



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

— ΙΔΡΥΘΕΝ ΤΟ 1837 —

Εισαγωγικές Διαλέξεις Βασικού Εργαστηρίου Φυσικής Ι - LLMs

Τι είναι ένα LLM;



LLM = Large Language Model = Μεγάλο Γλωσσικό Μοντέλο

- Μαθαίνει μοτίβα από πολλά κείμενα.
- Παράγει κείμενο, εξηγήσεις, περίληψη, κώδικα.
- Απαντά με βάση το πλαίσιο της ερώτησης.
- Δεν σκέφτεται και δεν παρατηρεί όπως ένας άνθρωπος.

Παραδείγματα εφαρμογών / μοντέλων

ChatGPT

Claude

Gemini

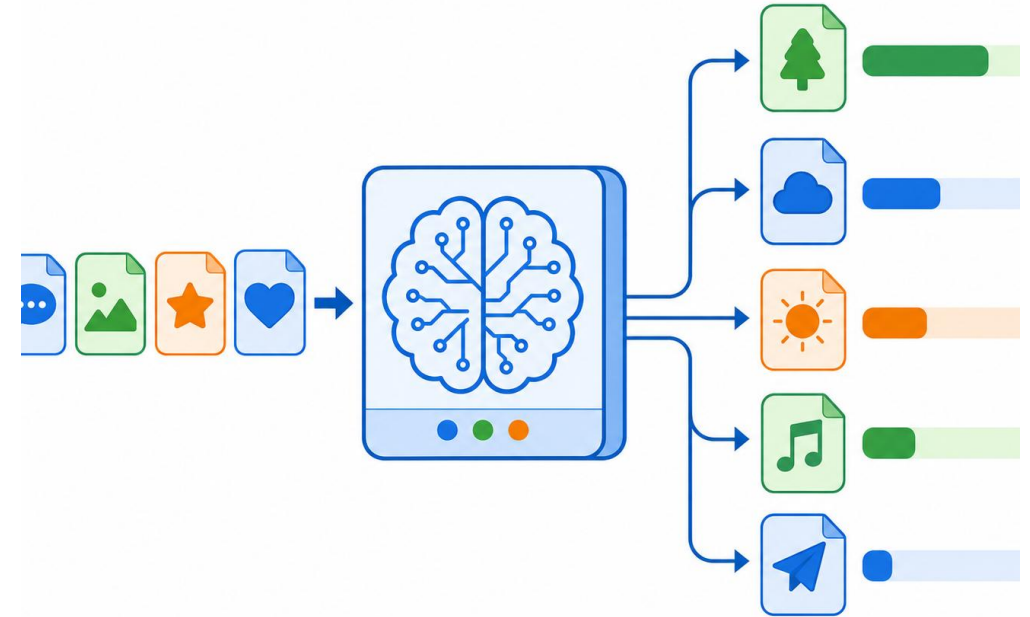
Copilot

Llama

**Mistral /
Le Chat**

Πώς δημιουργείται μία απάντηση;

- 1 Γράφουμε μια ερώτηση.
- 2 Το κείμενο χωρίζεται σε μικρά κομμάτια.
- 3 Το μοντέλο βλέπει το πλαίσιο.
- 4 Υπολογίζει ποια συνέχεια ταιριάζει καλύτερα.
- 5 Προσθέτει το επόμενο κομμάτι και συνεχίζει.



