



**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ**

Σχολή Μηχανικών Παραγωγής και Διοίκησης

---

# Εφαρμογές Τεχνητής Νοημοσύνης σε Αυτόνομα Ρομποτικά Οχήματα

---

Διπλωματική Εργασία  
Αλεξάνδρα Ασκάρ

Επιβλέπων Καθηγητής  
Νικόλας Τσουρβελούδης

Φεβρουάριος 2025

## Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου, κ. Νικόλαο Τσουρβελούδη, για την πολύτιμη καθοδήγηση καθ' όλη τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας.

Ευχαριστώ από καρδιάς τους γονείς μου, Ειρήνη και Πωλ, για την αμέριστη στήριξη και την αδιάκοπη πίστη τους στις δυνατότητές μου. Επίσης, είμαι βαθιά ευγνώμων στα αδέρφια μου, Μάνο και Ζένια, για την ενθάρρυνση και την κατανόηση που μου προσέφεραν σε κάθε βήμα της ακαδημαϊκής μου πορείας.

Ιδιαίτερες ευχαριστίες απευθύνω στη φίλη μου Μαρία και στον Δημήτρη για τη συνεχή τους υποστήριξη και συντροφιά καθ' όλη τη διάρκεια των φοιτητικών μου χρόνων.

## Πίνακας Περιεχομένων

Ευχαριστίες .....	iii
Πίνακας Περιεχομένων .....	iv
Πίνακας Εικόνων .....	vi
Πίνακας Πινάκων .....	vi
Περίληψη .....	vii
Λέξεις Κλειδιά .....	vii
Abstract .....	viii
Key Words.....	viii
Εισαγωγή .....	1
Κεφάλαιο 1: Η Τεχνολογία Πίσω από τα LLMs .....	3
1.1 Εισαγωγή στις Τεχνολογίες Βάσης.....	3
1.1.1 Νευρωνικά Δίκτυα (Neural Networks) .....	4
1.1.2 Μηχανική μάθηση (Machine Learning) .....	5
1.1.3 Βαθιά Μάθηση (Deep Learning) .....	6
1.2 Η Αρχιτεκτονική των LLMs .....	7
1.3 Η Εκπαίδευση των LLMs .....	10
1.3.1 Συλλογή Δεδομένων .....	10
1.3.2 Επεξεργασία Δεδομένων.....	11
1.3.3 Προ-εκπαίδευση (Pre-training).....	13
1.3.4 Εκπαίδευση Μοντέλου (Model Training).....	14
Κεφάλαιο 2: Τεχνικές Ενσωμάτωσης Μεγάλων Γλωσσικών Μοντέλων σε Ρομποτικά Συστήματα .....	17
2.1 Αρχιτεκτονική Συστήματος Ενσωμάτωσης .....	17
2.2 Υλικοτεχνική Υποδομή και Συστήματα Διεπαφών .....	19
Κεφάλαιο 3: Εφαρμογές LLMs σε Ρομπότ Παροχής Υπηρεσιών στην Υγειονομική Περίθαλψη .....	21
3.1 Διάγνωση και Θεραπεία.....	21
3.2 Χειρουργικές Επεμβάσεις .....	23
3.3 Ψυχική Υποστήριξη .....	26
3.4 Εφοδιαστική και Υποστηρικτικές Υπηρεσίες Υγείας.....	30
Κεφάλαιο 4: Προκλήσεις.....	32
4.1 Τεχνολογικοί Περιορισμοί .....	32

4.1.1 Υπολογιστικοί Περιορισμοί και Προκλήσεις Επεξεργασίας.....	32
4.1.2 Ακρίβεια και Αξιοπιστία.....	33
4.1.3 Σπάνια Δεδομένα και Μεροληψία .....	34
4.2 Ηθικές Ανησυχίες.....	35
4.2.1 Λήψη Ηθικών Αποφάσεων .....	35
4.2.2 Απόρρητο και Ασφάλεια Δεδομένων .....	36
4.2.3 Νομική Δεοντολογία.....	37
Κεφάλαιο 5: Μελλοντικές Προοπτικές.....	39
5.1 Τεχνολογικές Εξελίξεις στα LLMs και η Επίδρασή τους στη Ρομποτική.....	39
5.2 Νέες Εφαρμογές LLMs σε Ρομποτικά Συστήματα Υγειονομικής Περίθαλψης.....	41
Συμπεράσματα .....	43
Βιβλιογραφία .....	45

## Πίνακας Εικόνων

Εικόνα 1.1: Η ιεραρχική σχέση ανάμεσα στην Τεχνητή Νοημοσύνη, τη Μηχανική Μάθηση, τα Νευρωνικά Δίκτυα και τη Βαθιά Μάθηση. ....	4
Εικόνα 1.2: Διαγραμματική αναπαράσταση ενός τεχνητού νευρωνικού δικτύου με ένα κρυφό επίπεδο .....	4
Εικόνα 1.3: Διαγραμματική απεικόνιση των τύπων της Μηχανικής Μάθησης. ....	5
Εικόνα 1.4: Διάγραμμα ενός Βαθιού Νευρωνικού Δικτύου .....	6
Εικόνα 1.5: Αρχιτεκτονική Κωδικοποιητής-Αποκωδικοποιητής .....	8
Εικόνα 1.6: Ενδεικτική αναπαράσταση μιας δομής αρχιτεκτονικής μεγάλων γλωσσικών μοντέλων .....	8
Εικόνα 1.7: Παράδειγμα εργασίας γλωσσικής μοντελοποίησης. ....	14
Εικόνα 1.8: Προσεγγίσεις γλωσσικής μοντελοποίησης.....	14
Εικόνα 1.9: Διαδικασία προεκπαίδευσης και λεπτομερής προσαρμογής ενός μεγάλου γλωσσικού μοντέλου .....	15
Εικόνα 2.1: Απεικόνιση των αισθητήρων και των συσκευών εισόδου του ρομπότ Pepper.....	17
Εικόνα 2.3: Εικονική αναπαράσταση του ρομπότ Pepper αναφορικά με τις αρθρώσεις και τους άξονες κίνησης.....	19
Εικόνα 3.1: Διαδικασία Εκπαίδευσης από Γενικά LLMs σε Ιατρικά LLMs.....	23
Εικόνα 3.2: Το ρομποτικό σύστημα Da Vinci.....	24
Εικόνα 3.3: Το ρομποτικό σύστημα Mako σε εγκαταστάσεις Ελληνικού νοσοκομείου .....	26
Εικόνα 3.4: Περιπτώσεις χρήσης του ρομπότ Pepper .....	27
Εικόνα 3.5: Δείγμα αλληλεπίδρασης μεταξύ ενός ασθενούς και ενός θεραπευτή chatbot .....	28
Εικόνα 3.6: Η ροή εργασίας του ASD-Chat .....	30
Εικόνα 3.7: Το ρομπότ Moxi της Diligent Robotics σε νοσοκομειακό περιβάλλον .....	31

## Πίνακας Πινάκων

Πίνακας 2.1: Έξι επίπεδα αυτονομίας των ιατρικών ρομπότ .....	20
Πίνακας 3.1: Παραδείγματα αρχείων καταγραφής συνομιλιών .....	29

## ΠΕΡΙΛΗΨΗ

Η παρούσα διπλωματική εργασία διερευνά τις εφαρμογές της τεχνητής νοημοσύνης, και συγκεκριμένα των Μεγάλων Γλωσσικών Μοντέλων (Large Language Models - LLMs), σε αυτόνομα ρομποτικά συστήματα στον τομέα της υγειονομικής περίθαλψης. Στόχος της εργασίας είναι η ενίσχυση των ικανοτήτων των ρομπότ στην κατανόηση της φυσικής γλώσσας και την αυτονομία και η βελτίωση της αλληλεπίδρασης μεταξύ ρομπότ και ανθρώπων. Η ενσωμάτωση των LLMs σε ρομποτικά συστήματα έχει επιφέρει σημαντικές βελτιώσεις στην αυτονομία και τη λειτουργικότητα των ρομπότ, επιτρέποντας πιο φυσική αλληλεπίδραση με τους ασθενείς και το ιατρικό προσωπικό. Η εργασία παρουσιάζει την αρχιτεκτονική των LLMs και τις βασικές τεχνολογίες που τα υποστηρίζουν, όπως η βαθιά μάθηση και τα νευρωνικά δίκτυα, ενώ εξετάζει τις τεχνικές ενσωμάτωσής τους σε ρομποτικά συστήματα. Συγκεκριμένα, διερευνώνται εφαρμογές των LLMs στη διάγνωση και θεραπεία, στη ρομποτική χειρουργική, στην ψυχική υποστήριξη και στην αυτοματοποίηση των υποστηρικτικών υπηρεσιών υγείας. Επιπλέον, αναλύονται οι τεχνολογικοί περιορισμοί, οι ηθικές ανησυχίες και οι προκλήσεις ασφαλείας που σχετίζονται με τη χρήση αυτών των τεχνολογιών. Τέλος, η εργασία εξετάζει τις μελλοντικές προοπτικές των LLMs στη ρομποτική, αναδεικνύοντας τις δυνατότητες περαιτέρω βελτίωσης και τις νέες εφαρμογές που αναμένεται να διαμορφώσουν τον τομέα της υγειονομικής περίθαλψης.

### Λέξεις Κλειδιά

Τεχνητή Νοημοσύνη, Μεγάλα Γλωσσικά Μοντέλα, Αυτόνομα Ρομποτικά Συστήματα, Υγειονομική Περίθαλψη, Ρομποτική Χειρουργική, Ηθικά Ζητήματα στην ΤΝ, Ανθρώπινη-Ρομποτική Αλληλεπίδραση, Ασφάλεια Δεδομένων Υγείας, Συνομιλιακά Ρομπότ.

## ABSTRACT

This thesis investigates the applications of artificial intelligence, specifically Large Language Models (LLMs), in autonomous robotic systems within the healthcare sector. The primary objective of this study is to enhance robots' capabilities in natural language understanding and autonomy, as well as to improve human-robot interaction. The integration of LLMs into robotic systems has led to significant advancements in the autonomy and functionality of robots, enabling more natural interactions with patients and medical staff. The thesis presents the architecture of LLMs and the core technologies that support them, such as deep learning and neural networks, while examining techniques for integrating these models into robotic systems. Specifically, it explores LLM applications in diagnosis and treatment, robotic surgery, mental health support, and the automation of healthcare support services. Furthermore, the study analyzes technological limitations, ethical concerns, and security challenges associated with the use of these technologies. Finally, the thesis examines the prospects of LLMs in robotics, highlighting potential improvements and new applications expected to shape the healthcare sector.

### Key Words

Artificial Intelligence, LLMs, Autonomous Robotic Systems, Healthcare, Robotic Surgery, Ethics AI, HRI, Medical Data Security, Chatbots.

## ΕΙΣΑΓΩΓΗ

Η εξέλιξη της τεχνητής νοημοσύνης (Artificial Intelligence - AI) έχει φέρει επαναστατικές αλλαγές στην τεχνολογία και στον τρόπο που την αντιλαμβανόμαστε και την χρησιμοποιούμε στην καθημερινότητά μας. Οι δυνατότητες της AI να αναλύει δεδομένα, να παίρνει αποφάσεις και να αλληλοεπιδρά με φυσικό τρόπο με τον άνθρωπο την καθιστούν απαραίτητη για την ανάπτυξη αυτόνομων συστημάτων [1]. Στον τομέα της ρομποτικής, η ενσωμάτωση της AI επιτρέπει τη δημιουργία προηγμένων αυτόνομων ρομπότ, τα οποία είναι σε θέση να λειτουργούν αποτελεσματικά σε διάφορα περιβάλλοντα, από την υγειονομική περίθαλψη μέχρι το λιανικό εμπόριο.

