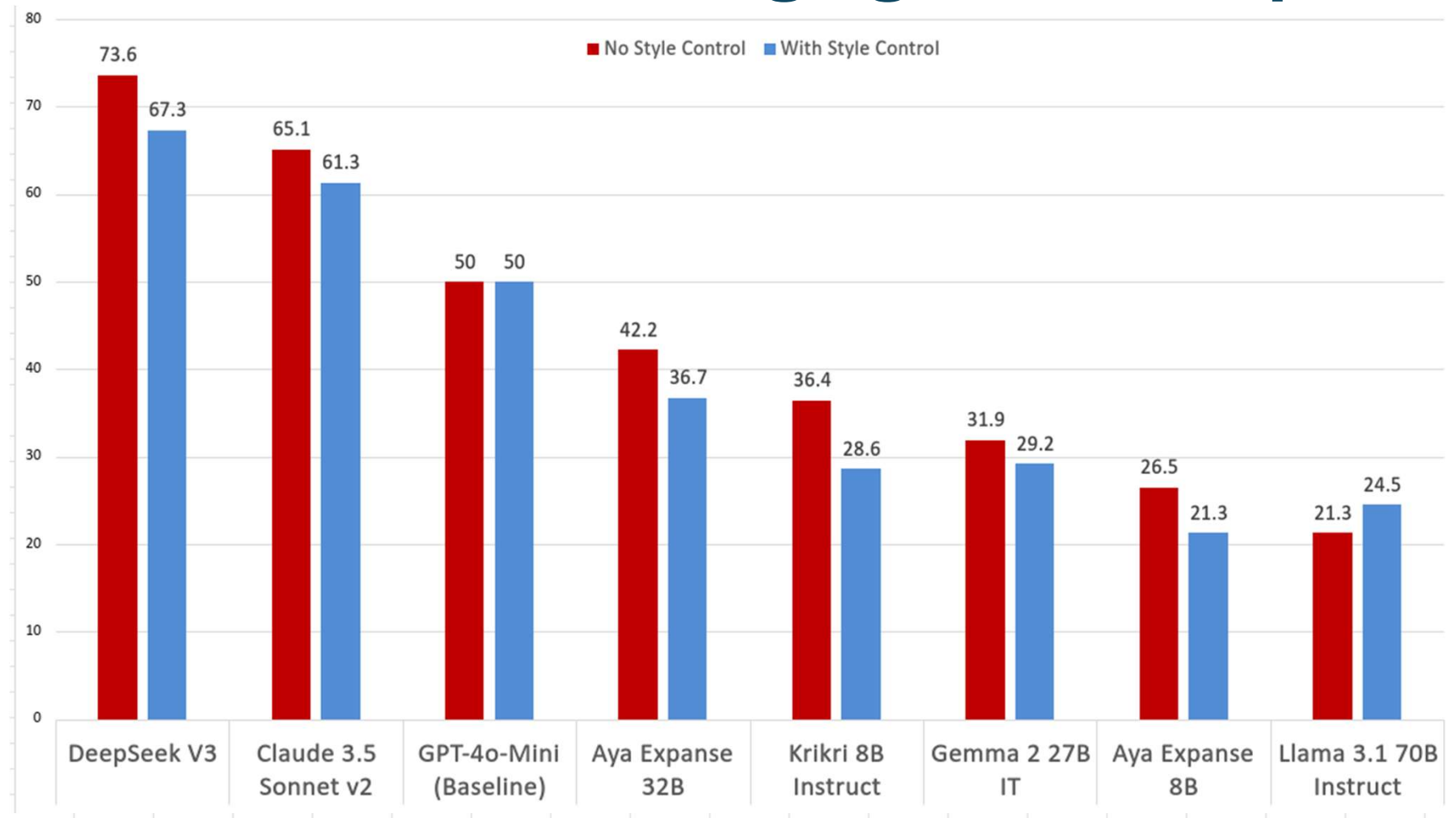


LLM evaluation datasets

Name	# Examples	Description
ARC	7.78K	ARC (Clark et al., 2018), a dataset of science exam questions (with typically four answer options) partitioned into a Challenge and an Easy Set of 2.6K and 5.2K questions
Truthful QA	817	Truthful QA (Lin et al., 2022), a dataset of questions that are crafted so that some humans would answer wrongly due to a false belief or misconception.
HellaSwag	59.8K	HellaSwag dataset (Zellers et al., 2019) for commonsense NLI
MMLU	15.9K	MMLU dataset (Hendrycks et al., 2021) of multiple-choice questions from 57 tasks including elementary mathematics, history, computer science, law, etc.
Belebele	900x122	Belebele (Bandarkar et al., 2023), a multiple-choice machine reading comprehension dataset covering 122 language variants.
Greek Medical Multiple Choice QA	2.03K	Multiple choice questions extracted from past medical exams of the Greek National Academic Recognition and Information Center available at https://www.doatap.gr



