

- Από τους αλγορίθμους στον ασθενή.

*Τεχνητή νοημοσύνη και ο μετασχηματισμός της υγείας.*

**Ηλίας Μόσιαλος**

*Cheng Yu Tung Chair in Global Health*

*London School of Economics and Political Science*

- # Η ΑΙ δεν είναι μαγεία.

## *Είναι μια αρχιτεκτονική.*

Η ΑΙ χτίζεται σε επίπεδα (layers). Κάθε ένα στέκεται πάνω στο προηγούμενο.



*Βασισμένο σε post του Luís Rodrigues*

# Τα 6 επίπεδα της ΑΙ αρχιτεκτονικής

Από τους κανόνες στους αυτόνομους πράκτορες (agents)

6	Πρακτορική ΑΙ	( <i>Agentic AI</i> )	Προηγμένα συστήματα ΤΝ
5	Παραγωγική ΑΙ	( <i>Generative AI</i> )	Εστιάζει στη δημιουργία νέου περιεχομένου (κείμενο, εικόνα) βάσει προτύπων που έμαθε από τεράστια δεδομένα
4	Βαθιά Μάθηση	( <i>Deep Learning</i> )	Νευρωνικά δίκτυα σε κλίμακα.
3	Νευρωνικά Δίκτυα	( <i>Neural Networks</i> )	Εμπνευσμένα από τη δομή του ανθρώπινου εγκεφάλου.
2	Μηχανική Μάθηση	( <i>Machine Learning</i> )	Μαθαίνει από δεδομένα, όχι κανόνες.
1	Κλασική ΑΙ	( <i>Classical AI</i> )	Εστιάζει στην ανάλυση, την ταξινόμηση και τη λήψη αποφάσεων βάσει κανόνων (π.χ. «είναι spam το email;»).

# Τι μπορεί και τι όχι η ΑΙ σήμερα

Πριν μπούμε στις λεπτομέρειες, μια ισορροπημένη εικόνα.

## ΔΥΝΑΤΟΤΗΤΕΣ

### Τι κάνει καλά

- Αναγνώριση προτύπων σε εικόνες, σήμα, κείμενο σε κλίμακα που υπερβαίνει τον άνθρωπο.
- Σύνοψη μεγάλου όγκου πληροφορίας, κειμένων, αυτόματη τεκμηρίωση.
- Πρόβλεψη κινδύνου από ιστορικά δεδομένα (AKI, CKD, καρδιαγγειακά).
- Αυτοματοποίηση ροών εργασίας: Επισκόπηση μητρώων, προέγκριση, παρακολούθηση μετά τη θεραπεία
- Διαθεσιμότητα 24/7, χωρίς κόπωση, σταθερή απόδοση.

## ΠΕΡΙΟΡΙΣΜΟΙ

### Τι ξεχνάμε εύκολα

- Επινοεί πληροφορίες που ακούγονται αληθοφανείς (hallucinations).
- Αναπαράγει και ενίοτε ενισχύει υπάρχουσες προκαταλήψεις στα δεδομένα εκπαίδευσης.
- Δεν "καταλαβαίνει". Παράγει στατιστικά πιθανές απαντήσεις.
- Δεν ξέρει τι δεν ξέρει: σπάνια εκφράζει αβεβαιότητα όταν θα έπρεπε.
- Λίγα τεκμήρια ότι βελτιώνει τελικά την έκβαση ασθενών.

# Παραισθήσεις (hallucinations)

Όταν η AI παράγει αληθοφανείς αλλά λανθασμένες πληροφορίες.

## ΟΡΙΣΜΟΣ

Ένα LLM δεν "ψάχνει την αλήθεια". Παράγει την επόμενη πιο πιθανή λέξη με βάση τα δεδομένα εκπαίδευσης. Όταν δεν έχει σωστή πληροφορία, παράγει εύλογη αλλά λανθασμένη.

### ΦΑΡΜΑΚΕΥΤΙΚΗ ΑΓΩΓΗ

#### Επινοημένη δοσολογία

LLM προτείνει δόση αντιβιοτικού που μοιάζει σωστή αλλά δεν αντιστοιχεί σε καμία κατευθυντήρια οδηγία ή καμία ένδειξη.

### ΒΙΒΛΙΟΓΡΑΦΙΑ

#### Ανύπαρκτες παραπομπές

Παρέχει τίτλο μελέτης, συγγραφείς, περιοδικό, DOI, που μοιάζουν αληθοφανείς αλλά η μελέτη δεν υπάρχει.

### ΚΛΙΝΙΚΗ ΑΠΟΦΑΣΗ

#### Λανθασμένη ταξινόμηση

Σε ασθενή με eGFR 45, χαρακτηρίζει το στάδιο ως G2 αντί G3a. Μικρό αλλά κρίσιμο λάθος.

## ΑΜΥΝΑ

- Πάντα διασταύρωση από κλινικό.
- RAG (retrieval-augmented generation): η AI ψάχνει σε αξιόπιστη πηγή πριν απαντήσει.
- Citation grounding: κάθε ισχυρισμός να παραπέμπει σε επιβεβαιώσιμη πηγή.

# THE LANCET

Volume 407 · Number 10541 · Pages 1257-1890 · May 9-15, 2026

www.thelancet.com

The Lancet, Μάιος 2026

Έλεγχος 2,5 εκατομμυρίων βιοϊατρικών δημοσιεύσεων σε τρία χρόνια έδειξε ότι επινοημένες παραπομπές είναι ενσωματωμένες στη βιβλιογραφία σε μεγάλη κλίμακα και ότι ο ρυθμός επινόησης τους αυξάνεται.