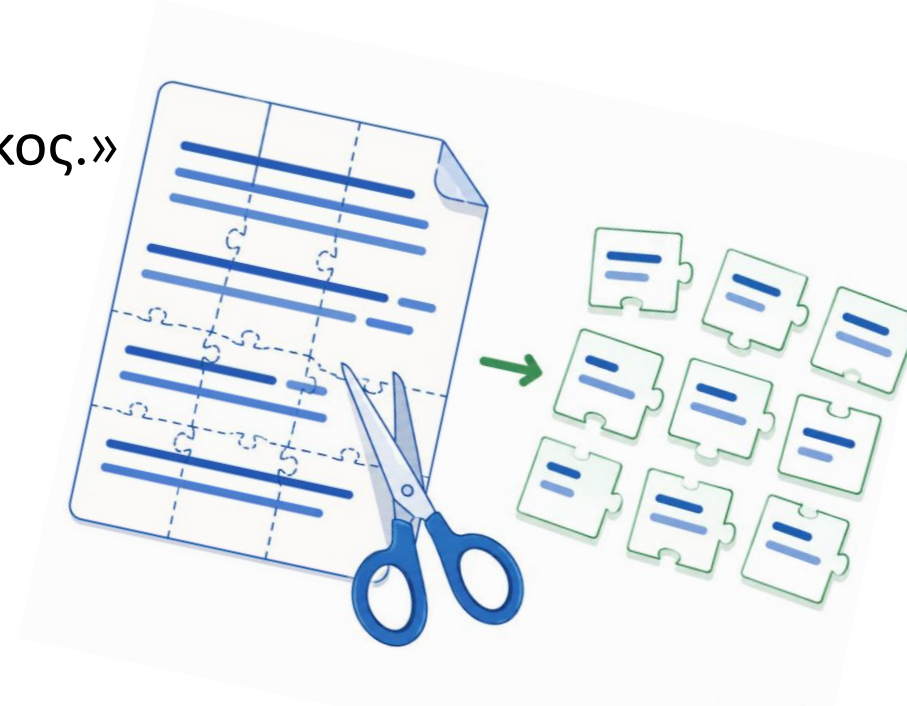
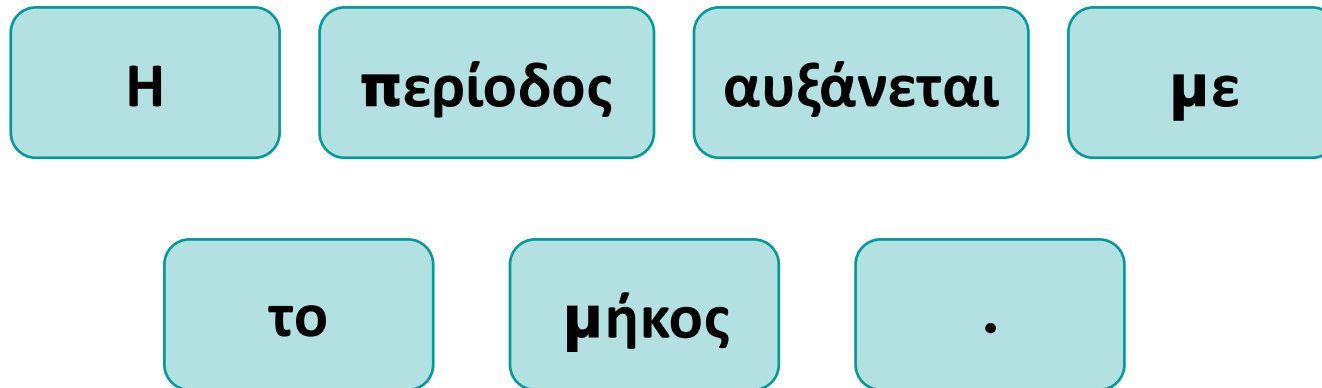
Ενα LLM προβλέπει την πιθανότερη συνέχεια του κειμένου, με βάση το πλαίσιο και όσα έχει μάθει.

Tokens: το κείμενο κόβεται σε μικρά κομμάτια

Δεν δουλεύει πάντα με ολόκληρες λέξεις όπως τις βλέπουμε εμείς.

Παράδειγμα

«Η περίοδος αυξάνεται με το μήκος.»



Το token μπορεί να είναι λέξη, μέρος λέξης, αριθμός ή σημείο στίξης.
Ο ακριβής χωρισμός εξαρτάται από το μοντέλο.

Το πλαίσιο και η σειρά αλλάζουν το νόημα

Το πλαίσιο της «συζήτησης» και η σειρά των λέξεων αλλάζουν το νόημα.
Το μοντέλο δεν βλέπει απλώς έναν «σύνολο λέξεων».

Ίδιες σχεδόν λέξεις - άλλο νόημα

- Το σώμα Α έχει υψηλότερη θερμοκρασία από το σώμα Β.
- Το σώμα Β έχει υψηλότερη θερμοκρασία από το σώμα Α.

Ίδια λέξη - άλλο πλαίσιο

- Περίοδος του εκκρεμούς.
- Περίοδος εξετάσεων.