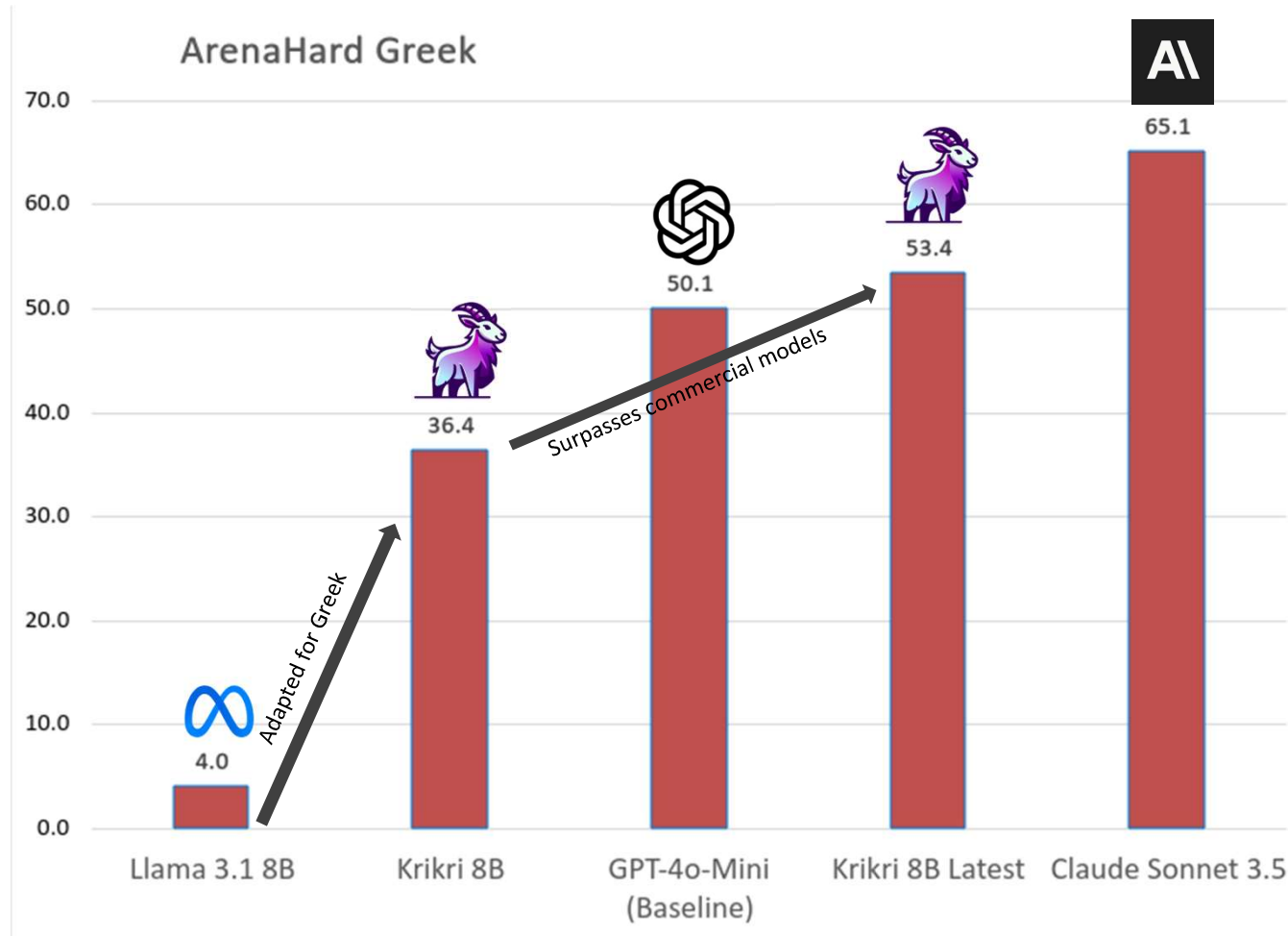
Arena Hard-500 challenging & diverse prompts