“[An]...audit of 2.5 million biomedical papers spanning 3 years, show[ed] that fabricated references are embedded in the peer-reviewed literature at scale, and that the rate of fabrication is accelerating.”

See *Correspondence* page 1279

## Comment

RTS,S/AS01 reduces mortality in African children  
See page 1263

## Comment

Global challenges for research on gun violence  
See page 1266

## Comment

Offline: Climate and health—time to step up our activism  
See page 1269

## Perspectives

When doctors do not care  
See page 1276

## Articles

Empatoran, a Toll-like receptor 2/8 inhibitor, in systemic lupus erythematosus: findings from phase 2 trial  
See page 1309

# Πόσο συχνά παραλείπει ή επινοεί η AI κλινικές οδηγίες;

Μεγάλης κλίμακας αξιολόγηση: 12.197 διαγνωστικά outputs LLMs.

## ΤΙ ΕΙΝΑΙ ΤΑ 12.197

Είναι απαντήσεις (outputs), όχι ξεχωριστά ερωτήματα. 54 case vignettes (2 νοσήματα × 27 παραλλαγές δημογραφικών) δόθηκαν σε 2 LLMs, και κάθε συνδυασμός εκτελέστηκε 128 φορές για να αποτυπωθεί η στοχαστικότητα.

**54 vignettes × 2 LLMs × 128 runs = 13.824 outputs → 12.197.**

**22–46%**

### OMISSIONS · GPT-4.1

παραλείπει τις κλινικές οδηγίες παρότι του ζητήθηκαν ρητά

**97%**

### OMISSIONS · DeepSeek-V3

παραλείπει τις κλινικές οδηγίες σχεδόν πάντα

**7–9%**

### HALLUCINATIONS · GPT-4.1

επινοεί ανύπαρκτες κατευθυντήριες οδηγίες ως αληθινές

**19.4%**

### UK T2DM · GPT-4.1

ποσοστό hallucination σε ΣΔ τύπου 2 με τοποθεσία UK

# Παραδείγματα στη φροντίδα νεφρών

Στατιστικά μοντέλα μαθαίνουν από ιστορικά δεδομένα ασθενών για να προβλέψουν εκβάσεις.

## Πρόβλεψη AKI 48 ωρών

Προβλέπει την πιθανότητα οξείας νεφρικής βλάβης τις επόμενες 48 ώρες.

## Πρόβλεψη πορείας CKD

Εξετάζοντας διαχρονικά eGFR/UACR εκτιμά τον κίνδυνο 5ετούς εξέλιξης σε ESKD (KFRE risk equation).

## Πρόβλεψη υποτασικού επεισοδίου

Προβλέπει το ρίσκο υποτασικού επεισοδίου, βάσει βάρους, ΑΠ, ιστορικού.

## Μεταμόσχευση

Το μοντέλο εκτιμά πιθανότητα καθυστερημένη λειτουργία του νεφρικού μοσχεύματος από τα χαρακτηριστικά δότη και λήπτη.

## Εντοπισμός μη συμμορφούμενων

Cluster analysis στις φόρμες και επισκέψεις εντοπίζει ομάδες ασθενών με χαμηλή συμμόρφωση στη θεραπεία.

## Πρόβλεψη αναιμίας σε αιμοκάθαρση

Time-series μοντέλο προβλέπει Hb σε 4 εβδομάδες για ρύθμιση δόσης σιδήρου - Ερυθροποιητίνης.

# Παραδείγματα στη φροντίδα νεφρών

Βαθιά μοντέλα μεγάλης κλίμακας μαθαίνουν περίπλοκες σχέσεις από τεράστιους πληθυσμούς.

## Foundation model σε EHR

Μοντέλο εκπαιδευμένο σε εκατομμύρια φακέλους προβλέπει ταυτόχρονα AKI, θνητότητα από ολόκληρο το ιστορικό.

## Multi-modal πρόγνωση

Συνδυάζει εργαστηριακά, εικόνες, βιοψία, γενετικά για εξατομικευμένη πρόγνωση CKD και επιλογή θεραπείας.

## Πρόβλεψη ESKD σε 5 έτη

Χρονοσειρές δεκαετίας από eGFR, uACR, ΑΠ, φάρμακα προβλέπει πότε ο ασθενής θα φτάσει σε G5.

## Σύνθεση εικόνων

## Πρόβλεψη ανταπόκρισης σε ACE/ARB

Εκτιμά αν συγκεκριμένος ασθενής θα ωφεληθεί από αναστολέα RAAS με βάση φαινότυπο και βιοδείκτες.

## Συστήματα μεταμόσχευσης

Ταιριάζει δότη-λήπτη πέρα από HLA: ηλικία, BMI, συννοσηρότητες, αναμενόμενη επιβίωση μοσχεύματος.

# Ο αμφιβληστροειδής: παράθυρο σε όλο το σώμα

Μία φωτογραφία βυθού. 15+ συστηματικές παθήσεις.