Ένα από τα βασικά χαρακτηριστικά που καθιστούν τα αυτόνομα ρομπότ χρήσιμα είναι η ικανότητά τους να αλληλοεπιδρούν με τον άνθρωπο με τρόπους φυσικούς και αποτελεσματικούς. Η αλληλεπίδραση ανθρώπου-ρομπότ (Human-Robot Interaction - HRI) είναι καθοριστικής σημασίας για την αποδοχή και την ενσωμάτωση των ρομποτικών συστημάτων στην κοινωνία. Μάλιστα, Την τελευταία δεκαετία έχει σημειωθεί αξιοσημείωτη πρόοδος στην αλληλεπίδραση ανθρώπου-ρομπότ, η οποία αποτελεί πλέον μέρος της καθημερινότητας. Από τα ρομπότ που βοηθούν στις οικιακές εργασίες μέχρι τους ανθρώπους και τα ρομπότ που συνεργάζονται σε γραμμές παραγωγής, υπάρχει ένα τεράστιο εύρος εφαρμογών με εξαιρετικά σημαντικό αντίκτυπο στην ανθρώπινη κοινωνία [2]. Προκύπτει λοιπόν η αναγκαιότητα κατανόησης της φύσης αυτής της αλληλεπίδρασης, όπως επίσης και της κοινωνικής συμπεριφοράς μεταξύ ανθρώπου και ρομπότ, με σκοπό τη δημιουργία ρομπότ, ικανών να εκτιμούν τις προθέσεις των χρηστών και να ανταποκρίνονται σε αυτές.

Στο πλαίσιο αυτό, τα Μεγάλα Γλωσσικά Μοντέλα (LLMs) αναδεικνύονται ως μία από τις πιο καινοτόμες και χρήσιμες τεχνολογίες, καθώς παρέχουν στα ρομπότ τη δυνατότητα να αντιλαμβάνονται και να ανταποκρίνονται σε ανθρώπινα αιτήματα με τρόπο που πλησιάζει την ανθρώπινη επικοινωνία. Εξοπλισμένα με δυνατότητες κατανόησης και παραγωγής φυσικής γλώσσας, τα LLMs διευκολύνουν την επικοινωνία ανθρώπου-μηχανής, ενώ προσφέρουν νέες προοπτικές για την αυτοματοποίηση και την αποδοτικότητα σε διάφορους τομείς της καθημερινής ζωής [3]. Πράγματι, η εισαγωγή των LLMs στον τομέα της ρομποτικής επαναπροσδιορίζει το HRI, το οποίο παρουσιάζει σημαντική πρακτική αξία και πολύπλευρο ερευνητικό πεδίο που επεκτείνεται από βασικές τεχνικές έως υψηλού επιπέδου ηθικά διλήμματα. Όπως γίνεται εύκολα αντιληπτό λοιπόν, η χρήση των LLMs σε ρομπότ παροχής υπηρεσιών δεν περιορίζεται μόνο στην τεχνολογική πρόοδο, αλλά αγγίζει ευρύτερα κοινωνικά και οικονομικά ζητήματα.

Η παρούσα εργασία διερευνά τη συμβολή των LLMs στη βελτίωση των αυτόνομων ρομποτικών οχημάτων και τις προκλήσεις που ανακύπτουν από την ενσωμάτωσή τους. Παρουσιάζονται οι εφαρμογές των ρομπότ παροχής υπηρεσιών στον τομέα της υγειονομικής περίθαλψης, αναδεικνύοντας την πρακτική τους αξία. Παράλληλα, εξετάζονται τα ηθικά διλήμματα και οι πιθανές επιπτώσεις αυτών των τεχνολογιών, θέτοντας τις βάσεις για περαιτέρω έρευνα και ανάπτυξη στον τομέα. Για την καλύτερη κατανόηση της τεχνολογίας και των εφαρμογών της, είναι ωφέλιμο να ξεκαθαριστούν ορισμένες βασικές έννοιες που αποτελούν το θεμέλιο της εργασίας.



### Αυτόνομο Ρομπότ

Ως αυτόνομο ρομπότ ορίζεται ένα σύστημα που δύναται να εκτελεί εργασίες και να λειτουργεί σε ένα περιβάλλον χωρίς ανθρώπινο έλεγχο ή παρέμβαση. Αυτά τα ρομπότ χρησιμοποιούν αισθητήρες για την αντίληψη του περιβάλλοντος, λαμβάνουν αποφάσεις μέσω προηγμένων αλγορίθμων και εκτελούν ενέργειες μέσω μηχανισμών δράσης (actuators). Τα αυτόνομα ρομπότ χρησιμοποιούν προηγμένα συστήματα ελέγχου και αλγόριθμους που τους επιτρέπουν να κινούνται, να ερμηνεύουν δεδομένα και να εκτελούν εργασίες σε δυναμικά και απρόβλεπτα περιβάλλοντα, όπως η βιομηχανία, η υγεία και η εφοδιαστική αλυσίδα [4].

### Τεχνητή Νοημοσύνη (AI – Artificial Intelligence)

Η τεχνητή νοημοσύνη είναι η τεχνολογία που επιτρέπει στους υπολογιστές και τις μηχανές να προσομοιώνουν την ανθρώπινη μάθηση, κατανόηση, επίλυση προβλημάτων, λήψη αποφάσεων, δημιουργικότητα και αυτονομία [5].

### Μεγάλα Γλωσσικά Μοντέλα (LLMs)

Ένα μεγάλο γλωσσικό μοντέλο (LLM - Large Language Model) είναι ένας τύπος προγράμματος τεχνητής νοημοσύνης (AI) που έχει ως στόχο την κατανόηση και τη δημιουργία φυσικής γλώσσας [6]. Τα LLMs λειτουργούν αναλύοντας τεράστια σύνολα γλωσσικών δεδομένων και βασίζονται σε έναν τύπο μηχανικής μάθησης (Machine Learning), που ονομάζεται βαθιά μάθηση (Deep Learning) και περιλαμβάνει την πιθανολογική ανάλυση μη δομημένων δεδομένων. Με απλούστερους όρους, ένα LLM είναι ένα πρόγραμμα υπολογιστή στο οποίο στο οποίο έχουν τροφοδοτηθεί αρκετά παραδείγματα ώστε να είναι σε θέση να αναγνωρίζει και να ερμηνεύει την ανθρώπινη γλώσσα ή άλλους τύπους σύνθετων δεδομένων.

## ΚΕΦΑΛΑΙΟ 1: Η ΤΕΧΝΟΛΟΓΙΑ ΠΙΣΩ ΑΠΟ ΤΑ LLMs

Όπως αναφέρθηκε και προηγουμένως τα μεγάλα γλωσσικά μοντέλα είναι μοντέλα βαθιάς μάθησης που έχουν ως στόχο την κατανόηση και τη δημιουργία φυσικής γλώσσας. Επί του παρόντος, βασίζονται κυρίως στην αρχιτεκτονική Μετασχηματιστή (Transformer) [7], σύμφωνα με την οποία, στρώματα προσοχής Πολλαπλών Κεφαλών (Multi-head Attention) στοιβάζονται σε ένα πολύ βαθύ νευρωνικό δίκτυο. Η ικανότητά αυτής της αρχιτεκτονικής να κλιμακώνεται αποτελεσματικά επιτρέπει την υλοποίηση της εκπαίδευσης των μοντέλων σε τεράστια σύνολα δεδομένων κειμένου. Έτσι προκύπτει και ο χαρακτηρισμός «μεγάλο» στα Μεγάλα Γλωσσικά Μοντέλα, ο οποίος αναφέρεται τόσο στο μέγεθος και τη πολυπλοκότητα του ίδιου του νευρωνικού δικτύου, όσο και στο μέγεθος του συνόλου δεδομένων στο οποίο εκπαιδεύτηκε. Για ορισμένα από αυτά τα μοντέλα γίνεται λόγος για τρισεκατομμύρια μονάδες κειμένου (tokens) από ένα σωρό δημόσια διαθέσιμες πηγές [8]. Αυτή είναι και η βασική διαφορά ανάμεσα στα Γλωσσικά Μοντέλα και τα Μεγάλα Γλωσσικά Μοντέλα: Η κλίμακα των παραμέτρων και η επεξεργαστική πολυπλοκότητα, και ως εκ τούτου η ακρίβεια και οι εφαρμογές.

Προκειμένου να κατανοηθεί πλήρως ο τρόπος με τον οποίο τα LLMs λειτουργούν και επεξεργάζονται τη φυσική γλώσσα, είναι χρήσιμο να επεξηγηθούν οι βασικές έννοιες και τεχνολογίες που τα θεμελιώνουν, όπως επίσης και να παρουσιαστεί μία λεπτομερής ανάλυση των πρωτοποριακών τεχνικών που τα απαρτίζουν. Στο κεφάλαιο αυτό λοιπόν, παρουσιάζονται οι θεμελιώδεις αρχές και τεχνολογίες που στηρίζουν τη λειτουργία των μεγάλων γλωσσικών μοντέλων, επεξηγείται η αρχιτεκτονική τους και αναλύονται οι μέθοδοι εκπαίδευσης των LLMs και η διάσταση των μεγάλων βάσεων δεδομένων.

### 1.1 Εισαγωγή στις Τεχνολογίες Βάσης

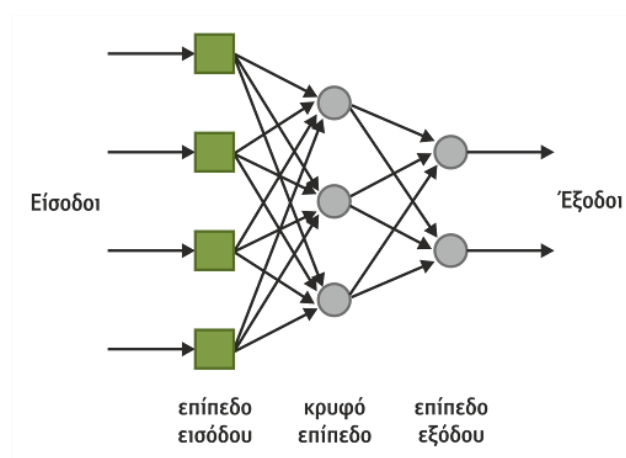
Τα LLMs χρησιμοποιούν έναν τύπο μηχανικής μάθησης που ονομάζεται βαθιά μάθηση, η οποία βασίζεται σε τεχνητά νευρωνικά δίκτυα για να αναγνωρίζει μοτίβα και να κάνει συσχετίσεις σε ακατέργαστα, μη δομημένα δεδομένα. Στη συνέχεια, γίνεται μία εισαγωγική ανάλυση των παραπάνω εννοιών.



Εικόνα 1.1: Η ιεραρχική σχέση ανάμεσα στην Τεχνητή Νοημοσύνη, τη Μηχανική Μάθηση, τα Νευρωνικά Δίκτυα και τη Βαθιά Μάθηση.

### 1.1.1 Νευρωνικά Δίκτυα (Neural Networks)

Τα νευρωνικά δίκτυα ή τεχνικά νευρωνικά δίκτυα είναι συστήματα που αποτελούνται από ένα μαζικά συνδεδεμένο δίκτυο υπολογιστικών «νευρώνων», οργανωμένων σε επίπεδα [9]. Πρόκειται για έναν τύπο υπολογιστικής αρχιτεκτονικής που βασίζεται σε ένα μοντέλο του τρόπου λειτουργίας του ανθρώπινου εγκεφάλου, χρησιμοποιώντας διαδικασίες που μιμούνται τον τρόπο με τον οποίο οι βιολογικοί νευρώνες συνεργάζονται για την λήψη αποφάσεων. Κατά αυτό τον τρόπο προκύπτει και η ονομασία «νευρωνικό».



Εικόνα 1.2: Διαγραμματική αναπαράσταση ενός τεχνητού νευρωνικού δικτύου με ένα κρυφό επίπεδο [136].

Κάθε νευρωνικό δίκτυο αποτελείται από στρώματα κόμβων ή τεχνητών νευρώνων, τα οποία περιλαμβάνουν ένα στρώμα εισόδου, ένα ή περισσότερα κρυφά στρώματα και ένα στρώμα εξόδου. Ως κρυφά στρώματα αναφέρονται οι κόμβοι μεταξύ των στρωμάτων εισόδου και εξόδου ενός νευρωνικού δικτύου. Κάθε κόμβος συνδέεται με άλλους κόμβους και έχει το δικό του σχετικό βάρος και κατώφλι. Ένας κόμβος ενεργοποιείται εφόσον η έξοδος του είναι πάνω από την καθορισμένη τιμή κατωφλίου και προβαίνει στην αποστολή δεδομένων στο επόμενο στρώμα του δικτύου. Σε άλλη περίπτωση, δεν διαβιβάζονται δεδομένα στο επόμενο επίπεδο του δικτύου [10].