Connect 2025 – Athens, June 5, 2025



Arena Hard-500 challenging & diverse prompts



Connect 2025 – Athens, June 5, 2025



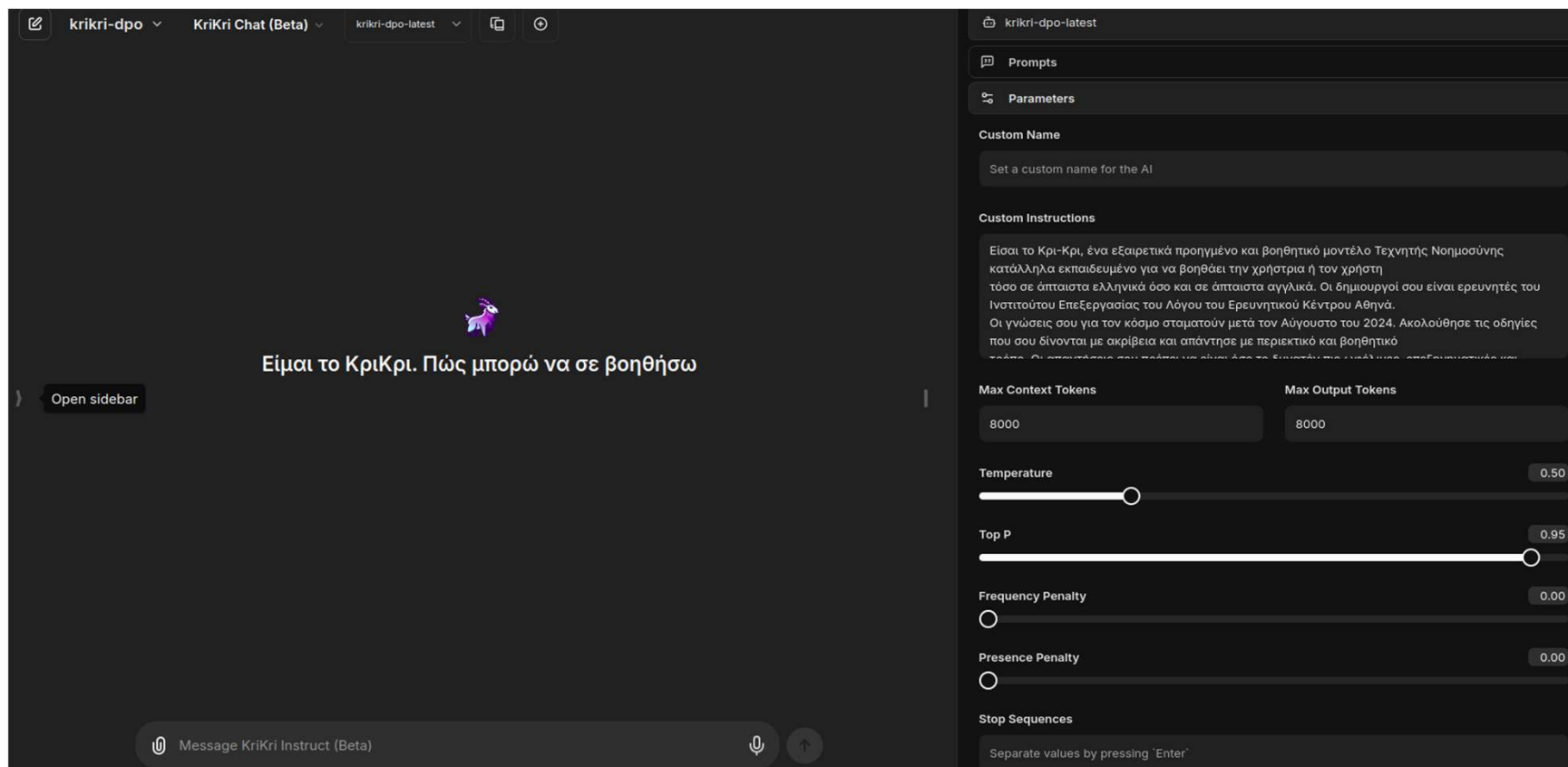
Outcomes



- / 2 LLMs: Meltemi-7B-Instruct-v1.5, Llama-Krikri-8B-Instruct
- / Krikri surpasses bigger LLMs for Greek & **competes with commercial LLMs, even of bigger size**
- / Released open versions of all models in Hugging Face.
Also provided quantized versions to run locally on consumer hardware
- / Created evaluation suite with 12+ test sets for Greek LLMs