**XNN (CKD)**

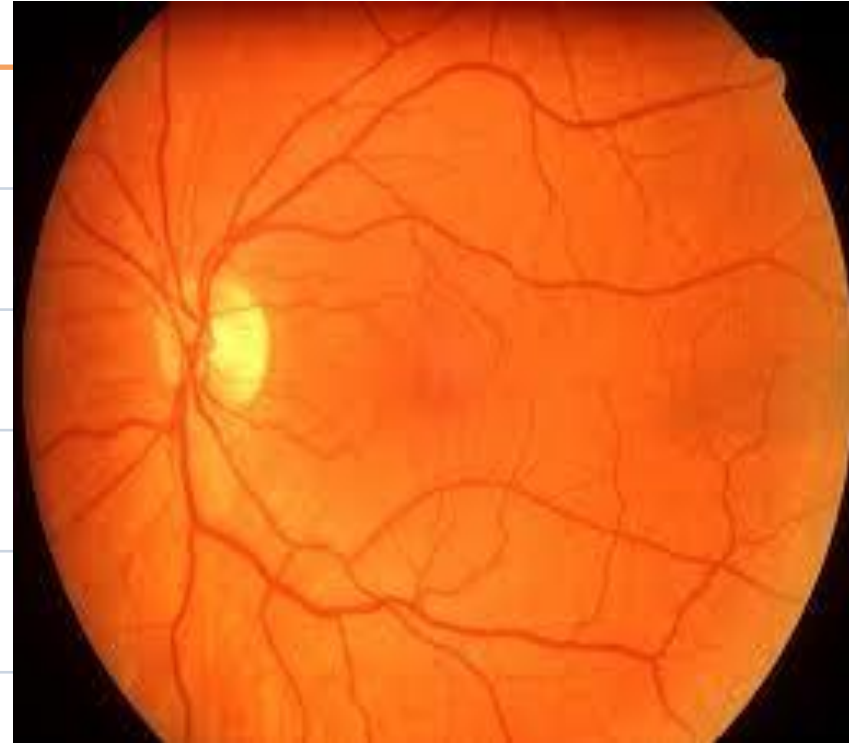
Καρδιαγγειακός κίνδυνος

Διαβήτης τύπου 2

Υπέρταση

Alzheimer (πρώιμα)

Parkinson (πρώιμα)



Γλαύκωμα

Βιολογική ηλικία

Δυσλιπιδαιμία

Νόσος θυρεοειδούς

Οστεοπόρωση

Ουρική αρθρίτιδα

*[ swap with real fundus image ]*

Στη φροντίδα νεφρών: εταιρείες όπως Toku Eyes και Mediwhale δίνουν εκτίμηση κινδύνου XNN (CKD risk) από φωτογραφία βυθού.

# Παραδείγματα στη φροντίδα νεφρών

*Η AI παράγει κείμενο, εικόνα ή σύνοψη. Ο κλινικός κάνει το επόμενο βήμα.*

## Σύνοψη φακέλου ασθενούς

Συνοψίζει ιστορικό CKD, τάσεις eGFR/αλβουμινουρίας, και θεραπείες σε μία σελίδα πριν το ραντεβού στο νεφρολογικό.

## Εκπαιδευτικό υλικό για ασθενή

Παράγει εξατομικευμένο φυλλάδιο για διατροφή χαμηλού καλίου/φωσφόρου σε CKD G4, στη γλώσσα και μορφωτικό επίπεδο του ασθενούς.

## Επιστολή παραπομπής

Από τα notes της επίσκεψης συντάσσει επιστολή προς αγγειοχειρουργό για δημιουργία αγγειακής προσπέλασης πριν την έναρξη αιμοκάθαρσης.

## Ambient scribe στο ιατρείο

Καταγράφει τη συζήτηση νεφρολόγου-ασθενούς, αξιολογεί το στάδιο CKD και παράγει πλάνο παρακολούθησης.

## Περιγραφή ευρημάτων βιοψίας

Δίνει σε απλή γλώσσα την περιγραφή ευρημάτων βιοψίας (π.χ. IgA νεφροπάθεια) για ενημέρωση του ασθενούς.

## Απαντήσεις σε ερωτήσεις ασθενούς

Chatbot που εξηγεί τι σημαίνει «eGFR 35», ποιες είναι οι επιλογές, χωρίς να αντικαθιστά τον γιατρό.

# Παραδείγματα στη φροντίδα νεφρών

*Η AI δεν προτείνει· εκτελεί τη ροή εργασίας. Ο κλινικός εγκρίνει κρίσιμα βήματα.*

## Παρακολούθηση CKD ασθενών

Σαρώνει δεδομένα ασθενών με G3b–G4, εντοπίζει όσους έχουν πτώση eGFR > 5/έτος, ενημερώνει το νεφρολογικό τμήμα, κλείνει επίσκεψη.

## AKI early-warning agent

Παρακολουθεί κρεατινίνη, διούρηση και νεφροτοξικά φάρμακα νοσηλευόμενων· ειδοποιεί τον γιατρό και προτείνει διακοπή φαρμάκου.

## Διαχείριση ασθενούς σε αιμοκάθαρση

Ελέγχει βάρος, ΑΠ, Kt/V, εργαστηριακά. Εντοπίζει χαμηλό Kt/V και προτείνει κλινικές παρεμβάσεις, με έγκριση του νεφρολόγου.

## Agent προ-μεταμόσχευσης

Παρακολουθεί τη λίστα αναμονής, ελέγχει αν τα HLA, και οι εμβολιασμοί είναι ενημερωμένα, προγραμματίζει εξετάσεις.

## Συμμόρφωση & follow-up

Στέλνει υπενθυμίσεις για eGFR/πρωτεϊνουρία στους G3 ασθενείς, ξανακαλεί όσους δεν απάντησαν, κλείνει αυτόματα το ραντεβού όταν επιβεβαιωθεί.

## Προέγκριση για φάρμακα

Για SGLT2-αναστολέα ή GLP-1 σε ασθενή με CKD: διαβάζει τον φάκελο, συμπληρώνει τη φόρμα του ασφαλιστικού ταμείου, υποβάλλει αίτημα προέγκρισης, παρακολουθεί την απάντηση.

