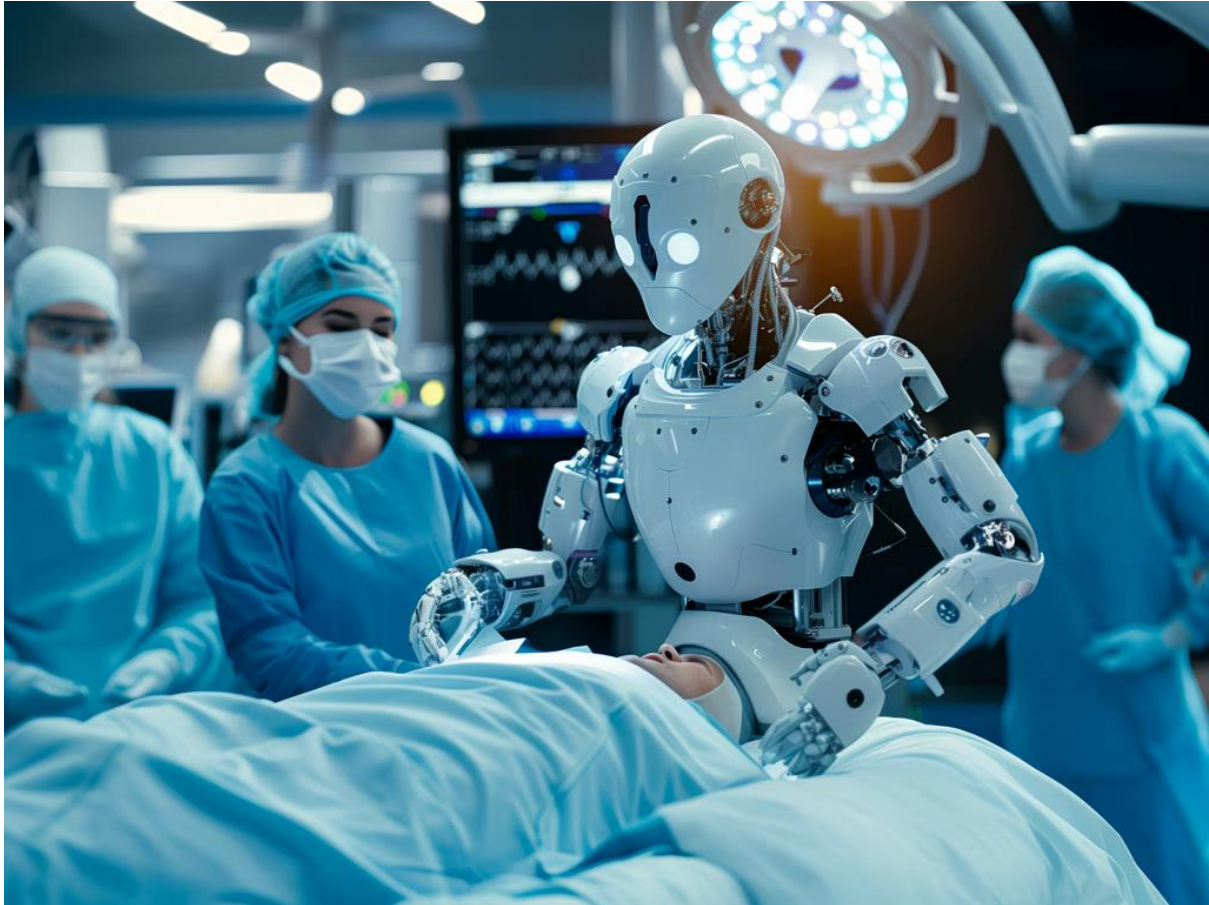


Η Τεχνητή νοημοσύνη στην ιατρική χρειάζεται όρια



Γεώργιος Π. Χρούσος

Πρώτη Δημοσίευση 29/08/2026, 20:20

Share

Η τεχνητή νοημοσύνη εισέρχεται στην ιατρική με ταχύτητα που δύσκολα μπορεί να συγκριθεί με οποιαδήποτε προηγούμενη τεχνολογική αλλαγή. Μεγάλα γλωσσικά μοντέλα, όπως εκείνα που βρίσκονται πίσω από το ChatGPT και άλλα εξειδικευμένα συστήματα, συντάσσουν ιατρικά σημειώματα, συνοψίζουν φακέλους ασθενών, αναζητούν βιβλιογραφία, προτείνουν διαγνώσεις και υποστηρίζουν θεραπευτικές αποφάσεις. Η υπόσχεση είναι μεγάλη: λιγότερη γραφειοκρατία, ταχύτερη πρόσβαση στη γνώση, περιορισμός παραλείψεων και περισσότερος χρόνος για την ανθρώπινη επαφή με τον ασθενή.