Οι κόμβοι έχουν τη δυνατότητα να αλληλοεπιδρούν σχεδόν με κάθε τρόπο, με αποτέλεσμα οι κατηγορίες των τύπων νευρωνικών δικτύων να επεκτείνεται συνεχώς. Παρ' όλα αυτά, είναι δυνατό να ταξινομηθούν γενικά σε ρηγά νευρωνικά δίκτυα και βαθιά νευρωνικά δίκτυα. Τα ρηγά

νευρωνικά δίκτυα έχουν συνήθως μόνο ένα κρυφό επίπεδο, είναι γρήγορα και απαιτούν λιγότερη ισχύ. Αντίθετα, τα βαθιά νευρωνικά δίκτυα έχουν πολλαπλά κρυφά στρώματα και μπορούν να εκτελέσουν εργασίες με μεγαλύτερη πολυπλοκότητα .

Τα νευρωνικά δίκτυα χρησιμοποιούν δεδομένα εκπαίδευσης για να βελτιώνονται και να αυξάνουν την ακρίβειά τους με την πάροδο του χρόνου. Με τη σωστή ρύθμιση, αποτελούν εξαιρετικά αποτελεσματικά εργαλεία που μπορούν να ταξινομούν και να οργανώνουν δεδομένα πολύ γρήγορα. Έτσι, εργασίες όπως η αναγνώριση ομιλίας ή εικόνας μπορούν να διαρκέσουν λίγα λεπτά αντί για ώρες που θα απαιτούσε η χειροκίνητη αναγνώριση από ανθρώπινους εμπειρογνώμονες. Ο αλγόριθμος αναζήτησης της Google αποτελεί ένα από τα πιο γνωστά παραδείγματα ενός νευρωνικού δικτύου είναι [11].

### 1.1.2 Μηχανική μάθηση (Machine Learning)

Η Μηχανική Μάθηση αποτελεί μία εφαρμογή της τεχνητής νοημοσύνης κατά την οποία οι μηχανές δεν προγραμματίζονται ρητά για να εκτελούν συγκεκριμένες εργασίες, αλλά μαθαίνουν και βελτιώνονται από την εμπειρία τους αυτόματα [9]. Συγκεκριμένα, αναφέρεται σε έναν τύπο στατιστικού αλγορίθμου που μπορεί να μάθει χωρίς συγκεκριμένες οδηγίες, μέσω της τροφοδότησης του προγράμματος με μεγάλο όγκο δεδομένων. Με αυτόν τον τρόπο επιτυγχάνεται η εκπαίδευσή του, ώστε να δύναται να εκτελεί ορισμένες εργασίες χωρίς ανθρώπινη παρέμβαση, γενικεύοντας από παραδείγματα.

Η μηχανική μάθηση μπορεί να χωριστεί σε τρεις τύπους αλγορίθμων: τη Μάθηση με Επίβλεψη (Supervised Learning), τη Μάθηση χωρίς Επίβλεψη (Unsupervised Learning) και την Ενισχυτική Μάθηση (Reinforcement Learning). Στη μάθηση με επίβλεψη, ο υπολογιστής εκπαιδεύεται μέσω παραδειγμάτων, λαμβάνοντας οδηγίες για το τι αντιπροσωπεύει κάθε παράδειγμα, ενώ κατά τη μάθηση χωρίς επίβλεψη, η μηχανή αναζητά μοτίβα ή ομάδες δεδομένων χωρίς ετικέτες. Τέλος, σύμφωνα με την ενισχυτική μάθηση, το μοντέλο μαθαίνει μέσω δοκιμής και λάθους, λαμβάνοντας ανταμοιβές για καλές προβλέψεις. Ο στόχος της μηχανής είναι να μάθει πώς να ενεργεί με τρόπο που μεγιστοποιεί τις μελλοντικές ανταμοιβές της ή ελαχιστοποιεί τις τιμωρίες κατά τη διάρκεια της ζωής της.



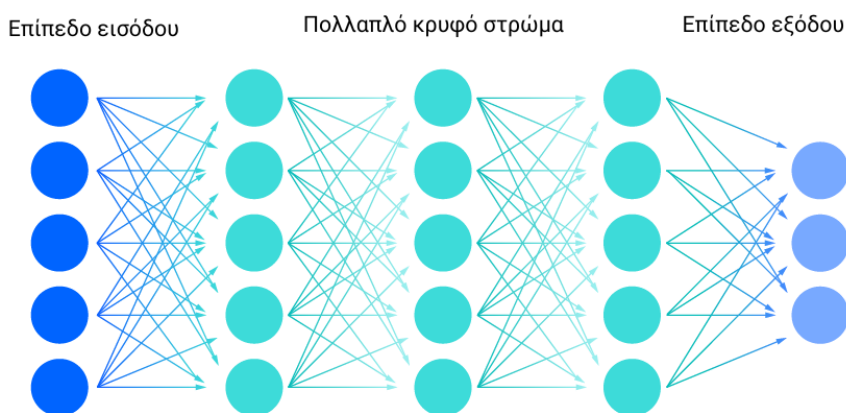
Εικόνα 1.3: Διαγραμματική απεικόνιση των τύπων της Μηχανικής Μάθησης.

Συνοψίζοντας, η μηχανική μάθηση βασίζεται σε εισόδους και εξόδους, καθώς ένας αλγόριθμος τροφοδοτείται με δεδομένα (είσοδος) τα οποία χρησιμοποιεί για να παράγει ένα αποτέλεσμα (έξοδος) [5]. Ένα σημαντικό πλεονέκτημα της μηχανικής μάθησης είναι η ικανότητα του υπολογιστή να αλλάζει τον τρόπο με τον οποίο προβλέπει τα πράγματα, καθώς οι πληροφορίες που του δίνονται αλλάζουν, μαθαίνοντας έτσι να βρίσκει νέα μοτίβα. Στο πεδίο της τεχνητής νοημοσύνης, η μηχανική μάθηση ασκεί σημαντική και ευρεία επίδραση σε διάφορους τομείς της τεχνολογίας και της επιστήμης [12].

### 1.1.3 Βαθιά Μάθηση (Deep Learning)

Η βαθιά μάθηση είναι ένας κλάδος της μηχανικής μάθησης που βασίζεται στην αρχιτεκτονική των τεχνητών νευρωνικών δικτύων για την αναγνώριση σύνθετων μοτίβων [9]. Όπως αναφέρθηκε προηγουμένως, ένα νευρωνικό δίκτυο αποτελείται από διασυνδεδεμένους κόμβους σε μια δομή με πολλά επίπεδα που συσχετίζουν τις εισόδους με τις επιθυμητές εξόδους και οι κόμβοι μεταξύ των στρώματων εισόδου και εξόδου αποτελούν τα κρυφά στρώματα. Ο όρος «βαθύς» αναφέρεται συνήθως στον αριθμό των κρυφών στρώματων του νευρωνικού δικτύου, καθώς τα μοντέλα βαθιάς μάθησης μπορεί να έχουν εκατοντάδες ή και χιλιάδες κρυφά στρώματα [13], επιτυγχάνοντας κορυφαία ακρίβεια σε προγνωστική ανάλυση.

#### Βαθύ Νευρωνικό Δίκτυο



Εικόνα 1.4: Διάγραμμα ενός Βαθιού Νευρωνικού Δικτύου που περιλαμβάνει ένα επίπεδο εισόδου, πολλαπλά κρυφά επίπεδα και ένα επίπεδο εξόδου [10].

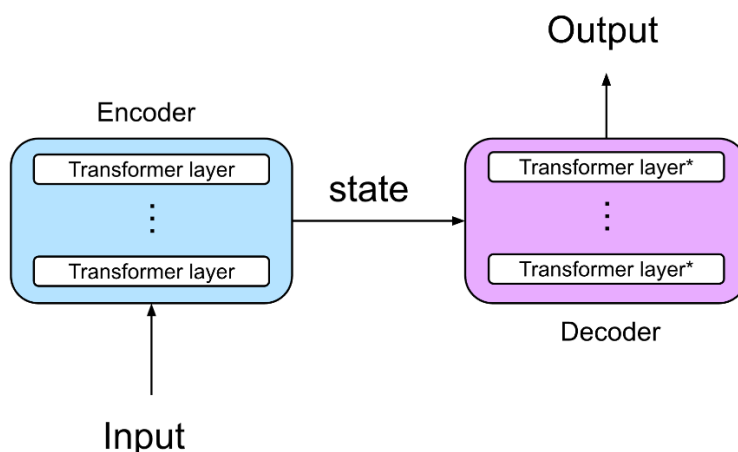
Οι αλγόριθμοι βαθιάς μάθησης εκπαιδεύονται σε μεγάλα σύνολα δεδομένων με ετικέτες, μαθαίνοντας να συσχετίζουν τα χαρακτηριστικά των δεδομένων με τις σωστές ετικέτες, ενώ κάθε κόμβος είναι υπεύθυνος για τη μάθηση ενός συγκεκριμένου χαρακτηριστικού των δεδομένων. Για παράδειγμα, σε μια εργασία αναγνώρισης εικόνας, ο αλγόριθμος μπορεί να μάθει να συσχετίζει ορισμένα χαρακτηριστικά με τη σωστή ετικέτα ως εξής: Το πρώτο στρώμα κόμβων μπορεί να μάθει να αναγνωρίζει ακμές, το δεύτερο στρώμα σχήματα και το τρίτο στρώμα μπορεί να μάθει να αναγνωρίζει χρώματα [14].

Ουσιαστικά, ένα μοντέλο βαθιάς μάθησης είναι ένα πρόγραμμα υπολογιστή που μπορεί να επιδείξει νοημοσύνη, χάρη στην πολύπλοκη και εξελιγμένη προσέγγισή του στην επεξεργασία δεδομένων. Αποτελεί τον πυρήνα της λειτουργίας πολλών μοντέλων τεχνητής νοημοσύνης, γύρω από τον οποίο πλαισιώνονται και τα μεγάλα γλωσσικά μοντέλα ώστε να μπορούν να μάθουν τη γλώσσα, να κατανοήσουν το πλαίσιο και να παράγουν απαντήσεις. Φυσικά, οι ικανότητές τους δεν περιορίζονται εκεί, καθώς τα LLMs με σύμμαχο τη βαθιά μάθηση παρουσιάζουν ένα εύρος δυνατοτήτων που κυμαίνεται από την αναγνώριση αντικειμένων σε μια φωτογραφία ή την αναγνώριση μιας φωνής μέχρι την οδήγηση ενός αυτοκινήτου.

## 1.2 Η Αρχιτεκτονική των LLMs

Η αρχιτεκτονική των LLMs λειτουργεί λαμβάνοντας δεδομένα κειμένου από πολλαπλές πηγές και στη συνέχεια προωθώντας τα στο επόμενο στάδιο για προ επεξεργασία. Έπειτα η διαδικασία εκπαίδευσης ολοκληρώνεται εκτελώντας μια σειρά από στάδια, όπως αρχικοποίηση τυχαίων παραμέτρων, εισαγωγή αριθμητικών δεδομένων, απώλεια υπολογισμού συνάρτησης, βελτιστοποίηση παραμέτρων και επαναληπτική εκπαίδευση [15]. Βασίζεται στην αρχιτεκτονική του Μετασχηματιστή (Transformer), ο οποίος θεωρείται ως το βασικό δομικό στοιχείο των LLMs, καθώς η ανάπτυξη της αρχιτεκτονικής αυτής έφερε επανάσταση στη μοντελοποίηση γλωσσών. Ο μετασχηματιστής είναι στην ουσία ένα μοντέλο βαθιάς μάθησης, βασιζόμενο σε έναν μηχανισμό προσοχής για την επεξεργασία δεδομένων ακολουθίας. Η ικανότητα του για παράλληλους υπολογισμούς καθώς και η πολυπλοκότητα του ίδιου του μοντέλου, του προσδίδουν μεγαλύτερη ακρίβεια και απόδοση, σε σύγκριση με τα προηγουμένως δημοφιλή επαναλαμβανόμενα νευρωνικά δίκτυα. Πράγματι, δύναται να επιλύσει αποτελεσματικά εξαιρετικά σύνθετα προβλήματα επεξεργασίας φυσικής γλώσσας. Η αρχιτεκτονική του Μετασχηματιστή αποτελείται κυρίως από δύο τμήματα, έναν Κωδικοποιητή (Encoder) και έναν Αποκωδικοποιητή (Decoder), καθώς και από τον Μηχανισμό Προσοχής, ο οποίος εφαρμόζεται μέσα σε κάθε ένα από αυτά τα τμήματα [16].