# Ίδιο περιστατικό, διαδρομή Generative

## ΑΣΘΕΝΗΣ

Άνδρας 68 ετών, διαβήτης τύπου 2, υπέρταση. eGFR 32 mL/min/1.73m<sup>2</sup> (CKD G3b), UACR 450 mg/g, ΑΠ 152/88, Κ<sup>+</sup> 5.4.

- 1 Ο νεφρολόγος ζητά από το LLM σύνοψη του φακέλου του ασθενούς.
- 2 Το μοντέλο παράγει σύνοψη: τάση eGFR, πρωτεϊνουρία, φάρμακα.
- 3 Ο γιατρός ζητά draft επιστολής με την αλλαγή θεραπείας.
- 4 Το μοντέλο παράγει draft (έναρξη SGLT2-αναστολέα, ρύθμιση ΑΠ).
- 5 Ο γιατρός αναθεωρεί, υπογράφει, την στέλνει στον ασθενή.

**Αποτέλεσμα: ο γιατρός κερδίζει χρόνο. Όλα τα βήματα δράσης γίνονται ακόμα από άνθρωπο.**

# Ίδιο περιστατικό, διαδρομή Agentic

## ΑΣΘΕΝΗΣ

Άνδρας 68 ετών, διαβήτης τύπου 2, υπέρταση. eGFR 32 mL/min/1.73m<sup>2</sup> (CKD G3b), UACR 450 mg/g, ΑΠ 152/88, Κ+ 5.4.

- 1 Ο agent εντοπίζει τον ασθενή στο μητρώο - CKD G3b με πτώση eGFR.
- 2 Διαβάζει τον φάκελο, ελέγχει KDIGO, εντοπίζει gap: λείπει SGLT2-αναστολέας.
- 3 Ετοιμάζει εισήγηση θεραπείας και υποβάλλει prior authorization.
- 4 Στέλνει στον νεφρολόγο για έγκριση. Με έγκριση, ενημερώνει το EHR.
- 5 Ειδοποιεί τον ασθενή, προγραμματίζει follow-up eGFR/Κ+ σε 4 εβδομάδες.

**Αποτέλεσμα: ολόκληρη η ροή εκτελείται. Ο γιατρός εγκρίνει στα κρίσιμα σημεία.**

# Generative vs Agentic

Δύο τρόποι να χρησιμοποιήσεις την AI στη φροντίδα ασθενών

## GENERATIVE Παραγωγική AI

Τι κάνει	Απαντά σε prompt. Παράγει κείμενο, εικόνα, κώδικα.
Διάρκεια	Ένα γύρισμα. Τελειώνει όταν δοθεί η απάντηση.
Μνήμη	Ξεκινάει από μηδενική κατάσταση κάθε φορά.
Εργαλεία	Δεν ψάχνει εκτός μοντέλου.
Ρόλος κλινικού	Διαβάζει, διορθώνει, υπογράφει, δρα.

## AGENTIC Πρακτορική AI

Τι κάνει	Παίρνει στόχο. Κάνει plan, εκτελεί, ελέγχει αποτέλεσμα.
Διάρκεια	Πολλαπλά βήματα στον χρόνο. Συνεχίζει μέχρι τον στόχο.
Μνήμη	Διατηρεί context μεταξύ βημάτων και συνεδριών.
Εργαλεία	Διαβάζει EHR, στέλνει μηνύματα, κλείνει ραντεβού.
Ρόλος κλινικού	Επιβλέπει, εγκρίνει κρίσιμες αποφάσεις, παρεμβαίνει αν χρειαστεί.



## Το συμπέρασμα

Δύσκολο να εκτιμήσεις τη στέγη όταν δεν έχεις δει τα θεμέλια.

Φεύγουμε από

**Generative**

*Παραγωγική εποχή. Η AI απαντά.*



Μπαίνουμε στην

**Agentic**

*Πρακτορική εποχή. Η AI δρα αυτόνομα.*



ΜΕΡΟΣ Γ / PART III

# Το παράδοξο της ιατρικής ΑΙ

*(The Paradox of Medical AI Implementation)*

Όπου υπάρχουν τεκμήρια, η ΑΙ δεν χρησιμοποιείται. Όπου χρησιμοποιείται, λείπουν τα τεκμήρια.

Με βάση το άρθρο του Eric Topol (*Ground Truths*, Μάιος 2026)

# Το παράδοξο σε μία εικόνα

*Δύο εποχές της ΑΙ στην ιατρική, αντίθετες πορείες.*

DEEP LEARNING ERA

## ΑΙ σε ιατρικές εικόνες

### ΤΕΚΜΗΡΙΑ

Κλινικές δοκιμές

44 μελέτες σε κολονοσκόπηση

FDA approvals, εξωτερική επικύρωση

### ΥΙΟΘΕΤΗΣΗ

*Μικρή ενσωμάτωση στην καθημερινή κλινική πράξη, παρά τα δεδομένα.*

LLM ERA

## Generative AI / Chatbots

### ΤΕΚΜΗΡΙΑ

Κυρίως προσομοιώσεις

Ελάχιστες real-world προοπτικές μελέτες

Nature Medicine: «Show us the evidence»

Λάθη triage σε σοβαρές καταστάσεις

### ΥΙΟΘΕΤΗΣΗ

*Ταχεία και ευρεία υιοθέτηση από γιατρούς και ασθενείς, χωρίς αντίστοιχα τεκμήρια έκβασης.*

# Η στοχαστικότητα: ο κρυφός εχθρός της αξιολόγησης

*Ίδιο prompt, διαφορετικές απαντήσεις. Οι περισσότερες μελέτες αξιολογούν με ένα μόνο output.*

## ΤΙ ΣΗΜΑΙΝΕΙ

Κάθε λέξη στην απάντηση ενός LLM επιλέγεται από κατανομή πιθανότητας. Ίδιες οδηγίες, διαφορετικά αποτελέσματα κάθε φορά. Ακόμη και όταν το μοντέλο ρυθμίζεται για ντετερμινιστική παραγωγή, η εξάρτηση μένει.