Chatting with Krikri: chat.ilsp.gr



<https://chat.ilsp.gr>

Connect 2025 – Athens, June 5, 2025



AI powered applications



AI-powered assistant in education

Teacher Mate

✕ Εργαλεία


Συνομιλία Με Τον Teacher Mate

Διαχείριση Τάξεων


Τα Τεστ Μου


Το Περιεχόμενό Μου


Προφίλ





Αποσύνδεση


**Δημιουργία τεστ αξιολόγησης**
Δημιουργήστε τεστ που ταιριάζουν στις ανάγκες σας και αναθέστε τα στην τάξη σας.


**Δημιουργία Κειμένων**
Δημιουργήστε ελκυστικά, πρωτότυπα κείμενα προσαρμοσμένα σε οποιοδήποτε θέμα, είδος ή επίπεδο μαθητών.


**Εξερεύνηση Εννοιών**
Δημιουργήστε υλικό για να βοηθήσετε τους μαθητές να κατανοήσουν, να εφαρμόσουν και να εξερευνήσουν έννοιες με κριτικό τρόπο μέσω ειδικά...


**Δημιουργία Σχεδίου Μαθήμα...**
Δημιουργήστε ολοκληρωμένα σχέδια μαθημάτων με λεπτομερείς δραστηριότητες και εκπαιδευτικό υλικό βάσει εκπαιδευτικών προσεγγίσεων.


**Δημιουργία Κουίζ**
Δημιουργήστε κουίζ με ποικιλία ερωτήσεων για την τάξη και για αυτοαξιολόγηση.


**Υποστήριξη Βαθμολόγησης ...**
Βαθμολογήστε και αξιολογήστε απαντήσεις μαθητών σε εργασίες με βάση εξατομικευμένα κριτήρια και με λεπτομερή ανατροφοδότηση.


**Σχεδιασμός Πρότζεκτ**
Δημιουργήστε ενδιαφέροντα, συνεργατικά πρότζεκτ για μαθητές με βάση οποιοδήποτε θέμα και για οποιαδήποτε τάξη.

**Κατάρτιση Κριτηρίων Αξιολό...**
Καθορίστε κριτήρια για την αξιολόγηση εργασιών.

**Σχεδιασμός Παρουσιάσεων**
Δημιουργήστε ελκυστικές και λεπτομερείς εκπαιδευτικές παρουσιάσεις με διαδραστικά στοιχεία.

**Προσαρμογή Κειμένου**
Προσαρμόστε οποιοδήποτε κείμενο στο επίπεδο, τις δεξιότητες και τις εκπαιδευτικές ανάγκες των μαθητών σας.

**Χρονολογική Σειρά Γεγονότων**
Βάλτε φανταστικά, ιστορικά ή επιστημονικά γεγονότα σε χρονολογική σειρά για να βοηθήσετε τους μαθητές να κατανοήσουν χρονολογίες και...



AI4EDU: Conversational AI assistant for teaching and learning. Co-funded by the European Commission, European Education and Culture Executive Agency Project: 101087451 – AI4EDU – ERASMUS-EDU-2022-PI-FORWARD

Connect 2025 – Athens, June 5, 2025

AI-powered assistant in education

Αποσύνδεση

Ai4edu

πικνα στο ακρο 5' και 3' (που ονομαζονται βροχοί ραστωτικών φιλίων (kissing stem loops)), κυκλοφορεί το mRNA χωρίς να εμπλέκονται πρωτεΐνες.

Τα γονιδιώματα του ιού RNA (των οποίων οι + κλώνιοι μεταφράζονται ως mRNA) κυκλοφορούν επίσης συνήθως. Κατά τη διάρκεια της αναπαραγωγής του γονιδιώματος, η κυκλοφορία δρα για να αυξήσει τις ταχύτητες αναπαραγωγής του γονιδιώματος, ανακυκλώνοντας πολύ την πολυμεράση RNA που εξαρτάται από ιικό RNA.