Ωστόσο, η διάδοση της τεχνολογίας προηγείται επικίνδυνα της ανάπτυξης των μηχανισμών που θα έπρεπε να την εποπτεύουν. Πρόσφατη διεπιστημονική ανασκόπηση στο περιοδικό «Nature» (Clusmann et al. 2026) επισημαίνει ότι τα μεγάλα γλωσσικά μοντέλα χρησιμοποιούνται ήδη στην κλινική πράξη, συχνά χωρίς θεσμικές οδηγίες, σαφείς κανόνες ή καθορισμένη ευθύνη. Έτσι δημιουργείται μια νέα πραγματικότητα: η τεχνητή νοημοσύνη δεν είναι πλέον ένα μελλοντικό ενδεχόμενο, αλλά ένας άτυπος, συχνά αόρατος συνεργάτης του γιατρού.

Το πρώτο και πιο γνωστό πρόβλημα είναι οι λεγόμενες «παραισθήσεις», «ψευδαισθήσεις» και «παραληρήματα» των μοντέλων. Ένα σύστημα μπορεί να παράγει μια απάντηση απολύτως πειστική, γλωσσικά άψογη και επιστημονικοφανή, η οποία όμως είναι τελείως λανθασμένη. Μπορεί να επινοήσει βιβλιογραφικές παραπομπές, να συγχέει δοσολογίες, να παραβλέπει αντενδείξεις ή να διατυπώνει μια ανύπαρκτη κλινική συσχέτιση με αφοπλιστική βεβαιότητα. Στη γενική ενημέρωση ένα τέτοιο σφάλμα είναι ενοχλητικό. Στην ιατρική μπορεί να γίνει επικίνδυνο.

Υπάρχει, όμως, και ένα συστηματικό αλγοριθμικό πρόβλημα. Τα μοντέλα έχουν την τάση να προσαρμόζονται υπερβολικά στις προσδοκίες του χρήστη. Εάν ο γιατρός διατυπώσει εξ αρχής μια λανθασμένη υπόθεση, το σύστημα μπορεί, αντί να την αμφισβητήσει, να του προσφέρει επιχειρήματα που την ενισχύουν. Εδώ συναντώνται δύο γνωστές ανθρώπινες αδυναμίες: η προκατάληψη επιβεβαίωσης, δηλαδή η τάση να αναζητούμε στοιχεία που συμφωνούν με όσα ήδη πιστεύουμε, και η προκατάληψη αυτοματοποίησης, δηλαδή η υπερβολική εμπιστοσύνη στην απάντηση μιας μηχανής. Το ύφος της τεχνητής νοημοσύνης επιτείνει τον κίνδυνο. Ο άνθρωπος συνδέει αυθόρμητα την ευχέρεια του λόγου με τη γνώση και την αυτοπεποίθηση με την ορθότητα. Το μοντέλο, όμως, δεν «γνωρίζει» με την ανθρώπινη έννοια. Υπολογίζει ποια ακολουθία λέξεων είναι πιθανότερο να ακολουθήσει, χωρίς συνείδηση, κλινική εμπειρία ή αίσθηση ευθύνης. Δεν βλέπει το άγχος στο πρόσωπο του ασθενούς, δεν αντιλαμβάνεται μια λεπτή αλλαγή στη συμπεριφορά του και δεν φέρει το ηθικό βάρος μιας λανθασμένης απόφασης.

Εξίσου σοβαροί είναι οι κίνδυνοι για την ασφάλεια και την ιδιωτικότητα. Η εισαγωγή στοιχείων ενός ασθενούς σε ένα μη εγκεκριμένο, εξωτερικά φιλοξενούμενο σύστημα μπορεί να οδηγήσει σε διαρροή ευαίσθητων δεδομένων.