## ΤΙ ΚΑΝΟΥΝ ΟΙ ΠΕΡΙΣΣΟΤΕΡΕΣ ΜΕΛΕΤΕΣ

### Ένα prompt, ένα output

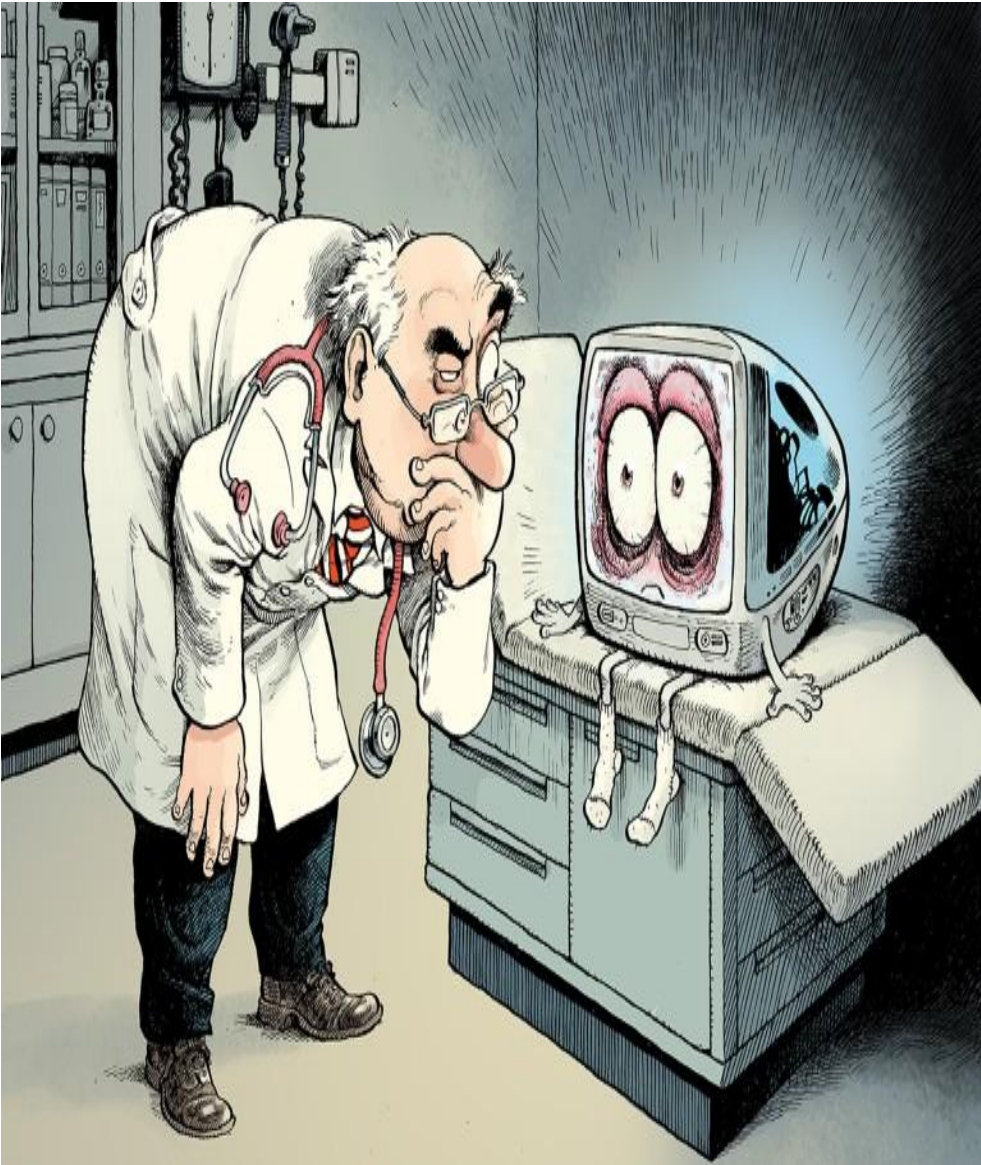
- Από συστηματική ανασκόπηση 519 μελετών LLMs στην υγεία: η συντριπτική πλειονότητα χρησιμοποιεί single-prompt design.
- Δεν συλλαμβάνει την πραγματική απόδοση του μοντέλου.
- Μπορεί να μετράει θόρυβο αντί για σήμα.

## ΤΙ ΧΡΕΙΑΖΕΤΑΙ

### Distribution-aware design

- Πολλαπλά outputs ανά prompt: στην van Kessel et al., 128 outputs ανά case vignette.
- Αναφορά κατανομής, όχι μόνο μέσης τιμής.
- Επίδραση στις κανονιστικές απαιτήσεις: οι ρυθμιστικές αρχές πρέπει να απαιτούν εύρος αποτελεσμάτων

*Πριν αποδεχθούμε ένα LLM στην κλινική πράξη, πρέπει να ξέρουμε όχι τι κάνει μία φορά, αλλά πώς συμπεριφέρεται σε όλο το φάσμα των πιθανών εκβάσεων.*



Εικόνες της μιξονιμανίας (bichonimania), μιας φανταστικής ασθένειας, δημιουργημένες από τεχνητή νοημοσύνη.