Το mRNA (αγγελιοφόρο RNA) είναι ένα σημαντικό μόριο στη διαδικασία της πρωτεϊνώνθεσης. Η δομή του ώριμου ευκαρυωτικού mRNA περιλαμβάνει διάφορα τμήματα: την καλόντρε 5', τις μη μεταφραζόμενες περιοχές 5' και 3', την κωδικοποιητική περιοχή και την ουρά πολυ(A).

- Κωδικοποιητικές Περιοχές:** Αυτές περιέχουν κωδικόνια, τα οποία είναι τριάδες νουκλεοτιδίων που καθορίζουν τα αμινοξέα μιας πρωτεΐνης. Ξεκινούν με το κωδικόνιο έναρξης (συνήθως AUG) και τελειώνουν με ένα από τα κωδικόνια διακοπής (UAG, UAA, UGA).
- Αμετάφραστες Περιοχές (UTRs):** Αυτές οι περιοχές δεν μεταφράζονται σε πρωτεΐνες αλλά παίζουν ρόλο στη σταθερότητα, τον εντοπισμό και την αποτελεσματικότητα της μετάφρασης του mRNA. Οι UTRs μπορούν να επηρεάσουν τη διάρκεια ζωής του mRNA και την ικανότητα του ριβοσώματος να συνδέεται με αυτό.
- Ουρά πολυ(A):** Αυτή είναι μια αλληλουχία αδενίνης στο 3' άκρο του mRNA που βοηθά στην εξαγωγή του από τον πυρήνα και προστατεύει το mRNA από την αποικοδόμηση.
- Μονοιστρωνικό και Πολυιστρωνικό mRNA:** Το μονοιστρωνικό mRNA περιέχει πληροφορίες για μία μόνο πρωτεΐνη, ενώ το πολυιστρωνικό περιέχει πληροφορίες για πολλές πρωτεΐνες.
- Κυκλοφορία του mRNA:** Το mRNA μπορεί να σχηματίσει κυκλικές δομές που προάγουν την αποδοτική μετάφραση και διασφαλίζουν ότι μόνο άθικτο mRNA μεταφράζεται.

Αυτά τα στοιχεία συνεργάζονται για να διασφαλίσουν ότι το mRNA μεταφέρει σωστά τις γενετικές πληροφορίες από το DNA στις πρωτεΐνες.

Επιλέξτε αναφορά

Συνομιλία με τον Study Buddy...

Βιολογία Β'-Γ' Γυμνασίου (Βιβλίο Μαθητή)

άλλων ελεύθερων δεσορυβονουκλεοτιδίων. Τα νουκλεοτίδια αυτά ενώνονται αφενός με τις αζεωργωμένες βάσεις της παλιάς νουκλεοτιδικής αλυσίδας και αφετέρου μεταξύ τους, σχηματίζοντας μία νέα συμπληρωματική αλυσίδα.

Το αποτέλεσμα της αντιγραφής είναι ο σχηματισμός δύο διπλών μορίων DNA, καθένα από τα οποία αποτελείται από μία παλιά και μία νέα αλυσίδα. Αυτά τα μόρια –σε περίπτωση που δεν έχει συμβεί κάποιο «λάθος» στη συμπληρωματικότητα των βάσεων– είναι πανομοιότυπα τόσο μεταξύ τους όσο και με το αρχικό μόριο, δηλαδή έχουν την ίδια αλληλουχία νουκλεοτιδίων και συνεπώς τις ίδιες γενετικές πληροφορίες.

Μεταγραφή, μετάφραση – Έκφραση της γενετικής πληροφορίας

Όπως έχουμε ήδη αναφέρει, η σειρά των αμινοξέων στις πρωτεΐνες καθορίζεται από τις γενετικές πληροφορίες που περιέχονται σε συγκεκριμένα τμήματα του DNA, τα γονίδια. Η σύνθεση των πρωτεϊνών γίνεται στα ριβοσώματα του κυττάρου. Δεν είναι δυνατόν όμως να μεταφερθεί ολόκληρο το DNA στα ριβοσώματα, κάθε φορά που το κύτταρο συνθέτει μία συγκεκριμένη πρωτεΐνη. Η συγκεκριμένη γενετική πληροφορία μεταφέρεται από το DNA στα ριβοσώματα μέσω του mRNA. Κάθε φορά, δηλαδή, που απαιτείται η σύνθεση μιας πρωτεΐνης, το τμήμα του DNA που φέρει την πληροφορία για τη σύνθεσή της αρχικά **μεταγράφεται** σε mRNA.