Γι' αυτό το γλωσσικό μοντέλο χρειάζεται ένα μηχανισμό που να λαμβάνει υπόψη τη θέση, τη σειρά και τις γειτονικές πληροφορίες.

Από λέξεις σε αριθμητικά μοτίβα



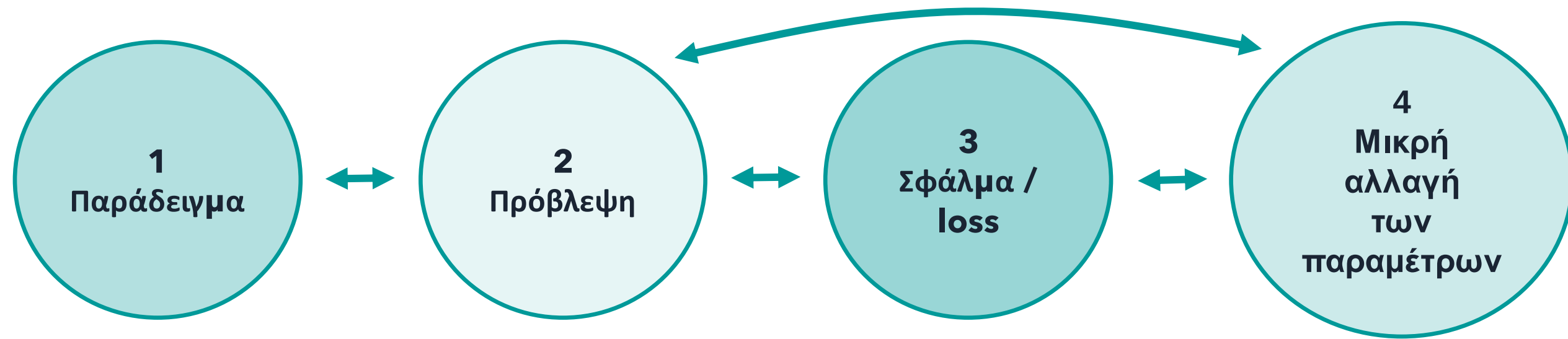
Το μοντέλο δεν κρατά λεξικό εννοιών - δουλεύει με αριθμητικές αναπαραστάσεις.



Όπως οι συντεταγμένες ενός σημείου δείχνουν τη θέση του σε έναν χάρτη, έτσι και οι αριθμοί ενός token δείχνουν τη θέση του σε έναν χώρο σχέσεων που έχει μάθει το μοντέλο.

Training: πώς μαθαίνει ένα νευρωνικό δίκτυο;

Training σημαίνει προσαρμογή των παραμέτρων, όχι απλή αποθήκευση παραδειγμάτων.



Παράδειγμα

«Η Γη περιστρέφεται γύρω από τον ...».

Αν το σωστό επόμενο token είναι «Ήλιο», το μοντέλο συγκρίνει την πρόβλεψή του με αυτό και διορθώνει ελαφρά τις παραμέτρους.

Το επόμενο κομμάτι επιλέγεται από υποψήφίες συνέχειες

Οι τιμές εδώ είναι υποθετικές και δείχνουν μόνο την ιδέα.



Ερώτηση

«Η μονάδα της αντίστασης είναι το ...»

Σχετική πιθανότητα σύμφωνα με το μοντέλο

ohm		70%
volt		15%
ampere		10%
joule		5%



Πιθανότερο δεν σημαίνει «αποδεδειγμένα αληθές». Σημαίνει ότι ταιριάζει καλύτερα ως συνέχεια στο συγκεκριμένο πλαίσιο.

Γιατί η ίδια ερώτηση μπορεί να δώσει διαφορετική απάντηση;

Η απάντηση δεν είναι πάντα απόλυτα ντετερμινιστική και το πλαίσιο αλλάζει εύκολα.

Διαφορετική διατύπωση

Οι λέξεις ή η σειρά
παρουσίασης, χωρίς να
αλλάζει απαραίτητα το
νόημα.

Διαφορετικό πλαίσιο

Το περιεχόμενο της
απάντησης, ανάλογα με
όσα έχουν προηγηθεί
στη συζήτηση.