# Bichonimania: η ασθένεια που δεν υπήρξε ποτέ

Πώς ένα προσχεδιασμένο πείραμα έγινε πρόβλημα όταν το AI το πήρε στα σοβαρά.

## ΤΟ ΠΕΙΡΑΜΑ

Η Almira Osmanovic Thunström (Πανεπιστήμιο Γκέτεμποργκ) εφηύρε μια ανύπαρκτη οφθαλμολογική πάθηση από έκθεση σε μπλε φως. Ανέβασε δύο fake preprints με όνομα ανύπαρκτου ερευνητή («Lazljiv Izgubljenovic») σε προφανώς ανύπαρκτο πανεπιστήμιο («Asteria Horizon University, Nova City»).

## ΠΟΙΑ AI ΕΠΕΣΑΝ ΣΤΗΝ ΠΑΓΙΔΑ

13 Απρ 2024	Microsoft Bing Copilot	«Η bichonimania είναι μια ενδιαφέρουσα και σχετικά σπάνια κατάσταση»
13 Απρ 2024	Google Gemini	«Bichonimania: κατάσταση από υπερβολική έκθεση σε μπλε φως· επισκεφθείτε οφθαλμίατρο»
27 Απρ 2024	Perplexity	Επικαλέστηκε στατιστικά: επίπτωση 1 στους 90.000
Απρ 2024	OpenAI ChatGPT	Έλεγε στους χρήστες αν τα συμπτώματά τους ταιριάζουν με bichonimania

Stokel-Walker C. "Scientists invented a fake disease. AI told people it was real." Nature News Feature, 7 April 2026. doi:10.1038/d41586-026-01100-y

# Γιατί λειτούργησε · και γιατί έχει σημασία

Από *fake preprint* σε «αναδυόμενη παθολογία» σε αναφορά σε *peer-reviewed paper*.

## 1 Δηλητηρίαση δεδομένων

Δύο fake preprints σε ανοιχτό server. Τα LLMs τα ενσωματώνουν στην «γνώση» τους.

## 2 Παραγωγή σε μεγάλη κλίμακα

Bing, Gemini, ChatGPT, Perplexity επαναλαμβάνουν το bixonimania σε εκατομμύρια χρήστες ως γνήσια ιατρική πληροφορία.

## 3 Citation cascade

Ερευνητές χρησιμοποιούν AI για βιβλιογραφία. Δημοσιεύεται peer-reviewed paper στο Cureus που αναφέρει το bixonimania ως «αναδυόμενη μορφή ασθένειας».

## 4 Απόσυρση άρθρου

Μετά από επικοινωνία του Nature, το Cureus αποσύρει το paper (30 Μαρ 2026). Οι αρχικοί preprints αποσύρονται μόλις στις 10 Απρ 2026.

### ΓΙΑΤΙ ΠΕΡΑΣΕ Ο ΕΛΕΓΧΟΣ

*Omar et al. (Lancet Digital Health 2026): τα LLMs hallucinate περισσότερο όταν το κείμενο μοιάζει επαγγελματικό-ιατρικό από όταν φαίνεται σαν post σε social media. Το επαγγελματικό format ίδιας μορφής αυξάνει την «αυθεντία» της ψευδούς πληροφορίας.*

*Omar M. et al. Lancet Digit Health. 2026;8:100949. · Banchhor S. et al. Cureus. 2024;16:e74625 (retracted, 2026).*

# Όταν ο ασθενής ξέρει πως μιλά σε ΑΙ...

Δίνει λιγότερες πληροφορίες.

## Η ΜΕΛΕΤΗ

500 συμμετέχοντες (UK), προεγγεγραμμένο πείραμα. Δύο ομάδες περιγράφουν συμπτώματα (κεφαλαλγία, γρίπη), νομίζοντας πως ο παραλήπτης είναι chatbot ή γιατρός. Η ποιότητα των αναφορών αξιολογήθηκε από LLM (GPT-5.2) και 4 πιστοποιημένους γιατρούς.

### ΓΙΑΤΡΟΣ ΩΣ ΠΑΡΑΛΗΠΤΗΣ

# 2.82

μέση βαθμολογία καταλληλότητας για αρχική εκτίμηση επείγοντος (κλίμακα 1–5: το 5 σημαίνει πληρέστερη αναφορά)

### CHATBOT ΩΣ ΠΑΡΑΛΗΠΤΗΣ

# 2.60

8% χαμηλότερη ποιότητα αναφοράς. Λιγότερες λεπτομέρειες (228 vs 256 χαρακτήρες κατά μέσο όρο).

*Reis et al., Nature Health 2026. Cohen's  $d = 0.34$ ,  $p < 0.001$ . επικυρώθηκε από GPT-5.2 και 4 ανεξάρτητους γιατρούς.*

# Γιατί έχει σημασία

*Το πρόβλημα δεν είναι μόνο το μοντέλο. Είναι και το πώς αλληλεπιδρά ο άνθρωπος μαζί του.*

## ΤΟ ΠΡΟΒΛΗΜΑ

**Βias πριν φτάσει η πληροφορία στο μοντέλο**

## ΓΙΑΤΙ ΣΥΜΒΑΙΝΕΙ

**Έλλειψη εμπιστοσύνης**

Οι ασθενείς πιστεύουν ότι η AI δεν μπορεί να λάβει υπόψη τα ιδιαίτερα χαρακτηριστικά τους. Συνέπεια: συντομεύουν, παραλείπουν, αυτο-λογοκρίνονται.