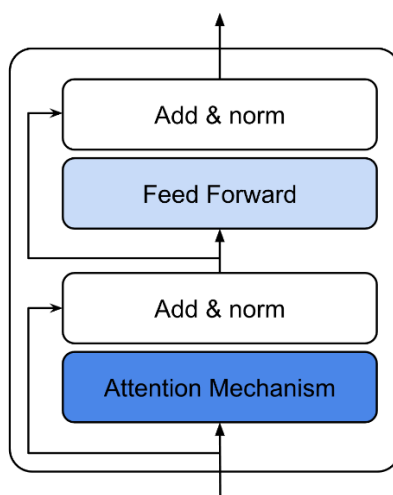
Ο Κωδικοποιητής δέχεται μια ακολουθία tokens (λεκτικών μονάδων) και παράγει μια διανυσματική αναπαράσταση σταθερού μεγέθους ολόκληρης της ακολουθίας, η οποία μπορεί στη συνέχεια να χρησιμοποιηθεί για ταξινόμηση. Ασχολείται μόνο με την ακολουθία εισόδου, χωρίς να παράγει καμία ακολουθία εξόδου. Ο Αποκωδικοποιητής λαμβάνει μια διανυσματική αναπαράσταση του πλαισίου σταθερού μεγέθους και τη χρησιμοποιεί για να παράγει μια ακολουθία λέξεων μία προς μία, με κάθε λέξη να εξαρτάται από τις προηγούμενες λέξεις που έχουν παραχθεί [17]. Η αρχιτεκτονική Κωδικοποιητή-Αποκωδικοποιητή (Transformer Encoder-Decoder) χρησιμοποιείται για εργασίες όπως η γλωσσική μετάφραση, όπου το μοντέλο πρέπει να λαμβάνει μια πρόταση σε μια γλώσσα και να εξάγει μια πρόταση σε μια άλλη. Στο συγκεκριμένο μοντέλο αρχιτεκτονικής ο Κωδικοποιητής δέχεται την πρόταση εισόδου και παράγει μια διανυσματική αναπαράστασή της σταθερού μεγέθους, η οποία στη συνέχεια τροφοδοτείται στον αποκωδικοποιητή για τη δημιουργία της πρότασης εξόδου [1]. Ένα από τα πιο δημοφιλή μοντέλα Κωδικοποιητή-Αποκωδικοποιητή Μετασχηματιστή είναι ο T5, ο οποίος συχνά χρησιμοποιείται σε εργασίες μετάφρασης.



Εικόνα 1.5: Αρχιτεκτονική Κωδικοποιητής-Αποκωδικοποιητής.

Η ακολουθία εισόδου (Input) κωδικοποιείται πρώτα σε ένα διάνυσμα κατάστασης (State), το οποίο στη συνέχεια χρησιμοποιείται για την αποκωδικοποίηση της ακολουθίας εξόδου (Output) [18].

Ο Μηχανισμό Προσοχής (Attention Mechanism) επιτρέπει στο μοντέλο να εστιάζει σε συγκεκριμένα τμήματα της ακολουθίας [19], ώστε να κατανοεί καλύτερα το πλαίσιο και τις σχέσεις των λέξεων μεταξύ τους. Στοχεύει στην επιλογή μιας μικρής ποσότητας σημαντικών πληροφοριών από έναν μεγάλο όγκο δεδομένων, καθώς εστιάζοντας σε αυτά τα σημαντικά κομμάτια αγνοεί την πλειονότητα των ασήμαντων πληροφοριών [16]. Συγκεκριμένα, το μοντέλο του μετασχηματιστή είναι χτισμένο γύρω από τον Μηχανισμό Αυτοπροσοχής, έναν από τους πιο σημαντικούς τύπους προσοχής [7], ο οποίος επιτρέπει σε κάθε λέξη να αλληλοεπιδρά με άλλες λέξεις της ίδιας ακολουθίας.



Εικόνα 1.6: Ενδεικτική αναπαράσταση μιας δομής αρχιτεκτονικής μεγάλων γλωσσικών μοντέλων. Περιλαμβάνει τους μηχανισμούς «Attention Mechanism» και «Feed Forward», οι οποίοι πλαισιώνονται από διαδικασίες κανονικοποίησης και προσθήκης («Add & Norm») [18].

Αναλυτικότερα, κατά την εκτέλεση του Μηχανισμού Αυτοπροσοχής (Self-Attention Mechanism), κάθε λέξη μιας πρότασης αναλύεται σε σχέση με τις υπόλοιπες λέξεις της ίδιας πρότασης, επιτρέποντας στο μοντέλο να υπολογίζει την αλληλεπίδραση κάθε λέξης με όλες τις άλλες λέξεις [7]. Έτσι, το μοντέλο αποκτά μια συνολική κατανόηση του νοήματος. Για την εφαρμογή αυτής της τεχνικής, κάθε λέξη της ακολουθίας μετατρέπεται σε τρία διαφορετικά διανύσματα: το ερώτημα (query), το κλειδί (key), και την τιμή (value). Αυτά τα διανύσματα είναι ουσιαστικά κωδικοποιήσεις που αντιπροσωπεύουν τις διαφορετικές σχέσεις της λέξης και παράγονται μέσω πολλαπλασιασμού της αναπαράστασης της λέξης με τρεις διαφορετικούς πίνακες βαρών.

Οι πίνακες βαρών είναι ουσιαστικά πίνακες παραμέτρων που προ εκπαιδεύονται, δηλαδή εκπαιδεύονται κατά τη διάρκεια της διαδικασίας εκπαίδευσης του μοντέλου και καθορίζουν πώς το μοντέλο «μετασχηματίζει» τα δεδομένα εισόδου ώστε να αναγνωρίζει σχέσεις μεταξύ λέξεων [20]. Πιο συγκεκριμένα η διαδικασία αφορά τρεις διαφορετικούς πίνακες, τον πίνακα ερωτήματος (query matrix), τον πίνακα κλειδιού (key matrix) και τον πίνακα τιμής (value matrix). Ο πίνακας ερωτήματος εφαρμόζεται σε κάθε λέξη, με σκοπό τη δημιουργία του διανύσματος ερωτήματός της και θέτει το ερώτημα γύρω από τη σχετικότητα της λέξης με τις υπόλοιπες λέξεις της ακολουθίας. Με τον πίνακα κλειδιού δημιουργείται το διάνυσμα κλειδιού της κάθε λέξης, το οποίο περιγράφει ποια πληροφορία προσδίδει η λέξη στην ακολουθία, ώστε να αντιστοιχιστεί με ένα ερώτημα από άλλες λέξεις. Τέλος, με την εφαρμογή του πίνακα τιμής, παράγεται τι διάνυσμα τιμής της κάθε λέξης, όπου περιέχεται η κύρια πληροφορία για την δημιουργία της τελικής αναπαράστασης της λέξης, καθοριζόμενη από τις συνδέσεις με άλλες λέξεις. Τα τρία αυτά διανύσματα πολλαπλασιάζονται ώστε να υπολογιστούν οι βαθμολογίες προσοχής της λέξης ως προς τις άλλες λέξεις στην ακολουθία [7]. Οι βαθμολογίες που προκύπτουν αποτελούν μια αναπαράσταση για κάθε λέξη, που περιλαμβάνει πληροφορίες από τις άλλες λέξεις της ακολουθίας, με βαρύτητα ανάλογη της συνάφειάς τους. Ως αποτέλεσμα, οι πίνακες βαρών εκπαιδεύονται με σκοπό να μαθαίνουν τις συνάφειες μεταξύ των λέξεων και να παρέχουν την κατάλληλη πληροφορία στο μοντέλο για την βέλτιστη κατανόηση της ακολουθίας.

Ο Μηχανισμός Αυτοπροσοχής αποτελεί το κύριο πλεονέκτημα του μετασχηματιστή [6], καθώς επιτρέπει στα μεγάλα γλωσσικά μοντέλα όχι μόνο να κατανοούν το πλήρες πλαίσιο μιας πρότασης αλλά και να εστιάζουν σε λέξεις που επηρεάζουν άμεσα το νόημα των υπολοίπων. Χάρη σε αυτή τη διαδικασία, οι Μετασχηματιστές μπορούν να μάθουν πολύπλοκες γλωσσικές δομές και ως εκ τούτου να παράγουν νοηματικά πλούσιες και ακριβείς απαντήσεις [21]. Ασφαλώς, τα LLMs με εργαλείο την Αυτοπροσοχή χρησιμοποιούνται με επιτυχία σε ποικίλες εργασίες, όπως η κατανόηση της ανάγνωσης, η αφηρημένη περίληψη, κειμενική συνεπαγωγή και εκμάθηση αναπαραστάσεων προτάσεων ανεξάρτητων από την εργασία.

Αναμφίβολα, η Αυτοπροσοχή Πολλαπλών Κεφαλών (Multi-Head Self-Attention) αποτελεί θεμελιώδες βάση της αρχιτεκτονικής του Μετασχηματιστή, καθώς αφορά μια τεχνική που επιτρέπει την παράλληλη επεξεργασία πολλαπλών προοπτικών προσοχής. Συγκεκριμένα, Ο Μηχανισμός Αυτοπροσοχής πολλαπλών κεφαλών χωρίζει το διάνυσμα της κάθε λέξης σε πολλά υπό-διανύσματα, ώστε κάθε κεφαλή να παίρνει ένα μέρος του αρχικού διανύσματος και να επεξεργάζεται διαφορετικές σχέσεις [15]. Για κάθε κεφαλή, εκτελείται ξεχωριστός υπολογισμός



της προσοχής μέσω των τριών διανυσμάτων *query*, *key*, και *value*, και έτσι κάθε κεφαλή εστιάζει σε διαφορετικές συνάφειες μεταξύ των λέξεων της ακολουθίας. Τα αποτελέσματα από όλες τις κεφαλές προσοχής ενοποιούνται σε ένα ενιαίο διάνυσμα, το οποίο αποτυπώνει μια πιο ολοκληρωμένη και σύνθετη αναπαράσταση των σχέσεων μεταξύ των λέξεων και έπειτα εισάγεται στα επόμενα επίπεδα του μοντέλου [7]. Αυτή η διαδικασία επιτρέπει στο μοντέλο να παρακολουθεί πληροφορίες από πολλές διαφορετικές προοπτικές, καθώς κάθε κεφαλή μπορεί να εντοπίζει διαφορετικού τύπου σχέσεις. Η καινοτομία του Μηχανισμού Αυτοπροσοχής Πολλαπλών Κεφαλών έγκειται στην ικανότητά του αυτή να αναγνωρίζοντας σύνθετες και πολυδιάστατες σχέσεις μεταξύ των λέξεων της ακολουθίας, γεγονός που βελτιώνει την κατανόηση της γλώσσας, ειδικά σε περίπλοκες προτάσεις.

### 1.3 Η Εκπαίδευση των LLMs

Η εκπαίδευση των μεγάλων γλωσσικών μοντέλων περιλαμβάνει τη χρήση τεράστιων συνόλων δεδομένων κειμένου και την εφαρμογή περίπλοκων αλγορίθμων νευρωνικών δικτύων ώστε να μπορέσει το μοντέλο να μάθει τη γλώσσα και να αποκτήσει τη δυνατότητα παραγωγής νοηματικών απαντήσεων. Αποτελεί μία πολύπλοκη διαδικασία που μπορεί να χωριστεί σε γενικές γραμμές στα παρακάτω στάδια: Το πρώτο βήμα περιλαμβάνει τη συλλογή των δεδομένων και έπειτα ακολουθεί η επεξεργασία τους. Το επόμενο στάδιο αφορά τη διαδικασία προ-εκπαίδευσης, η οποία περιλαμβάνει τον καθορισμό των αντίστοιχων εργασιών και τη χρήση κατάλληλων αλγορίθμων παράλληλης εκπαίδευσης. Τελικά, πραγματοποιείται η λεπτομερή ρύθμιση και ευθυγράμμιση [16]. Σε αυτή την ενότητα, γίνεται μια εισαγωγή στα σχετικά σύνολα δεδομένων εκπαίδευσης, την προετοιμασία και την προ-επεξεργασία δεδομένων, τις ειδικές μεθοδολογίες εκπαίδευσης, την αξιολόγηση του μοντέλου και τα συνήθως χρησιμοποιούμενα πλαίσια εκπαίδευσης των Μεγάλων Γλωσσικών Μοντέλων.