Το mRNA που συντίθεται πρέπει να περιέχει τη γενετική πληροφορία για τη σειρά των αμινοξέων της συγκεκριμένης πρωτεΐνης. Για να γίνει αυτό, το συγκεκριμένο τμήμα του DNA ξετυλίγεται και η μία αλυσίδα απομακρύνεται από την άλλη. Απέναντι από τις αζεωργωμένες πλέον αζωτούχες βάσεις των δεσορυβονουκλεοτιδίων της μιας αλυσίδας τοποθετούνται, κατά μία συγκεκριμένη φορά, ελεύθερα ριβονουκλεοτίδια που διαθέτουν τις συμπληρωματικές αζωτούχες βάσεις. Ετσι, απέναντι από την αδενίνη, τη θυμίνη, τη γουανίνη και την κυτοσίνη του DNA τοποθετούνται ελεύθερα ριβονουκλεοτίδια που φέρουν αντίστοιχα ουρακίλη, αδενίνη, κυτοσίνη και γουανίνη. Στη συνέχεια, τα ριβονουκλεοτίδια ενώνονται μεταξύ τους, σχηματίζοντας ένα μόριο mRNA στο οποίο έχει πλέον καταγραφεί η γενετική πληροφορία ως αλληλουχία ριβονουκλεοτιδίων. Το μόριο αυτό απομακρύνεται και οι συμπληρωματικές αζωτούχες βάσεις των δύο αλυσίδων του DNA ενώνονται και πάλι.

Με τη μεταγραφή όμως δεν παράγεται μόνο mRNA. Διάφορα τμήματα του DNA μεταγράφονται με τον ίδιο τρόπο, για να συντεθούν άλλα μόρια RNA, το βασικότερο

Εικ. 5.10 Η αντιγραφή του DNA.

Εικ. 5.11 Η διαδικασία της μεταγραφής του DNA.

100 ΒΙΟΛΟΓΙΑ Β' & Γ' ΓΥΜΝΑΣΙΟΥ



AI-powered chatbot for digital archives

The screenshot displays the NT Archive website with a chatbot overlay. The website header includes the NT logo and navigation links: ΠΑΡΑΣΤΑΣΕΙΣ, ΕΡΓΑ, ΠΡΟΣΩΠΑ, ΡΟΛΟΙ, ΣΚΗΝΕΣ, ΘΕΑΤΡΑ, ΣΥΛΛΟΓΕΣ. The chatbot, named Melina, introduces herself as a theater lover and assistant, offering information on plays and reminding users to log in for quality and education. It provides example questions like requesting photos of a specific play or listing plays an actor has performed in. Below the chatbot, a section titled 'Παραστάσεις' (Performances) includes a timeline from 1932 to 1944 and a search bar.

ΨΗΦΙΑΚΟ ΑΡΧΕΙΟ

ΠΑΡΑΣΤΑΣΕΙΣ ΕΡΓΑ ΠΡΟΣΩΠΑ ΡΟΛΟΙ ΣΚΗΝΕΣ ΘΕΑΤΡΑ ΣΥΛΛΟΓΕΣ ▾

ΣΥΝΘΕΤΗ ΑΝΑΖΗΤΗΣΗ

ΕΛ ▾

Παραστάσεις

Καλώς ήρθατε στην κατηγορία Παραστάσεις του ψηφιακού αρχείου του Εθνικού Θεάτρου. Χρησιμοποιήστε την αλφαβητική ή τη χρονολογική πλοήγηση για να βρείτε την παράσταση που σας ενδιαφέρει.

Χρονολογικά

1932 1934 1936 1938 1940 1942 1944

Type your message...

Είμαι η Μελίνα, λάτρης του θεάτρου και βοηθός στο αρχείο του Εθνικού Θεάτρου. Είμαι εδώ για να σας παρέχω πληροφορίες για παλαιότερες παραστάσεις. Παρακαλώ σημειώστε ότι όλες οι συνομιλίες καταγράφονται για λόγους ποιότητας και εκπαίδευσης. Μην μοιράζεστε προσωπικά στοιχεία. Πώς μπορώ να σας εξυπηρετήσω;

Παραδείγματα ερωτήσεων:

- Θέλω να δω φωτογραφίες από την παράσταση 'Ξύπνα Βασίλη'
- Σε ποιες παραστάσεις έχει παίξει η Παξινού