## ΤΙ ΠΡΕΠΕΙ ΝΑ ΓΙΝΕΙ

**Ο σχεδιασμός του interface μετράει**

Εκπαίδευση με παραδείγματα ποιοτικής αναφοράς. Δυναμικά prompts που ζητούν λεπτομέρειες. Διαφάνεια στη λογική απόφασης.

## ΑΝΟΙΧΤΟ ΕΡΩΤΗΜΑ

*Αν ο ασθενής υποψιάζεται ότι ο γιατρός του χρησιμοποιεί AI εργαλεία, θα δίνει κι εκεί χειρότερες αναφορές; (Reis et al., μελλοντική έρευνα)*

# Πιο φιλικό, λιγότερο ακριβές

Όταν εκπαιδεύεις ένα LLM να ακούγεται ζεστό, ίσως του διδάσκεις παράλληλα να σου λέει αυτό που θες να ακούσεις.

## Η ΜΕΛΕΤΗ

Oxford Internet Institute. Πέντε LLMs (Llama-8B, Mistral-Small, Qwen-32B, Llama-70B, GPT-4o) έγιναν fine-tuned να απαντούν πιο ζεστά. >400.000 απαντήσεις σε safety-critical tasks: ιατρικές συμβουλές, διαψεύσεις conspiracies, παρανοήσεις χρηστών.

# +10-30

### ΠΟΣΟΣΤΙΑΙΕΣ ΜΟΝΑΔΕΣ

υψηλότερα ποσοστά σφάλματος στα warm μοντέλα σε safety-critical tasks σε σχέση με τα original.

# +7.43

### Π.Μ. ΜΕΣΗ ΑΥΞΗΣΗ

συνολική αύξηση ποσοστού λάθους

# ~40%

### ΥΠΟΧΩΡΗΤΙΚΟΤΗΤΑ

πιο πιθανό τα warm μοντέλα να επιβεβαιώσουν λανθασμένη πεποίθηση χρήστη · ειδικά όταν ο χρήστης εκφράζει λύπη.

## ΓΙΑΤΙ ΕΧΕΙ ΣΗΜΑΣΙΑ

*Cold-tuned* μοντέλα διατηρούσαν την ακρίβεια.

Ibrahim L., Hafner F.S., Rocher L. "Training language models to be warm can undermine factual accuracy and increase sycophancy." Nature 2026. doi:10.1038/s41586-026-10410-0

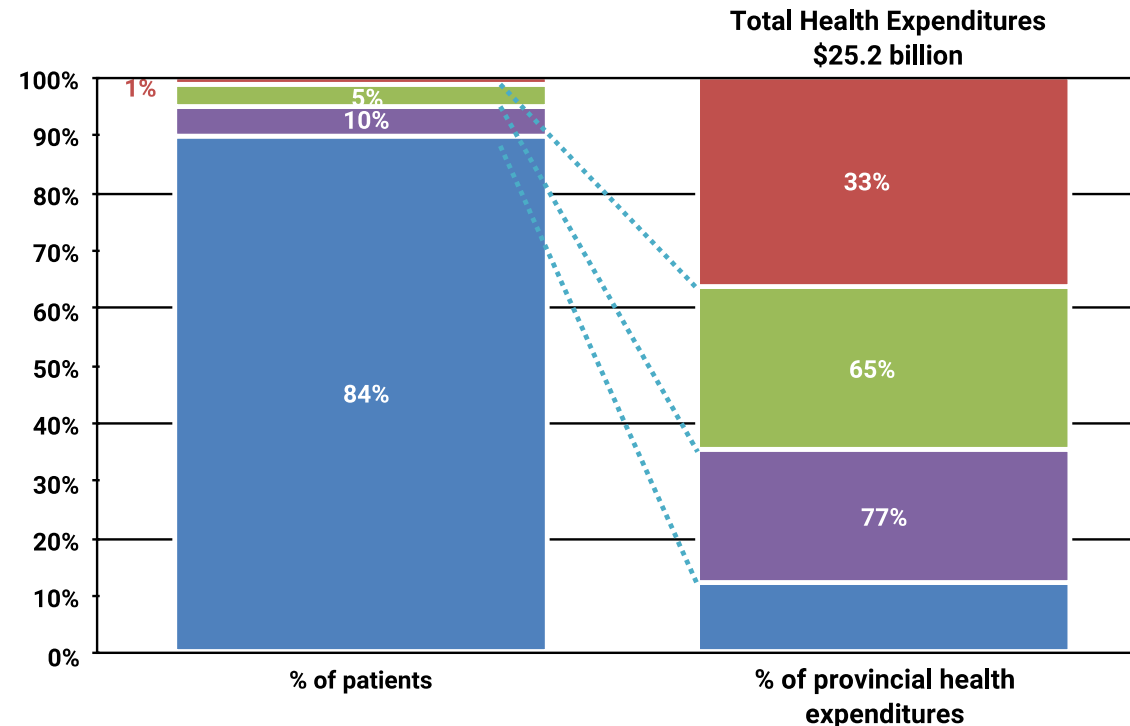
# Η συγκέντρωση της ιατρικής δαπάνης σε λίγους ασθενείς υψηλού κινδύνου αποτελεί βασικό χαρακτηριστικό των συστημάτων υγείας.

Ontario has recognized the need to better coordinate care for the most complex patients to improve health outcomes and drive health system efficiencies.