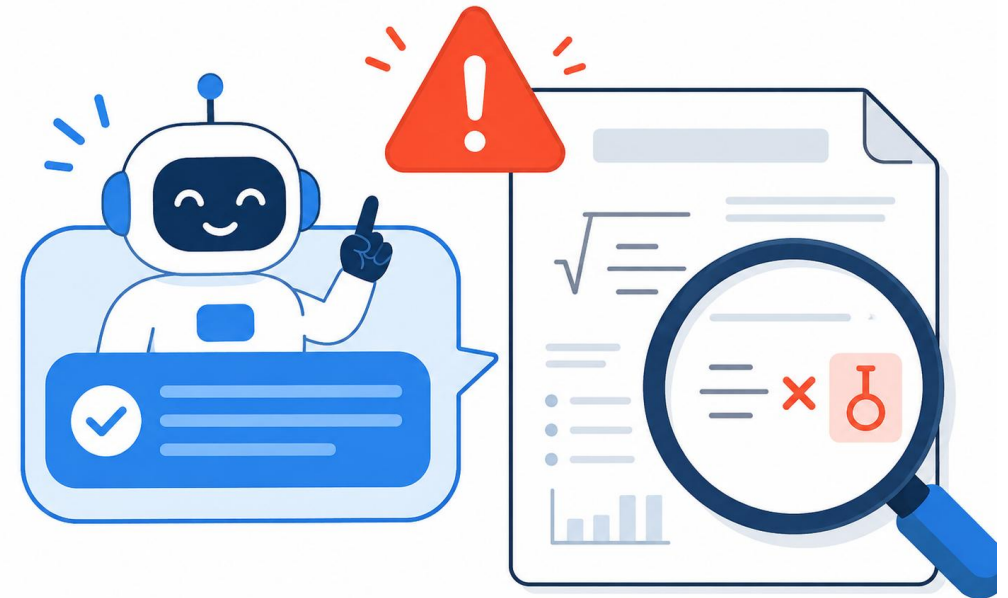
Στοχαστική παραγωγή

Η επιλογή της
επόμενης λέξης ή
φράσης, ακόμη και με
την ίδια ακριβώς
είσοδο.

Γιατί κάνει λάθη;

Η καλή γλώσσα δεν είναι απόδειξη επιστημονικής ορθότητας.

- Δεν καταλαβαίνει όπως άνθρωπος.
- Μπορεί να μπερδέψει τύπους ή μονάδες.
- Μπορεί να δώσει παρωχημένη πληροφορία.
- Μπορεί να εφεύρει πηγές ή λεπτομέρειες.



Σε μελέτη με φοιτητές φυσικής, γλωσσικά πειστικές απαντήσεις ChatGPT δυσκόλεψαν την αναγνώριση επιστημονικών λαθών, ιδίως όταν το θέμα δεν ήταν καλά γνωστό.

Παράδειγμα: πειστικό αλλά λάθος στη φυσική



Ερώτηση

Δύο σώματα έχουν ίδια κινητική ενέργεια. Ποιο έχει μεγαλύτερη ορμή;

Πειστική λανθασμένη απάντηση

«Αφού έχουν ίδια κινητική ενέργεια, έχουν και ίδια ορμή.»

Έλεγχος με τύπο

$$K = p^2 / 2m$$

Άρα για ίδιο K :

$$p \propto \sqrt{m}$$



Η σωστή πρακτική: βρίσκω τον κατάλληλο τύπο, ελέγχω μονάδες και συμπέρασμα, όχι απλώς αν η απάντηση ακούγεται ωραία.

Τι μπορεί να κάνει καλά - και πού θέλει προσοχή

Χρήσιμος βοηθός, όχι αλάνθαστος κριτής.