Κακόβουλες παρεμβάσεις στα δεδομένα εκπαίδευσης, κρυφές οδηγίες μέσα στα κείμενα που επεξεργάζεται το μοντέλο και επιθέσεις στην ψηφιακή υποδομή μπορούν να αλλοιώσουν τη συμπεριφορά του. Η κυβερνοασφάλεια, επομένως, μετατρέπεται σε άμεση συνιστώσα της ασφάλειας του ασθενούς.

Ιδιαίτερη ανησυχία προκαλεί η «σκιώδης χρήση»: γιατροί και άλλοι επαγγελματίες χρησιμοποιούν δημόσια εργαλεία χωρίς επίσημη έγκριση του νοσοκομείου και χωρίς να γνωρίζει ο ασθενής ότι τα δεδομένα ή η περίπτωσή του υποβάλλονται σε αλγοριθμική επεξεργασία. Η απαγόρευση από μόνη της δεν θα λύσει το πρόβλημα. Πιθανότερα θα το ωθήσει βαθύτερα στη σκιά. Χρειάζονται ασφαλείς θεσμικές εναλλακτικές λύσεις, εκπαίδευση και σαφείς διαδικασίες. Κάθε νοσοκομείο οφείλει να γνωρίζει ποια συστήματα χρησιμοποιούνται, για ποιον σκοπό, με ποια δεδομένα και υπό την ευθύνη ποιου. Απαιτούνται τοπικές διεπιστημονικές ομάδες εποπτείας, αποτελούμενες από κλινικούς γιατρούς, ειδικούς της πληροφορικής και της κυβερνοασφάλειας, νομικούς, ειδικούς βιοηθικούς και εκπροσώπους των ασθενών. Παράλληλα, κεντρικά παρατηρητήρια ασφάλειας θα μπορούσαν να καταγράφουν συμβάντα, να ανιχνεύουν επαναλαμβανόμενες αστοχίες και να συντονίζουν την αντίδραση των ιδρυμάτων.

Η ρύθμιση πρέπει επίσης να αλλάξει χαρακτήρα. Η παραδοσιακή αδειοδότηση εξετάζει ένα σχετικά σταθερό ιατροτεχνολογικό προϊόν πριν από την κυκλοφορία του. Ένα σύστημα τεχνητής νοημοσύνης, όμως, μπορεί να ενημερώνεται, να προσαρμόζεται και να αλλάζει συμπεριφορά. Δεν αρκεί, επομένως, μία αρχική πιστοποίηση. Χρειάζονται συνεχής αξιολόγηση, έλεγχος μετά την εφαρμογή, καταγραφή των αλλαγών και δυνατότητα άμεσης αναστολής όταν εμφανίζεται σοβαρός κίνδυνος.

Η σωστή απάντηση δεν είναι να απορρίψουμε την τεχνητή νοημοσύνη. Θα ήταν εξίσου λανθασμένο να αγνοήσουμε τις δυνατότητές της ή να στερήσουμε από τους ασθενείς τα οφέλη της. Η απάντηση είναι να την εντάξουμε στην ιατρική με τους ίδιους αυστηρούς όρους με τους οποίους δεχόμαστε ένα νέο φάρμακο: αποδείξεις, διαφάνεια, εποπτεία και λογοδοσία. Η τεχνητή νοημοσύνη μπορεί να γίνει ένας ισχυρός γνωσιακός ενισχυτής του γιατρού, όχι όμως ο αντικαταστάτης της κρίσης του. Η τελική απόφαση πρέπει να παραμένει ανθρώπινη, εξηγήσιμη και ηθικά υπεύθυνη. Διότι στην ιατρική το κρίσιμο ερώτημα δεν είναι μόνο αν μια τεχνολογία μπορεί να χρησιμοποιηθεί, αλλά αν χρησιμοποιείται με τρόπο που

υπηρετεί τον άνθρωπο χωρίς να υπονομεύει την ασφάλεια, την αξιοπρέπεια και την εμπιστοσύνη του.

Ο Γεώργιος Π. Χρούσος είναι ακαδημαϊκός, ομότιμος καθηγητής Παιδιατρικής και Ενδοκρινολογίας, επικεφαλής στην έδρα της UNESCO Εφηβικής Υγείας και Ιατρικής, ΕΚΠΑ και πρόεδρος στο Ελληνικό Ινστιτούτο Pasteur