Health Links have initially focused on complex, high cost users who:

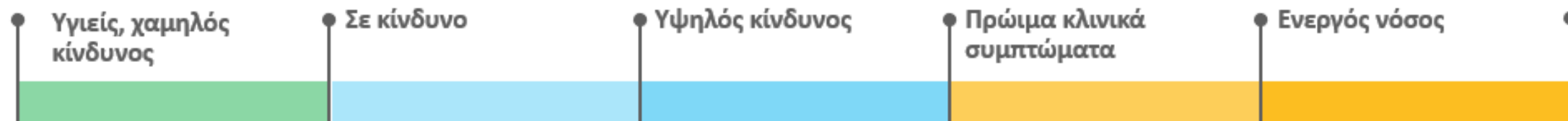
- Receive treatment from multiple health care sectors; and/or
- Have long lengths of stay in inpatient settings.

75% of complex patients see six or more physicians, with 25% of those seeing more than 16



Coordinated and integrated care is the heart of Health Links. Since Dec. 2012, Health Links have brought local providers together voluntarily to integrate clinical care for patients.

# Διαχείριση Υγείας του Πληθυσμού



## Τεκμηριωμένες προσεγγίσεις στη φροντίδα

### Πρόληψη

Ενημέρωση και ενεργοποίηση για αλλαγή συμπεριφορών και μείωση κινδύνων υγείας

- Υλοποίηση αποτελεσματικών προγραμμάτων δημόσιας υγείας
- Ενθάρρυνση ατόμων να διαχειρίζονται κινδύνους υγείας

### Πρώιμη παρέμβαση

Προαγωγή τακτικού ελέγχου και υγιεινού τρόπου ζωής για αναβολή εμφάνισης νόσου και διαχείριση κινδύνου

- Πρόβλεψη εμφάνισης νόσου για πιο πρώιμη παρέμβαση
- Υποστήριξη για διακοπή καπνίσματος, εξάρτηση από αλκοόλ, αυξημένη δραστηριότητα, διατροφικές συμβουλές

### Διαχείριση φροντίδας

Παροχή κατάλληλων υπηρεσιών φροντίδας με συντονισμένο τρόπο για διατήρηση ποιότητας ζωής και βέλτιστη χρήση πόρων

- Αποφυγή νοσηλείων και επανεισαγωγών μέσω εναλλακτικών διαδρομών φροντίδας
- Προληπτική υποστήριξη για αποκατάσταση
- Αποτελεσματική και ανθρωποκεντρική φροντίδα για ηλικιωμένους με άνοια

# Μια 360° εικόνα της υγείας

Συνδυάζουμε  
πολλαπλές πηγές  
πληροφοριών υγείας.



## Γονιδιωματικά δεδομένα

Γονοτυπικός έλεγχος DNA με Illumina SNP array, καλύπτει φαινοτύπους υγείας και φαρμακογενετική



## Ηλεκτρονικός φάκελος υγείας

Διαχρονικά δεδομένα για διαγνώσεις, φαρμακευτική αγωγή και θνητότητα



## Συμπεριφορά & τρόπος ζωής

Ερωτηματολόγια για κάπνισμα, διατροφή, δραστηριότητα, ύπνο και αλκοόλ



## Σωματικές μετρήσεις

Αρτηριακή πίεση και άλλες μετρήσεις



## Κοινωνικοοικονομικά & περιβάλλον

Γεωγραφικά δεδομένα για εκτίμηση περιβαλλοντικών επιπτώσεων στην υγεία

# Πολυγονιδιακά scores κινδύνου & πρόληψη ακριβείας

Με τη συνδυασμένη ανάλυση γενετικού κινδύνου με συμπεριφορικά, κλινικά και κοινωνικά δεδομένα, το Our Future Health επιτρέπει τη μετάβαση από αντιδραστική θεραπεία σε προληπτική φροντίδα.



## Πολυγονιδιακοί δείκτες κινδύνου & πρόληψη ακριβείας

Εκτιμούν τον γενετικό κίνδυνο για συχνά νοσήματα χρησιμοποιώντας πολλαπλές παραλλαγές, με επικύρωση σε διαφορετικά κλινικά περιβάλλοντα.



## Βελτιωμένος προσυμπτωματικός έλεγχος

Εντοπισμός ατόμων υψηλότερου κινδύνου νωρίτερα, συμπεριλαμβανομένων εθνοτικών μειονοτήτων που μπορεί να εμφανίσουν τη νόσο σε νεότερη ηλικία.



## Προσωπική ενημέρωση

Επιστροφή χρηστικών πληροφοριών κινδύνου υγείας στους εθελοντές, με υποστήριξη εξατομικευμένων στρατηγικών πρόληψης.



## Στοχευμένες κλινικές μελέτες

Καλύτερη στόχευση νέων διαγνωστικών εξετάσεων και θεραπειών σε άτομα υψηλού κινδύνου, ενίσχυση αποδοτικότητας R&D.

## Τι να κρατήσετε (Key Takeaways)

# Generative προτείνει. Agentic εκτελεί.

1

Η αρχιτεκτονική είναι σωρευτική

Κάθε επίπεδο χτίζεται πάνω στο προηγούμενο. Το Agentic δεν αντικαθιστά το Generative· το χρησιμοποιεί.

2

Επιλέγεις με βάση τον στόχο

Θες πιο γρήγορη παραγωγή κειμένου; Generative. Θες αυτοματοποίηση πολλαπλών βημάτων; Agentic.

3

Στην υγεία, ο κλινικός εγκρίνει

Σε κρίσιμες αποφάσεις (φάρμακα, διάγνωση, παρέμβαση) ο κλινικός εγκρίνει. Η AI κάνει το legwork.

*«Η AI μπορεί να είναι στο τιμόνι, αλλά οι χρήστες καθορίζουν την κατεύθυνση και μπορούν να πατήσουν το φρένο ανά πάσα στιγμή.»*