#### 1.3.1 Συλλογή Δεδομένων

Όπως αναφέρθηκε προηγουμένως, η εκπαίδευση των LLMs απαιτεί τεράστιες ποσότητες δεδομένων κειμένου και φυσικά η ποιότητά τους έχει σημαντική επίδραση στην απόδοση των Μεγάλων Γλωσσικών Μοντέλων. Τα σώματα κειμένων φυσικής γλώσσας αντλούνται από διάφορες πηγές. Οι πιο συνήθεις περιλαμβάνουν κείμενα από τον ιστό, δεδομένα συνομιλιών και βιβλία, ενώ ακόμη ορισμένες ερευνητικές προσπάθειες εισάγουν εξειδικευμένα δεδομένα από επαγγελματικούς τομείς, όπως κώδικες ή επιστημονικά δεδομένα, με σκοπό την ενίσχυση των δυνατοτήτων των LLMs σε αυτούς τους τομείς [16].

Συγκεκριμένα, το BookCorpus και το Gutenberg αποτελούν δύο ευρέως χρησιμοποιούμενα σύνολα δεδομένων βιβλίων, απαρτίζοντας βιβλιοθήκες 7000 και 70000 ηλεκτρονικών βιβλίων αντίστοιχα [22]. Βεβαίως, τα βιβλία είναι μια πλούσια πηγή τόσο φυσικής γλώσσας και λεξιλογίου όσο και λεπτομερών πληροφοριών, όπως η όψη ενός αντικειμένου και σημασιολογίας υψηλού επιπέδου, προσφέροντας μια πιο ολοκληρωμένη κατανόηση της γλώσσας σε διάφορους τομείς. Όσον αφορά την άντληση δεδομένων από τον ιστό, το CommonCrawl συνιστά βασική πηγή εκπαιδευτικού υλικού [23], καθώς διαχειρίζεται ένα αποθετήριο δεδομένων εντοπισμού ιστοσελίδων, διαθέσιμο ελεύθερα για χρήση από άτομα και οργανισμούς. Αποτελείται από μια τεράστια συλλογή που περιλαμβάνει περισσότερες από 250 δισεκατομμύρια ιστοσελίδες, και διευρύνεται συνεχώς με την προσθήκη 3-5 δισεκατομμυρίων νέων ιστοσελίδων κάθε μήνα.

Παρουσιάζει ιδιαίτερη σημασία στον κλάδο επεξεργασίας της φυσικής γλώσσας, όπου λειτουργεί ως βασικό εκπαιδευτικό σύνολο για πολλά μεγάλα γλωσσικά μοντέλα [16]. Σε αυτό το πλαίσιο, η Βικιπαίδεια (Wikipedia), η οποία συνιστά ένα τεράστιο αποθετήριο εγκυκλοπαιδικού περιεχομένου υψηλής ποιότητας, καλύπτοντας ένα ευρύ φάσμα θεμάτων, χρησιμοποιείται εκτενώς στην εκπαίδευση των LLMs [24]. Καθώς η ποιότητα των δεδομένων επηρεάζει σε μεγάλο βαθμό την απόδοση των Μεγάλων Γλωσσικών Μοντέλων, τα συγκεκριμένα βρήκαν σημαντική εφαρμογή σε εργασίες γλωσσικής κατανόησης και παραγωγής. Αναφορικά με τα δεδομένα συνομιλιών, τα κοινωνικά δίκτυα αποτελούν αναμφισβήτητα έναν αντιπροσωπευτικό πόρο καθημερινών συζητήσεων και ανθρώπινης αλληλεπίδρασης και ειδικότερα, οι σύνδεσμοι Reddit είναι οι πλέον κατάλληλες πηγές τέτοιων δειγμάτων εκπαίδευσης. Πρόκειται για μία αμερικανική ιστοσελίδα συγκέντρωσης ειδήσεων, περιεχομένου και συζήτησης, όπου τα εγγεγραμμένα μέλη υποβάλλουν περιεχόμενο στον ιστότοπο, όπως εικόνες, συνδέσμους, κείμενα, τα οποία συζητούνται με άλλους χρήστες. Οι χρήστες έχουν τη δυνατότητα να ψηφίσουν υπέρ ή κατά των δημοσιεύσεων, τα οποία καλύπτουν μια ποικιλία θεμάτων, γεγονός που αυξάνει την ποιότητα των αναρτήσεων και ως εκ τούτου των δεδομένων εκπαίδευσης [25]. Τέλος, για την εκπαίδευση των LLMs χρησιμοποιούνται κομμάτια κώδικα, με τις κύριες πηγές να περιλαμβάνουν το Github και το Stack Overflow. Τροχοπέδη στην άντληση αυτών των δεδομένων αποτελεί η περιορισμένη διαθεσιμότητα δημοσίως προσβάσιμων συνόλων δεδομένων κώδικα [16].

Όπως γίνεται αντιληπτό λοιπόν, υπάρχουν διαθέσιμοι διάφοροι τύποι προ-εκπαιδευμένων συνόλων δεδομένων. Σαφώς, η αξιοποίηση ποικίλων πηγών δεδομένων κειμένου για την εκπαίδευση ενός LLM μπορεί να ενισχύσει σημαντικά τις ικανότητες γενίκευσης του μοντέλου [26], ενώ θεωρητικά, όσο περισσότερο μαθαίνει ένα γλωσσικό μοντέλο, τόσο περισσότερες παραμέτρους αποκτά, γεγονός που του επιτρέπει κατανοεί τη φυσική γλώσσα σε βάθος και να αναπτύσσει ισχυρές δυνατότητες [27]. Η θεωρία αυτή επιβεβαιώνεται και από έρευνες που έχουν δείξει ότι ένα πλουσιότερο σύνολο δεδομένων για την εκμάθηση γλωσσικών μοντέλων μπορεί να οδηγήσει σε σωστές απαντήσεις ποικίλων ερωτήσεων [26]. Τα σύνολα δεδομένων μπορούν να κατηγοριοποιηθούν σύμφωνα με είδος της γνώσης που παρέχουν στους εξής διάφορους τύπους: Τα δεδομένα βασικής σημασιολογικής γνώσης, που εστιάζουν στην κατανόηση του γλωσσικού νοήματος, τα δεδομένα κοινής λογικής, που περιγράφουν γενικές πληροφορίες και καθημερινές αλήθειες και τα δεδομένα γνώσης επαγγελματικού πεδίου, που αφορούν τεχνικές και επιστημονικές πληροφορίες [3].

### 1.3.2 Επεξεργασία Δεδομένων

Μόλις συλλεχθεί ένα επαρκές σώμα δεδομένων, ακολουθεί η διαδικασία επεξεργασίας, η οποία συμβάλλει στην διασφάλιση της ποιότητάς τους και έχει άμεση επίδραση στην απόδοση και την ασφάλεια του μοντέλου. Η επεξεργασία περιλαμβάνει δύο βασικά στάδια: το φιλτράρισμα και καθαρισμό δεδομένων, όπου απομακρύνεται τοξικό ή χαμηλής ποιότητας περιεχόμενο, και την προετοιμασία δεδομένων για την προσαρμογή τους στο μοντέλο, όπου τα δεδομένα μετατρέπονται σε μορφή κατάλληλη για να κατανοηθούν από το μοντέλο.

Κατά το στάδιο φιλτραρίσματος και καθαρισμού, πραγματοποιούνται οι παρακάτω διαδικασίες με σκοπό την εξασφάλιση της ποιότητας και της ουδετερότητας των δεδομένων, ώστε αυτά να ευθυγραμμίζονται με τα ανθρώπινα ηθικά πρότυπα [16]. Αρχικά, πραγματοποιείται μία

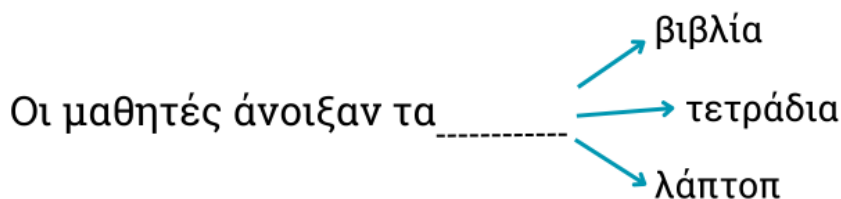
διαλογή ανάμεσα στα αξιόπιστα και μη δεδομένα, απομακρύνοντας κείμενα χαμηλής ποιότητας ή ακατάλληλα για το μοντέλο, όπως είναι οι διαφημίσεις και οι σπασμένες προτάσεις. Οι κύριες τεχνικές που χρησιμοποιούνται για την υλοποίηση αυτής της διαδικασίας βασίζονται σε Ευρετικές Μεθόδους (Heuristic-Based Methods) και Ταξινομητές (Classifier-Based Methods) [28]. Οι Ευρετικές Μέθοδοι στηρίζονται σε χειροκίνητα καθορισμένους κανόνες για την ανίχνευση δεδομένων χαμηλής ποιότητας. Για παράδειγμα οι κανόνες αυτοί μπορεί να αφορούν την απομάκρυνση κειμένων με ασυνήθιστη δομή ή χαμηλή αναγνωσιμότητα. Εφαρμόζοντας μεθόδους με Ταξινομητές, γίνεται χρήση μοντέλων μηχανικής μάθησης που έχουν εκπαιδευτεί να εντοπίζουν δεδομένα χαμηλής ποιότητας [29]. Λειτουργούν βασιζόμενα σε χαρακτηριστικά σχετικά με την ποιότητα, όπως είναι η σύνταξη, η πληρότητα και τα τυπικά σφάλματα, και αποφασίζουν αυτόματα ποια δεδομένα δεν πληρούν τα ποιοτικά κριτήρια. Εξίσου σημαντικό παρουσιάζεται και το φιλτράρισμα τοξικού και μεροληπτικού περιεχομένου, όπου εφαρμόζονται ισχυρές τεχνικές μετριασμού περιεχομένου, όπως η χρήση αλγορίθμων ανάλυσης συναισθήματος, ανίχνευσης ρητορικής μίσους και αναγνώρισης προκατάληψης. Στο πλαίσιο του καθαρισμού των δεδομένων, ιδιαίτερα σημαντικό παρουσιάζεται και το στάδιο αποδελτίωσης, που αφορά την επιλογή χρήσιμων κειμένων και την απομάκρυνση αντιγράφων, περιττών και επαναλαμβανόμενων δεδομένων στο σύνολο εκπαίδευσης [1]. Τέλος, η εκκαθάριση απορρήτου κρίνεται ζωτικής σημασίας, ώστε να αντιμετωπιστούν οι ανησυχίες που σχετίζονται με την ιδιωτικότητα, με τη συστηματική αφαίρεση τυχόν ευαίσθητων πληροφοριών [19]. Η προστασία των προσωπικών δεδομένων εξασφαλίζεται με τη χρήση τεχνικών όπως η ανωνυμοποίηση, και η τεχνική Tokenization (που θα αναλυθεί παρακάτω) για την εξάλειψη των προσωπικών στοιχείων, της γεωγραφικής θέσης και άλλων εμπιστευτικών δεδομένων [16].

Αφού τα δεδομένα καθαριστούν και φιλτραριστούν, περνούν στη φάση της προετοιμασίας, όπου υποβάλλονται σε μια σειρά διαδικασιών, ώστε να γίνουν κατανοητά από το γλωσσικό μοντέλο. Αρχικά εφαρμόζεται μία μέθοδος απομάκρυνσης ειδικών χαρακτήρων και συμβόλων, emojis, περιττών κενών ή άλλων ανεπιθύμητων στοιχείων που δεν εξυπηρετούν το μοντέλο. Η διαδικασία της Κανονικοποίησης (Data Normalization) και του καθαρισμού των δεδομένων κρίνεται απαραίτητη για την προσαρμογή τους στην κατάλληλη μορφή [19]. Περιλαμβάνει επίσης την ενοποίηση των μορφολογιών λέξεων, όπως πεζά και κεφαλαία, τονισμοί ή γραμματικές διαφορές και την αφαίρεση τυπογραφικών σφαλμάτων, ώστε να υπάρχει συνέπεια. Ένα ακόμη κρίσιμο στάδιο στην διαδικασία επεξεργασίας είναι η διαδικασία Tokenization, δηλαδή η διαδικασία κατά την οποία το κείμενο διασπάται σε διακριτά μέρη, τα οποία ονομάζονται tokens [30]. Αυτά τα tokens μπορούν να είναι λέξεις, υπολέξεις, σύμβολα ή χαρακτήρες, ανάλογα με τη μέθοδο που χρησιμοποιείται. Μέσω αυτής της τμηματοποίησης το μοντέλο δύναται να επεξεργάζεται και να κατανοεί το κείμενο με έναν τρόπο συμβατό με τα δεδομένα του και τη μορφή που μπορεί να αναλύσει, ενώ ταυτόχρονα ευθυγραμμίζεται με τη γλωσσική γνώση του ανθρώπου [1]. Το τελευταίο βήμα της φάσης προετοιμασίας περιλαμβάνει την μετατροπή των tokens σε αριθμητικά διανύσματα, συνήθως μέσω τεχνικών embeddings [31]. Τα embeddings είναι διανύσματα που αναπαριστούν τα tokens σε ένα αριθμητικό χώρο, και με αυτόν τον τρόπο το μοντέλο δύναται να κατανοήσει τη σημασιολογία των λέξεων σε αριθμητική μορφή [27].