<https://www.nt-archive.gr/>

Connect 2025 – Athens, June 5, 2025




AI-powered chatbot for digital archives

ΑΠΟΤΕΛΕΣΜΑΤΑ: 20


Έτος: 2022 X

9 v


ΗΜΕΡΟΜΗΝΙΑ v




Ο εχθρός της τάξης (2021)
Επανάληψη
Σκηνοθεσία: Γεωργία Μαυραγάνη




Φουέντε Οβεχούνα (2021)
Επανάληψη
Σκηνή: Κεντρική Σκηνή
Σκηνοθεσία: Ελένη Ευθυμίου




Λεωφορείον ο Πόθος (2021)
Επανάληψη
Σκηνή: Νέα Σκηνή
Σκηνοθεσία: Θανάσης Σαράντος



Η φάρμα των ζώων (2021)



Αιολική γη (2021)



Ματωμένος γάμος (2022)

Ποιος ηθοποιός έχει τους περισσότερους πρωταγωνιστικούς ρόλους στο αρχείο;

Ο ηθοποιός με τους περισσότερους πρωταγωνιστικούς ρόλους στο αρχείο είναι ο Αλέξης Μινωτής, με 26 ρόλους.

Θέλω να δω φωτογραφίες από παραστάσεις μετά το 1970 στις οποίες έχει παίξει ο Μινωτής

Βρήκα τις παρακάτω παραστάσεις μετά το 1970 στις οποίες έχει παίξει ο Μινωτής:

- Οιδίπους επί Κολωνών (1989) [link](#)
- Η θυσία του Αβραάμ (1989-1990) [link](#)
- Φοίνισσες (1988) [link](#)
- Ο πατέρας (1987-1989) [link](#)
- Το τέλος του παινιδιού - Πράξη χωρίς λόγια (1986-1987) [link](#)
- Ο καρδινάλιος της Ισπανίας (1981-1982) [link](#)
- Ευτυχισμένες μέρες - Το τέλος του παινιδιού (1979-1987) [link](#)
- Προμηθεύς δεσμώτης (1979-1981) [link](#)
- Βούδας (1978) [link](#)
- Φοίνισσες (1978-1981) [link](#)

Type your message...

<https://www.nt-archive.gr/>



SciNoBo: AI assistant for science discovery and analysis



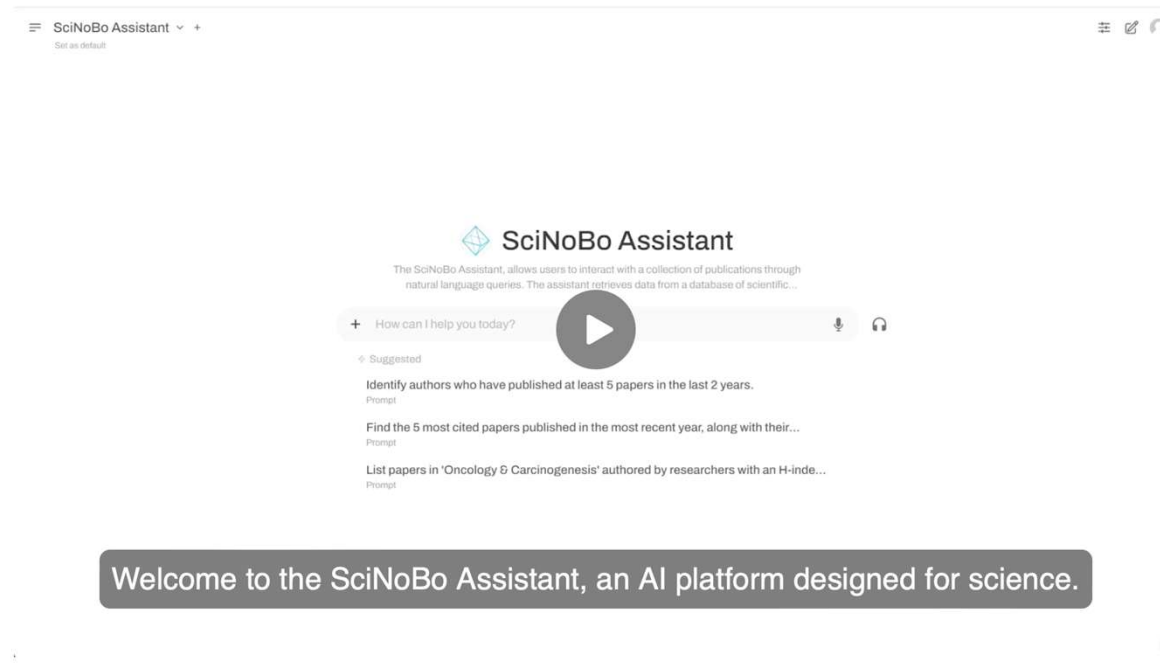
- Home
- About
- Toolkit
- Research
- Documentation
- Contact