Μπορεί να βοηθήσει

- ✓ να εξηγήσει μια έννοια
- ✓ να προτείνει βήματα
- ✓ να ελέγξει μονάδες
- ✓ να σχολιάσει κείμενο



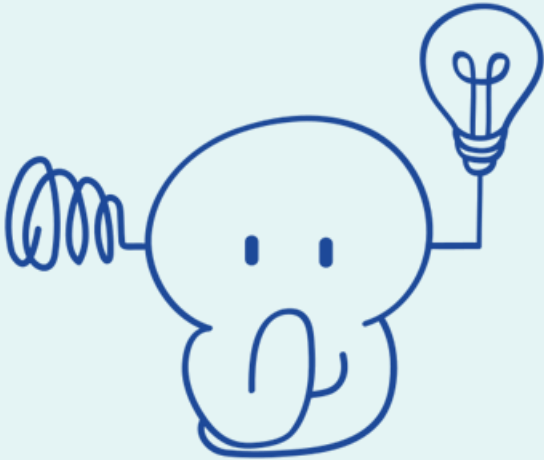
Θέλει προσοχή

- ! μπορεί να κάνει αριθμητικά λάθη
- ! να μπερδέψει προϋποθέσεις
- ! να επινοήσει πηγές
- ! δεν είδε το πείραμά σου

  Κάθε απάντηση πρέπει να περνά από επιστημονικό έλεγχο πριν χρησιμοποιηθεί σε εργασία.

Πριν από το εργαστήριο

Χρήση για προετοιμασία, όχι για έτοιμη αναφορά.



- Εξήγηση της φυσικής αρχής.
- Ποια μεγέθη θα μετρηθούν;
- Τι γράφημα περιμένουμε;
- Ερωτήσεις αυτοελέγχου.
- Προϋποθέσεις τύπων.

Καλή εντολή:

«Προετοιμάζομαι για το απλό εκκρεμές. Εξήγησέ μου τι θα μετρήσουμε και γιατί.
Στο τέλος κάνε μου ερωτήσεις.»

Μετά το εργαστήριο

Το LLM βοηθά στον έλεγχο και στη βελτίωση της δικής σου εργασίας.



1. Δική μου πρώτη επεξεργασία
2. Έλεγχος με LLM
3. Ανεξάρτητη επαλήθευση
4. Δική μου τελική διατύπωση

Ζητάμε έλεγχο υπολογισμών, μονάδων, γραφήματος, αβεβαιοτήτων και σαφήνειας - όχι να γραφτεί η αναφορά από την αρχή.

Στοιχεία για μία καλή εντολή

Μια καλή εντολή μοιάζει με καλή εκφώνηση εργαστηριακής άσκησης.

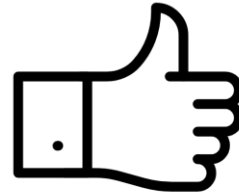
- 1 Πλαίσιο:** Σε ποιο μάθημα / άσκηση είμαι;
- 2 Στόχος:** Τι ακριβώς θέλω;
- 3 Δεδομένα:** Ποια στοιχεία δίνω;
- 4 Περιορισμοί:** Τι δεν θέλω να κάνει;
- 5 Μορφή:** Πώς θέλω την απάντηση;
- 6 Έλεγχος:** Τι πρέπει να ελεγχθεί;

Εντολές για να μάθω θεωρία

Στόχος: να καταλάβω, όχι να απομνημονεύσω κείμενο.



Εξήγησε την ενέργεια.



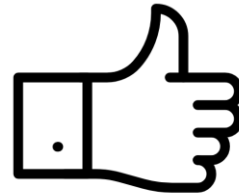
Εξήγησέ μου τη διαφορά κινητικής και δυναμικής ενέργειας σε επίπεδο πρωτοετούς. Δώσε ένα καθημερινό παράδειγμα και ένα εργαστηριακό παράδειγμα.

Εντολές για εργαστηριακή αναφορά

Το LLM σχολιάζει το δικό σου κείμενο - δεν γράφει την αναφορά αντί για εσένα.



Γράψε μου τη συζήτηση αποτελεσμάτων.



Παρακάτω είναι το δικό μου προσχέδιο. Μην το ξαναγράψεις. Εντόπισε ισχυρισμούς που δεν υποστηρίζονται από τα δεδομένα, ασάφειες, λάθος μονάδες και υπερβολικά απόλυτες διατυπώσεις.

Ζητάμε σχόλια ως ερωτήσεις: «Ποια πρόταση χρειάζεται καλύτερη τεκμηρίωση;»

Εντολές για έλεγχο δικής μου λύσης

Καλή εντολή:

«Κάνε μου προφορική εξέταση στις δυνάμεις. Μία ερώτηση κάθε φορά. Μετά από κάθε απάντηση πες τι είναι σωστό, τι λείπει και κάνε μία επόμενη ερώτηση.»