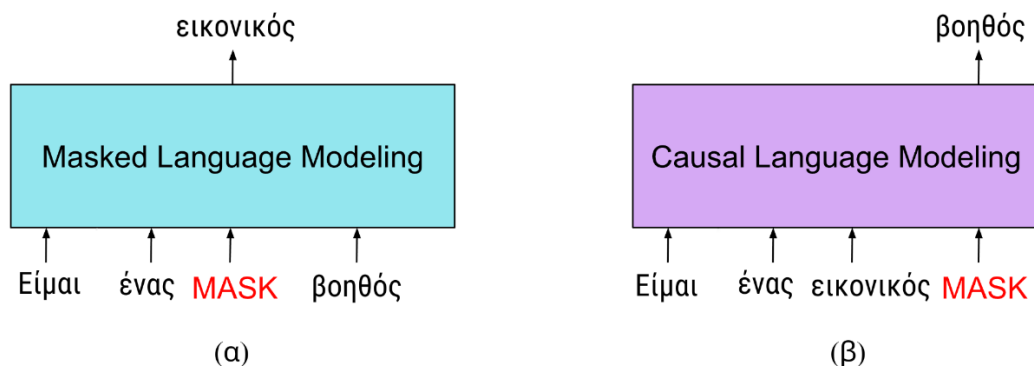
### 1.3.3 Προ-εκπαίδευση (Pre-training)

Αφού ολοκληρωθεί η διαδικασία συλλογής και επεξεργασίας των δεδομένων ακολουθεί το στάδιο της προ-εκπαίδευσης του συστήματος, όπου τα μεγάλα γλωσσικά μοντέλα μαθαίνουν πλούσιες γλωσσικές αναπαραστάσεις. Το γλωσσικό μοντέλο BERT, που αναπτύχθηκε από την Google, εισήγαγε την έννοια της προ-εκπαίδευσης σε συνδυασμό με τη Λεπτομερή Ρύθμιση (Fine-Tuning) για εργασίες γλωσσικής κατανόησης, επιτυγχάνοντας εξαιρετικά αποτελέσματα σε εργασίες όπως η απάντηση ερωτήσεων και η ταξινόμηση κειμένων [3]. Πολλοί αποδίδουν την ικανότητα συλλογισμού των μεγάλων γλωσσικών μοντέλων στην επίδραση των πρωτοπόρων αυτών μηχανισμών [32].

Στο αρχικό στάδιο προ εκπαίδευσης, τα LLMs χρησιμοποιούν μια προσέγγιση μάθησης με αυτό-επίβλεψη (Self-Supervised Learning), που τους επιτρέπει να μαθαίνουν από μεγάλες ποσότητες δεδομένων χωρίς σχολιασμό, δηλαδή χωρίς να απαιτείται χειροκίνητη προσθήκη ετικετών από ανθρώπους [6]. Η μέθοδος αυτή ονομάζεται Γλωσσική Μοντελοποίηση (Language Modeling) και είναι μια συνηθισμένη μορφή εργασίας αυτό-επιβλεπόμενης μάθησης, στην οποία το μοντέλο προβλέπει την επόμενη λέξη σε ένα δεδομένο πλαίσιο. Μέσω αυτής της εργασίας, το μοντέλο αποκτά την ικανότητα να συλλαμβάνει πληροφορίες που σχετίζονται με το λεξιλόγιο, τη γραμματική, τη σημασιολογία και τη δομή του κειμένου [16]. Αποτελεί την κεντρική εργασία για τη δημιουργία γλωσσικών μοντέλων, όπως τα GPT, BERT και άλλα που βασίζονται σε δεδομένα κειμένου [33]. Η γλωσσική μοντελοποίηση στην διαδικασία προ-εκπαίδευσης χωρίζεται σε τρεις κατηγορίες: την Αυτοπαλινδρομική Γλωσσική Μοντελοποίηση (Autoregressive ή Causal Language Modeling), την Μοντελοποίηση Γλώσσας με Μάσκα (Masked Language Modeling) και την Μοντελοποίηση Κωδικοποιητή-Αποκωδικοποιητή (Encoder-Decoder Modeling) [15]. Η Αυτοπαλινδρομική Μοντελοποίηση είναι η επικρατέστερη μέθοδος προ-εκπαίδευσης του GPT και των περισσότερων LLM. Η λειτουργία της περιλαμβάνει την εκπαίδευση του μοντέλου, ώστε να προβλέπει την πιθανότητα μιας λέξης με βάση τις προηγούμενες ή τις γειτονικές. Στην Μοντελοποίηση Γλώσσας με Μάσκα, το μοντέλο μαθαίνει να προβλέπει λέξεις που έχουν καλυφθεί μέσα σε μια πρόταση, επιτρέποντας τη χρήση και των δύο κατευθύνσεων του κειμένου (πριν και μετά τη λέξη), κάτι που κάνει τα μοντέλα όπως το BERT ιδιαίτερα αποτελεσματικά για εργασίες κατανόησης [16]. Στην τεχνική Κωδικοποιητή-Αποκωδικοποιητή, το μοντέλο μαθαίνει να μετατρέπει μια ακολουθία εισόδου σε μια ακολουθία εξόδου και συχνά χρησιμοποιείται σε εργασίες μετάφρασης ή περίληψης. Ολοκληρώνοντας τη διαδικασία πρόβλεψης και μετατροπής, στο τελικό στάδιο των Γλωσσικών Μοντελοποιήσεων εφαρμόζεται ένας αλγόριθμος βελτιστοποίησης για την ελαχιστοποίηση της διαφοράς μεταξύ των προβλέψεων του μοντέλου και των πραγματικών δεδομένων [34]. Συνοψίζοντας, η προ-εκπαίδευση είναι το βήμα όπου το μοντέλο αποκτά μια ευρεία και γενική κατανόηση της γλώσσας, η οποία του επιτρέπει να προσαρμόζεται εύκολα σε εξειδικευμένες εργασίες στο επόμενο στάδιο.



Εικόνα 1.7: Παράδειγμα εργασίας γλωσσικής μοντελοποίησης.



Εικόνα 1.8: Προσεγγίσεις γλωσσικής μοντελοποίησης.

(α) Μοντελοποίηση γλώσσας με μάσκα, όπου προβλέπει κρυμμένες λέξεις στην ακολουθία. (β) Αυτοπαλινδρομική Γλωσσική Μοντελοποίηση, όπου γίνεται πρόβλεψη της επόμενης λέξης στην ακολουθία [18].

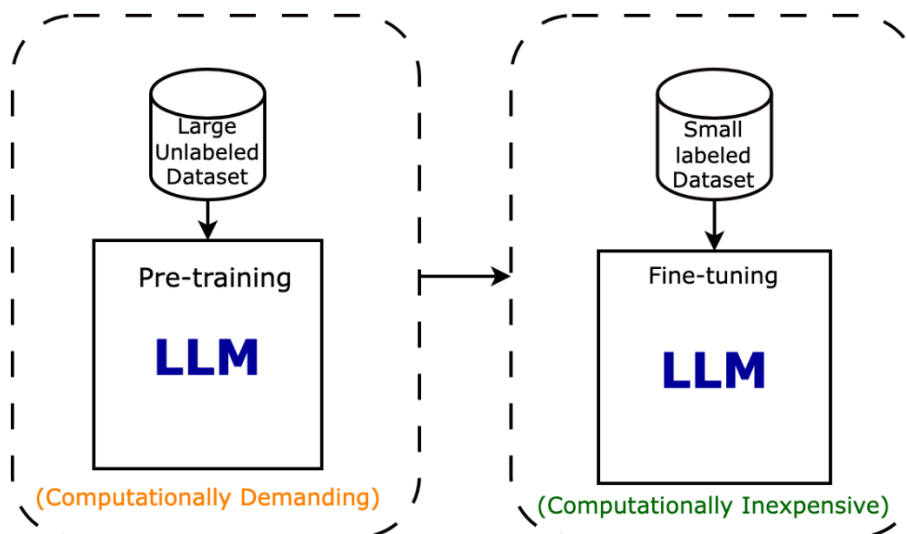
#### 1.3.4 Εκπαίδευση Μοντέλου (Model Training)

Έχοντας ολοκληρώσει το στάδιο της προ-εκπαίδευσης, τα καθαρισμένα και προετοιμασμένα δεδομένα κειμένου είναι έτοιμα για να χρησιμοποιηθούν στην εκπαίδευση του μοντέλου. Η διαδικασία ξεκινά τροφοδοτώντας το μοντέλο με μια σειρά λέξεων ή tokens. Στη συνέχεια, το μοντέλο στοχεύει στην πρόβλεψη της επόμενης λέξης στη δεδομένη ακολουθία, αξιοποιώντας τις γλωσσικές γνώσεις που απέκτησε κατά την προ-εκπαίδευση [35]. Στο στάδιο προώθησης (Forward Pass), το νευρωνικό δίκτυο παίρνει ως είσοδο τα δεδομένα και, μέσω των βαρών και των συνδέσεών του, υπολογίζει πιθανότητες για την επόμενη λέξη, παράγοντας ένα αποτέλεσμα (output). Έπειτα, μέσω μιας Συνάρτησης Απώλειας (Loss Function), υπολογίζεται το σφάλμα ανάμεσα στην πρόβλεψη του δικτύου και στην πραγματική λέξη. Το σφάλμα αυτό διαδίδεται προς τα πίσω μέσω μίας διαδικασίας που ονομάζεται Backpropagation, επιτρέποντας την προσαρμογή των βαρών του δικτύου [36]. Τέλος, ένας αλγόριθμος βελτιστοποίησης, όπως ο Gradient Descent, ενημερώνει τα βάρη λαμβάνοντας μικρά βήματα προς τη μείωση της συνάρτησης απώλειας.

Αυτός ο κύκλος επαναλαμβάνεται εκατομμύρια ή ακόμη και δισεκατομμύρια φορές, ανάλογα με το μέγεθος του συνόλου δεδομένων, έως ότου το μοντέλο επιτύχει τη βέλτιστη απόδοση.

Δεδομένου του τεράστιου όγκου δεδομένων και μεγάλου αριθμού παραμέτρων, απαιτείται τεράστια υπολογιστική ισχύς για την εκπαίδευση του μοντέλου. Για να μειωθεί ο χρόνος εκπαίδευσης και να αυξηθεί η αποδοτικότητα, εφαρμόζονται διάφορες τεχνικές παραλληλισμού [19]: Ο Παραλληλισμός Δεδομένων (Data Parallelism) μοιράζει το σύνολο δεδομένων σε πολλαπλές συσκευές, επιτρέποντας παράλληλη επεξεργασία. Με παρόμοιο τρόπο λειτουργεί και ο Παραλληλισμός Αγωγών (Pipeline Parallelism), ο οποίος διανέμει τους υπολογισμούς σε στάδια, δημιουργώντας μια γραμμή παραγωγής. Διαφορετική προσέγγιση χρησιμοποιείται στον Παραλληλισμό Τανυστών (Tensor Parallelism) που διαιρεί μεγάλους τανυστές σε μικρότερα κομμάτια και στον Παραλληλισμό Μοντέλων (Model Parallelism), όπου επιτρέπεται ο διαχωρισμός του μοντέλου σε διάφορες συσκευές [16]. Τέλος, για τη βελτιστοποίηση, χρησιμοποιείται ο Παραλληλισμός Βελτιστοποιητών (Optimizer Parallelism), εστιάζοντας στη διαχείριση των βαρών και των ενημερώσεών τους [37]. Οι τεχνικές αυτές μπορούν να εφαρμοστούν μεμονωμένα ή σε συνδυασμό, ανάλογα με τις ανάγκες και την πολυπλοκότητα του μοντέλου, επιτυγχάνοντας εξαιρετική βελτίωση της ταχύτητας της εκπαιδευτικής διαδικασίας.