Community



<https://scinobo.ilsp.gr>


Demonstration of the SciNoBo Assistant:





Connect 2025 – Athens, June 5, 2025



Talking to a video

 Τι κινδύνους μπορεί να ενέχει η τεχνητή νοημοσύνη;

 **Απάντηση**
Η τεχνητή νοημοσύνη μπορεί να ενέχει κινδύνους όταν δεν ελέγχεται σωστά ή όταν χρησιμοποιείται σε λάθος πλαίσια. Μπορεί να γίνει επικίνδυνη αν δεν υπάρχουν κατάλληλες δικλείδες ασφαλείας, καθώς μπορεί να οδηγήσει σε ανεπιθύμητες συνέπειες. Επιπλέον, η χρήση της σε διάφορους τομείς χωρίς επαρκή κατανόηση των δυνατοτήτων και των περιορισμών της μπορεί να προκαλέσει προβλήματα.

 **Δείτε σχετικά βίντεο** ⇄


Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:32

Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:21

Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:35


Δείτε σχετικές διαφάνειες

Πλοήγηση στη διαφάνεια 1.2.1-Levels-of-Intelligence.pdf: 2

Τι θέλετε να μάθετε; 



Talking to a video

 **Δείτε σχετικά βίντεο**


Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:32


Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:21

Πλοήγηση στο βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:35

Δείτε σχετικές διαφάνειες


Πλοήγηση στη διαφάνεια 1.2.1-Levels-of-Intelligence.pdf: 2

 Λήψη του βίντεο 1.4.1_Failure-of-AI.mp4: 00:02:32.309000



Ethics of AI

- **AI “failures”**
Microsoft AI Chatbot “TAY” (2016) was attacked by Internet trolls and in less than 24 hours the chatbots “personality” was corrupted





Krikri using recent knowledge from news sources

Check recent news with Krikri and other LLMs

Enter your question:

Τι συμβαίνει με το Πακιστάν και την Ινδία;

Select an LLM:

krikri-dpo-latest

Get Answer

Answer from krikri-dpo-latest:

Η Ινδία και το Πακιστάν βρίσκονται σε σοβαρή κλιμάκωση των σχέσεών τους με αφορμή μια σειρά γεγονότων που αφορούν κυρίως τη διαφιλονικούμενη περιοχή του Κασμίρ και τη διαχείριση των υδάτινων πόρων.

Συγκεκριμένα:

- Μετά από μια αιματηρή επίθεση σε τουριστικό σημείο στο ινδικό Κασμίρ με 26 νεκρούς (βλ. <https://www.in.gr/2025/04/28/world/kasmir-pos-nero-ginetai-pio-isxyro-oplo-sti-nea-entasi-metaksy-dyo-pyrinikon-dynameon>),
- η Ινδία κατηγορήσε το Πακιστάν για υποστήριξη τρομοκρατικών ενεργειών και προχώρησε σε:
 - Ανάκληση διπλωματών
 - Ακύρωση έκδοσης βίζας
 - Πάγωμα εμπορικών συναλλαγών
 - Κλείσιμο χερσαίων συνόρων (με ανταλλαγή πυρών σε κάποια σημεία)
 - Αναστολή της "Συνθήκης Υδάτων του Ινδού" (IWT), μονομερώς, προκαλώντας διακοπή στη διαχείριση των υδάτινων πόρων του ποταμού Ινδού, που είναι ζωτικής σημασίας για το Πακιστάν (80% των υδάτων του για άρδευση και υδροηλεκτρική ενέργεια - βλ.

Connect 2025 – Athens, June 5, 2025



The team



Leon Voukoutis



George Paraskevopoulos



Prokopis Prokopidis



Dimitris Roussis



Sokratis Sofianopoulos



Vassilis Katsouros

Connect 2025 – Athens, June 5, 2025



Acknowledgements