Γιατί δουλεύει;

Σε αναγκάζει να σκεφτείς πριν δεις την απάντηση.
Μην ζητάς μόνο περίληψη. Ζήτησε ερωτήσεις, κενά γνώσης και ανατροφοδότηση.

Εντολές για αυτοεξέταση και ενεργητική ανάκληση

Να ελέγχει τη σκέψη μας χωρίς να την αντικαθιστά.

Πρότυπο εντολής

«Ακολουθεί τη δική λύση μου. Μην την αντικαταστήσεις με άλλη. Βρες το πρώτο σημείο όπου υπάρχει λάθος ή ασάφεια. Εξήγησε γιατί και περίμενε να το διορθώσω πριν συνεχίσεις.»



Πριν παραδώσουμε την εργασία πρέπει να μπορούμε να εξηγήσουμε κάθε βήμα χωρίς το LLM.

Εντολές για πηγές και βιβλιογραφία

Μια παραπομπή δεν χρησιμοποιείται αν δεν την ανοίξουμε και δεν την ελέγχουμε.

Καλή εντολή:

«Βρες επίσημες ή ακαδημαϊκές πηγές. Για κάθε ισχυρισμό δείξε ποια πηγή τον υποστηρίζει. Μην κατασκευάσεις παραπομπές. Δήλωσε τι δεν μπόρεσες να επαληθεύσεις.»

Έλεγχος από εμάς



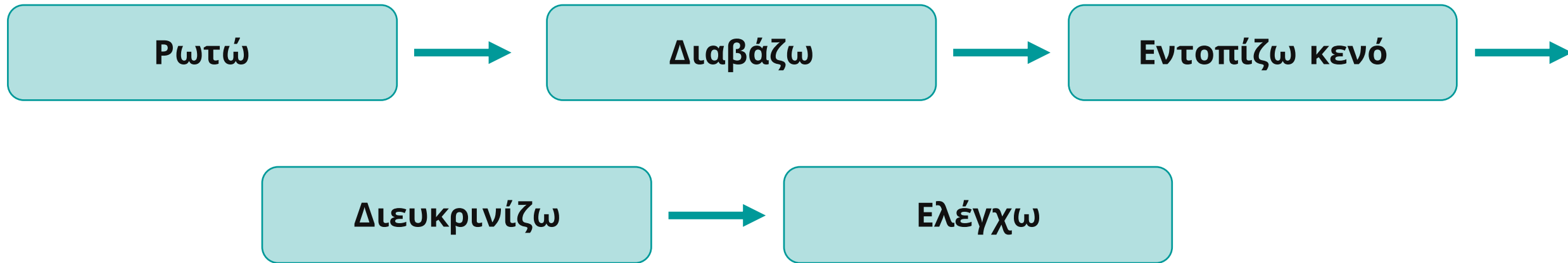
- Υπάρχει η πηγή;
- Ταιριάζει ο τίτλος;
- Υπάρχει DOI / επίσημη σελίδα;
- Υποστηρίζει πράγματι τον ισχυρισμό;



Χωρίς αναζήτηση, ένα LLM μπορεί να μην έχει πρόσφατη πληροφορία.

Η καλή εντολή είναι συχνά διάλογος

Δεν χρειάζεται να πετύχουμε την τέλεια ερώτηση με την πρώτη.



Παράδειγμα

1. «Εξήγησέ το απλά.» → 2. «Δεν κατάλαβα το βήμα 2.» → 3. «Δώσε αντιπαράδειγμα.» → 4. «Κάνε μου μία ερώτηση ελέγχου.» → 5. «Έλεγχξε τη δική μου εξήγηση.»