Μετά την ολοκλήρωση της παραπάνω διαδικασίας, το μοντέλο εισέρχεται στη φάση της Λεπτομερούς Ρύθμισης, όπου προσαρμόζεται για συγκεκριμένες εργασίες χρησιμοποιώντας μικρότερα, σχολιασμένα σύνολα δεδομένων [6]. Αυτή η διαδικασία επιτρέπει στο μοντέλο να εξειδικεύσει τις γενικές γνώσεις που απέκτησε κατά την προ-εκπαίδευση σε πιο στοχευμένες εφαρμογές [15], όπως η κατηγοριοποίηση κειμένων, ανάλυση συναισθήματος ή η απάντηση ερωτήσεων. Το Fine-Tuning εκμεταλλεύεται την αρχιτεκτονική και τα βάρη που ήδη διαθέτει το μοντέλο και τα αναπροσαρμόζει ώστε να επιτύχει υψηλή ακρίβεια στη συγκεκριμένη εργασία.



Εικόνα 1.9: Διαδικασία προεκπαίδευσης και λεπτομερής προσαρμογής ενός μεγάλου γλωσσικού μοντέλου [38]. Κατά την προεκπαίδευση, το μοντέλο μαθαίνει από ένα μεγάλο μη επισημασμένο σύνολο δεδομένων, ενώ στη φάση της λεπτομερής προσαρμογής (fine-tuning) εκπαιδεύεται με μικρότερο επισημασμένο σύνολο δεδομένων. Η προεκπαίδευση είναι υπολογιστικά απαιτητική, ενώ η λεπτομερής προσαρμογή είναι πιο αποδοτική σε πόρους.

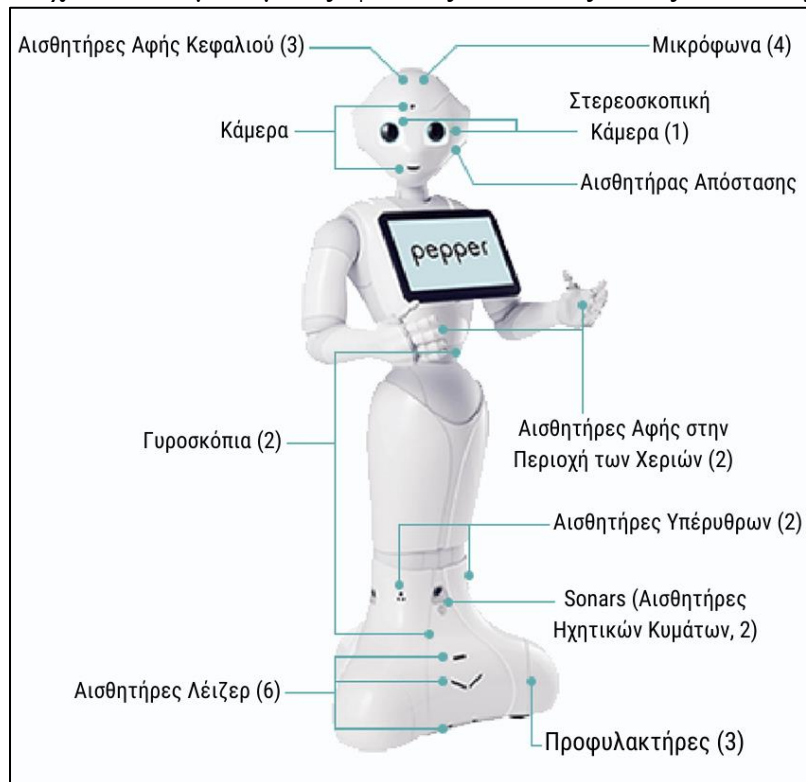
Για περαιτέρω βελτιστοποίηση, μετά την κύρια εκπαίδευση ακολουθεί η εφαρμογή Ενισχυτικής Μάθησης (Reinforcement Learning), κατά την οποία το μοντέλο εκπαιδεύεται μέσω της ανθρώπινης ανατροφοδότησης [39]. Συγκεκριμένα, άνθρωποι αξιολογητές παρέχουν οδηγίες, ή διορθώσεις για την απόκριση του μοντέλου, οι οποίες χρησιμοποιούνται για τη διαμόρφωση ενός Μηχανισμού Ανταμοιβής (Reward Mechanism). Η μέθοδος αυτή, αποτελεί την πιο γνωστή προσέγγιση και ονομάζεται Reinforcement Learning from Human Feedback (RLHF). Μόλις το μοντέλο ανταμοιβής εκπαιδευτεί, μπορεί να αντικαταστήσει τον άνθρωπο στην επισήμανση δεδομένων και στη συνέχεια να βελτιώνει τις προβλέψεις του προσπαθώντας να μεγιστοποιήσει την ανταμοιβή αυτή [40]. Ως αποτέλεσμα το μοντέλο κατανοεί σε μεγαλύτερο βαθμό τις προθέσεις του χρήστη, διαμορφώνοντας μια πιο χρήσιμη και κατανοητή συμπεριφορά προς αυτόν.

## ΚΕΦΑΛΑΙΟ 2: ΤΕΧΝΙΚΕΣ ΕΝΣΩΜΑΤΩΣΗΣ ΜΕΓΑΛΩΝ ΓΛΩΣΣΙΚΩΝ ΜΟΝΤΕΛΩΝ ΣΕ ΡΟΜΠΟΤΙΚΑ ΣΥΣΤΗΜΑΤΑ

Αδιαμφισβήτητα, η ενσωμάτωση των παραπάνω τεχνολογιών σε ρομπότ υγειονομικής περίθαλψης έχει φέρει επανάσταση στην κλινική πρακτική. Τα ρομπότ που υποστηρίζονται από μεγάλα γλωσσικά μοντέλα είναι ικανά να αλληλεπιδρούν με τους ασθενείς και τους παρόχους υγειονομικής περίθαλψης, να συλλέγουν πολύτιμες πληροφορίες, να τις αναλύουν χρησιμοποιώντας προηγμένη τεχνητή νοημοσύνη και να εκτελούν εργασίες που βελτιώνουν την αποτελεσματικότητα και την ποιότητα της περίθαλψης [41]. Σε αυτό το κεφάλαιο διερευνάται η ενσωμάτωση των μεγάλων γλωσσικών μοντέλων στα ρομποτικά συστήματα, εστιάζοντας στην φυσική υλοποίηση και υποδομή αλλά και στον τρόπο με τον οποίο αυτές οι εξελίξεις επιτρέπουν προηγμένες αλληλεπιδράσεις ανθρώπου-ρομπότ.

### 2.1 Αρχιτεκτονική Συστήματος Ενσωμάτωσης

Η ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων σε ρομποτικά συστήματα υγείας πραγματοποιείται μέσω μιας σύνθετης, πολυεπίπεδης αρχιτεκτονικής όπου κάθε επίπεδο επιτελεί συγκεκριμένη λειτουργία και επικοινωνεί με τα γειτονικά του. Στο κατώτερο επίπεδο (δηλαδή το επίπεδο που βρίσκεται πιο κοντά στο φυσικό κόσμο), βρίσκονται οι αισθητήρες και τα συστήματα συλλογής δεδομένων. Τα ρομπότ έχουν ενσωματωμένες φυσικές συσκευές όπως οπτικούς αισθητήρες υψηλής ανάλυσης, μικρόφωνα πολλαπλών καναλιών, αισθητήρες αφής και πίεσης, καθώς και εξειδικευμένους ιατρικούς αισθητήρες όπως θερμομέτρα και οξύμετρα, που μετατρέπουν τα φυσικά φαινόμενα σε ψηφιακά δεδομένα. Σε πολλές περιπτώσεις η συλλογή δεδομένων πραγματοποιείται μέσω καμερών, που συλλαμβάνουν οπτικές πληροφορίες όπως εκφράσεις προσώπου, χειρονομίες ή ιατρικές εικόνες [42]. Επιπλέον, οι ασθενείς ή οι πάροχοι υγειονομικής περίθαλψης μπορούν να εισάγουν απευθείας πληροφορίες, όπως συμπτώματα, μέσω διεπαφών κειμένου.



Εικόνα 2.1: Απεικόνιση των αισθητήρων και των συσκευών εισόδου του ρομπότ Pepper [135].

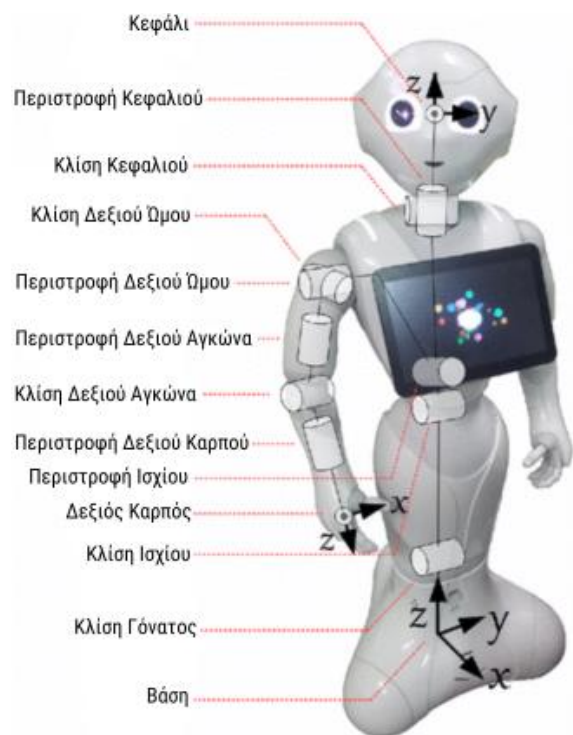


Τα δεδομένα από τους αισθητήρες διέρχονται από το επίπεδο προεπεξεργασίας, όπου πραγματοποιείται φιλτράρισμα θορύβου, συμπίεση και βελτιστοποίηση των δεδομένων. Επιπλέον, σε αυτό το επίπεδο εκτελείται ο συγχρονισμός των πολυτροπικών πληροφοριών (multimodal synchronization), δηλαδή ο συντονισμός και η χρονική ευθυγράμμιση των διαφορετικών τύπων δεδομένων όπως εικόνα, ήχος και μετρήσεις αισθητήρων [43]. Στη συνέχεια, στο επίπεδο επεξεργασίας LLM, τα δεδομένα υφίστανται επεξεργασία με δύο τρόπους: είτε τοπικά στο ίδιο το ρομπότ μέσω edge computing για απλές λειτουργίες, είτε μέσω cloud computing, σε απομακρυσμένους servers, για πιο σύνθετες διεργασίες που απαιτούν μεγαλύτερη υπολογιστική ισχύ. Παράλληλα, το σύστημα υποστηρίζει λειτουργίες caching (προσωρινή αποθήκευση), όπου τα συχνά χρησιμοποιούμενα δεδομένα και αποτελέσματα αποθηκεύονται προσωρινά για ταχύτερη πρόσβαση και βελτιστοποίηση της απόδοσης. Τέλος, το επίπεδο ελέγχου του ρομπότ μεταφράζει την έξοδο του LLM σε συγκεκριμένες ρομποτικές εντολές, διαχειρίζεται τον έλεγχο κίνησης και συντονισμού, και εφαρμόζει μηχανισμούς ασφαλείας και διόρθωσης σφαλμάτων [44].