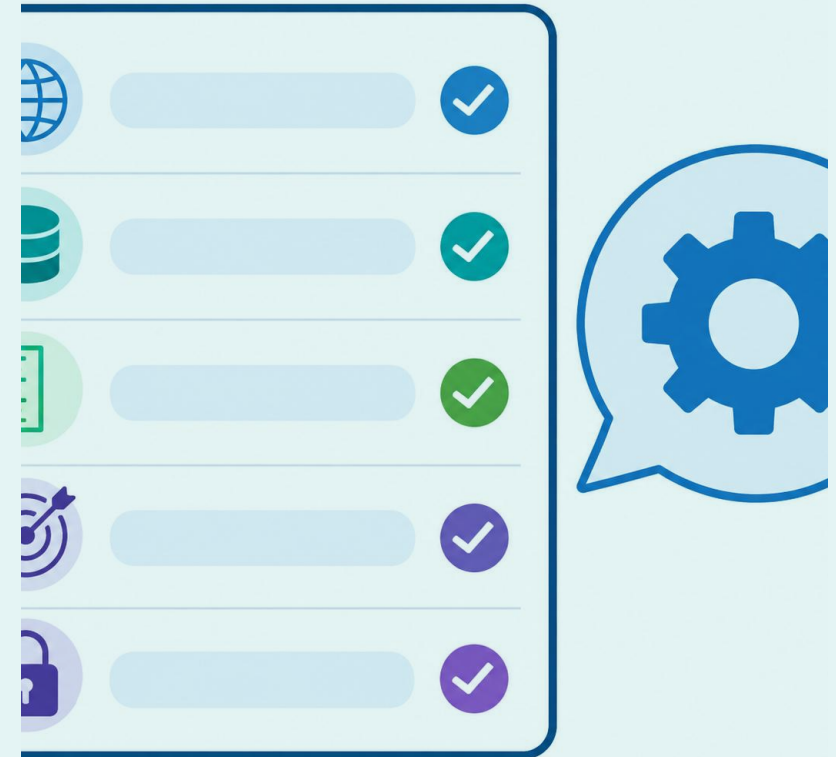


Καλύτερο αποτέλεσμα δεν σημαίνει μόνο καλύτερη απάντηση.
Σημαίνει καλύτερη κατανόηση μετά τη συνομιλία.

Πρωτόκολλο ελέγχου απάντησης

Πριν χρησιμοποιήσουμε αποτέλεσμα LLM σε εργασία, κάνουμε αυτόν τον έλεγχο.

- ☐ Οι αριθμοί προέρχονται από τα δικά μου δεδομένα;
- ☐ Οι μονάδες και οι μετατροπές είναι σωστές;
- ☐ Ο τύπος είναι κατάλληλος για τις προϋποθέσεις;
- ☐ Η τάξη μεγέθους είναι λογική;
- ☐ Η γραφική παράσταση έχει σωστούς άξονες;
- ☐ Οι πηγές υπάρχουν και είναι αξιόπιστες;
- ☐ Μπορώ να εξηγήσω τα βήματα μόνος/μόνη μου;



«Κάθε αποτέλεσμα που παρουσιάζεται πρέπει να μπορεί να εξηγηθεί και να το αιτιολογηθεί.»

Ακαδημαϊκή ακεραιότητα και δεδομένα

Πάντα υπερισχύουν οι οδηγίες του διδάσκοντα και της συγκεκριμένης εργασίας.

Δεν επιτρέπεται



- Επινόηση ή αλλαγή δεδομένων
- Ολόκληρη εργασία χωρίς κατανόηση
- Ανύπαρκτες πηγές
- Προσωπικά / εμπιστευτικά δεδομένα

Πρέπει να δηλώνεται με σαφήνεια πώς αξιοποιήθηκε η Τεχνητή Νοημοσύνη.
Η ευθύνη για την ορθότητα και την τεκμηρίωση της εργασίας παραμένει δική μας.

Τα LLMs είναι ισχυρά εργαλεία. Η επιστημονική κρίση, ο έλεγχος και η ακαδημαϊκή ευθύνη παραμένουν στον άνθρωπο.

Βιβλιογραφία

1. Vaswani et al. (2017), "Attention Is All You Need", NeurIPS.
2. OpenAI, "How ChatGPT and our foundation models are developed".
3. Dahlkemper, Lahme & Klein (2023), Phys. Rev. Phys. Educ. Res. 19, 010142.
4. Polverini & Gregorcic (2024), European Journal of Physics 45, 025701.
5. UNESCO (2023), "Guidance for Generative AI in Education and Research".
6. OpenAI, "Prompt engineering best practices for ChatGPT".