Μέσω αυτής της διαδικασίας τα μοντέλα αναλύουν την είσοδο και την μετατρέπουν σε δομημένες γνώσεις. Ειδικότερα, το LLM συγκρίνει την πληροφορία που λαμβάνει με την τεράστια βάση γνώσης του, η οποία μπορεί να περιλαμβάνει ιατρική βιβλιογραφία, βάσεις δεδομένων συμπτωμάτων και κλινικές κατευθυντήριες γραμμές, για να παράγει σχετικές απαντήσεις ή πιθανές διαγνώσεις και να εκτελέσει τις απαιτούμενες ενέργειες. Οι ενέργειες αυτές μπορεί να περιλαμβάνουν λεκτικές απαντήσεις, όπως η παροχή συμβουλών υγείας ή ψυχολογική υποστήριξη, και σωματικές εργασίες, όπως η μέτρηση του σφυγμού ενός ασθενούς ή η ανάκτηση ιατρικών προμηθειών [45]. Για παράδειγμα, όταν ένας ασθενής λέει: «Έχω βήχα εδώ και μια εβδομάδα», το LLM προσδιορίζει το σύμπτωμα (βήχας) και τη διάρκειά του (μια εβδομάδα) και το ρομπότ μπορεί να ενημερώσει τον ασθενή: «Με βάση τα συμπτώματά σας, θα πρέπει να παρακολουθείτε τη θερμοκρασία σας και να συμβουλευτείτε έναν γιατρό εάν η κατάστασή σας επιδεινωθεί». Μπορεί επίσης να δημιουργήσει μια έκθεση που συνοψίζει τα δεδομένα του ασθενούς για έναν πάροχο υγειονομικής περίθαλψης ή να καθοδηγήσει έναν φοιτητή ιατρικής σε ένα προσομοιωμένο διαγνωστικό σενάριο. Περισσότερες πληροφορίες για τις εφαρμογές των μεγάλων γλωσσικών μοντέλων στα ιατρικά ρομπότ και τη συμβολή τους σε κάθε τομέα της ιατρικής παρουσιάζονται στο επόμενο κεφάλαιο.

## 2.2 Υλικοτεχνική Υποδομή και Συστήματα Διεπαφών

Αναφορικά με τη φυσική υλοποίηση των LLM στα ρομποτικά συστήματα υγείας, πραγματοποιείται μέσω διαφόρων τύπων πλατφορμών, καθεμία σχεδιασμένη για συγκεκριμένες



Εικόνα 2.2: Εικονική αναπαράσταση του ρομπότ Pepper αναφορικά με τις αρθρώσεις και τους άξονες κίνησης [137].

εφαρμογές στο χώρο της υγείας. Τα ρομπότ φυσικής παρουσίας αποτελούν την πιο εξελιγμένη μορφή ρομποτικών συστημάτων με μεγάλο εύρος εφαρμογών, που περιλαμβάνει τα κοινωνικά ρομπότ, τα χειρουργικά ρομποτικά συστήματα και τα ρομπότ προσωπικής φροντίδας. Πρόκειται για ανθρωπόμορφα ρομπότ όπως το Pepper της SoftBank Robotics και το Moxi της Diligent Robotics, τα οποία διαθέτουν προηγμένους βαθμούς ελευθερίας (degrees of freedom), δηλαδή τη δυνατότητα κίνησης σε πολλαπλούς άξονες, επιτρέποντάς τους να μιμούνται ανθρώπινες κινήσεις και να αλληλεπιδρούν φυσικά με ασθενείς και προσωπικό [46], [47]. Ταυτόχρονα, συνδυάζουν τις εξελιγμένες δυνατότητες των μεγάλων γλωσσικών μοντέλων για αποτελεσματική επικοινωνία με τον άνθρωπο, άμεση εκτίμηση της κατάστασης και συντονισμό των απαραίτητων ενεργειών.

Παράλληλα με τα φυσικά ρομποτικά συστήματα, σημαντικό ρόλο διαδραματίζουν οι εικονικοί βοηθοί και τα chatbots που ενσωματώνουν LLMs. Το SERMO και το Woebot αποτελούν παραδείγματα προηγμένων συστημάτων εικονικής βοήθειας που χρησιμοποιούν LLMs για την αρχική αξιολόγηση συμπτωμάτων και την καθοδήγηση ασθενών [48] [49]. Τα ιατρικά chatbots λειτουργούν ως πρώτο σημείο επαφής για τους ασθενείς, παρέχοντας άμεση πρόσβαση σε ιατρικές πληροφορίες και προκαταρκτική διαλογή. Αυτά τα συστήματα μπορούν να λειτουργήσουν είτε αυτόνομα είτε ως συμπληρωματικά εργαλεία των φυσικών ρομποτικών συστημάτων, ενισχύοντας τη συνολική αποτελεσματικότητα της ρομποτικής υποστήριξης στην υγειονομική περίθαλψη [50].

Μάλιστα, γίνεται λόγος για ρομπότ που μπορεί ακόμη και να είναι σε θέση να ολοκληρώνουν εργασίες αυτόνομα, επιτρέποντας στους ιατρούς να επιβλέπουν τις διαδικασίες από μια κονσόλα [51]. Σε αυτό το πλαίσιο, η ιατρική ρομποτική στηρίζεται σε ένα μοντέλο έξι επιπέδων αυτονομίας, που παρέχει διαφορετικούς βαθμούς αυτονομίας ανάλογα με την εφαρμογή και την πολυπλοκότητα της επέμβασης. Τα επίπεδα αυτά περιγράφονται συνοπτικά στον Πίνακα 2.1. Τα τρέχοντα συστήματα έχουν φτάσει στο επίπεδο 2, που περιλαμβάνει την αυτονομία σε συγκεκριμένες εργασίες, και στο επίπεδο 3, όπου υπάρχει υπό όρους αυτονομία με επίτηρηση από χειρουργούς. Παρότι πλήρως αυτόνομα ρομποτικά συστήματα που μπορούν να ολοκληρώσουν

επεμβάσεις από την αρχή ως το τέλος παραμένουν προς το παρόν σε θεωρητικό στάδιο, η έρευνα συνεχίζει να επεκτείνει τις δυνατότητες του πεδίου [52].

<b>Βαθμός αυτοματοποίησης</b>	<b>περιγραφή</b>
<b>Επίπεδο 0</b>	Καμία αυτονομία – ο ανθρώπινος χειριστής εκτελεί όλες τις λειτουργίες
<b>Επίπεδο 1</b>	Υποβοήθηση από ρομπότ – ο άνθρωπος διατηρεί συνεχή έλεγχο
<b>Επίπεδο 2</b>	Αυτονομία σε εργασίες – ο άνθρωπος έχει διακριτικό έλεγχο, με συγκεκριμένες εργασίες να εκτελούνται από το ρομπότ
<b>Επίπεδο 3</b>	Υπό όρους αυτονομία – ο άνθρωπος επιλέγει τη στρατηγική που εκτελείται αυτόνομα από το ρομπότ
<b>Επίπεδο 4</b>	Υψηλή αυτονομία – το ρομπότ λαμβάνει αποφάσεις αλλά υπό την επίβλεψη του ανθρώπου
<b>Επίπεδο 5</b>	Πλήρης αυτονομία – η διαδικασία εκτελείται πλήρως από το ρομπότ χωρίς την ανάγκη επίβλεψης από άνθρωπο

Πίνακας 2.1: Έξι επίπεδα αυτονομίας των ιατρικών ρομπότ, προσαρμοσμένα από την ταξινόμηση της αυτονομίας των οχημάτων, που κυμαίνονται από καμία αυτονομία έως πλήρη αυτονομία [52].

Η αλληλεπίδραση των ρομποτικών συστημάτων με το ιατρικό προσωπικό και τους ασθενείς επιτυγχάνεται μέσω πολλαπλών διεπαφών (interfaces) που διασφαλίζουν την αποτελεσματική επικοινωνία. Οι οθόνες αφής υψηλής ανάλυσης, ενσωματωμένες στα ρομποτικά συστήματα, προσφέρουν ένα διαισθητικό περιβάλλον αλληλεπίδρασης όπου οι χρήστες μπορούν να εισάγουν δεδομένα και να λαμβάνουν πληροφορίες. Τα συστήματα φωνητικής αναγνώρισης και σύνθεσης επιτρέπουν την προφορική επικοινωνία, με το ρομπότ να μπορεί να κατανοεί και να απαντά σε φωνητικές εντολές, ενώ τα προηγμένα συστήματα αναγνώρισης χειρονομιών επιτρέπουν την αλληλεπίδραση μέσω κινήσεων του σώματος [41]. Ιδιαίτερη έμφαση δίνεται στη διασύνδεση με τα υπάρχοντα κλινικά συστήματα, και συγκεκριμένα με τα συστήματα Ηλεκτρονικών Ιατρικών Φακέλων (Electronic Medical Record - EMR) και Ηλεκτρονικών Φακέλων Υγείας (Electronic Health Record - EHR), τα οποία περιέχουν το πλήρες ιατρικό ιστορικό των ασθενών, συμπεριλαμβανομένων διαγνώσεων, θεραπειών και εξετάσεων. Η διασύνδεση αυτή πραγματοποιείται μέσω αυστηρών πρωτοκόλλων ασφαλείας που διασφαλίζουν την προστασία των ευαίσθητων ιατρικών δεδομένων [53].

Συνοψίζοντας, η ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων στα ρομποτικά συστήματα υγείας αποτελεί μια πολύπλοκη τεχνολογική πρόκληση που απαιτεί την αρμονική συνεργασία πολλαπλών επιπέδων επεξεργασίας, από τη συλλογή δεδομένων μέσω αισθητήρων μέχρι την τελική εκτέλεση ενεργειών. Η αρχιτεκτονική αυτή επιτρέπει στα ρομποτικά συστήματα να λειτουργούν ως αποτελεσματικοί συνεργάτες στο χώρο της υγείας, καθιστώντας την κλινική πρακτική πιο αποτελεσματική, διαδραστική και ασθενοκεντρική. Αυτή η τεχνολογική υποδομή αποτελεί το θεμέλιο για την ανάπτυξη εξειδικευμένων εφαρμογών σε διάφορους τομείς της ιατρικής, από τη διάγνωση και θεραπεία μέχρι την εκπαίδευση και την ψυχική υγεία, οι οποίες θα αναλυθούν διεξοδικά στο επόμενο κεφάλαιο.

## ΚΕΦΑΛΑΙΟ 3: ΕΦΑΡΜΟΓΕΣ LLMS ΣΕ ΡΟΜΠΟΤ ΠΑΡΟΧΗΣ ΥΠΗΡΕΣΙΩΝ ΣΤΗΝ ΥΓΕΙΟΝΟΜΙΚΗ ΠΕΡΙΘΑΛΨΗ

Ο τομέας της υγειονομικής περίθαλψης βρίσκεται στην πρώτη γραμμή της υιοθέτησης προηγμένων τεχνολογιών, με στόχο τη βελτίωση της ποιότητας των υπηρεσιών και της εμπειρίας του ασθενούς. Την τελευταία δεκαετία, η τεχνητή νοημοσύνη γίνεται όλο και πιο επωφελής για την συγκεκριμένο κλάδο [9]. Είναι γεγονός πως η εμφάνιση των Μεγάλων Γλωσσικών Μοντέλων έχει προκαλέσει μια μετασχηματιστική αλλαγή στο τοπίο της υγειονομικής περίθαλψης, προσφέροντας πρωτοφανείς ευκαιρίες για καινοτομία και πρόοδο [54]. Τα αυτόνομα ρομπότ παροχής υπηρεσιών, σε συνδυασμό με την δυνατότητα των LLMS να αποκτούν χρήσιμες γνώσεις κωδικοποιημένες σε ιατρικές βάσεις δεδομένων [55], έχουν επιδείξει αξιοσημείωτες δυνατότητες σε ένα ευρύ φάσμα εφαρμογών υγειονομικής περίθαλψης. Συχνά χρησιμοποιείται ο όρος ιατρικά ρομπότ, ο οποίος αναφέρεται σε έναν τύπο ρομπότ επαγγελματικών υπηρεσιών που χρησιμοποιούνται εντός και εκτός νοσοκομείων για τη βελτίωση του συνολικού επιπέδου περίθαλψης των ασθενών [56]. Βεβαίως τα συστήματα αυτά προσφέρουν καινοτόμες λύσεις που ενισχύουν τη διαγνωστική ακρίβεια, την ανάκτηση γνώσεων και τη φροντίδα ασθενών [57], ενώ παράλληλα δημιουργούν σημαντικές λειτουργικές αποδοτικότητες και μειώσεις κόστους για τις εγκαταστάσεις υγειονομικής περίθαλψης. Στο παρόν κεφάλαιο, διερευνώνται οι σημαντικότερες εφαρμογές των ρομπότ στον τομέα της υγειονομικής περίθαλψης, αναδεικνύοντας τον ρόλο τους στη διαμόρφωση του μέλλοντος της ιατρικής.

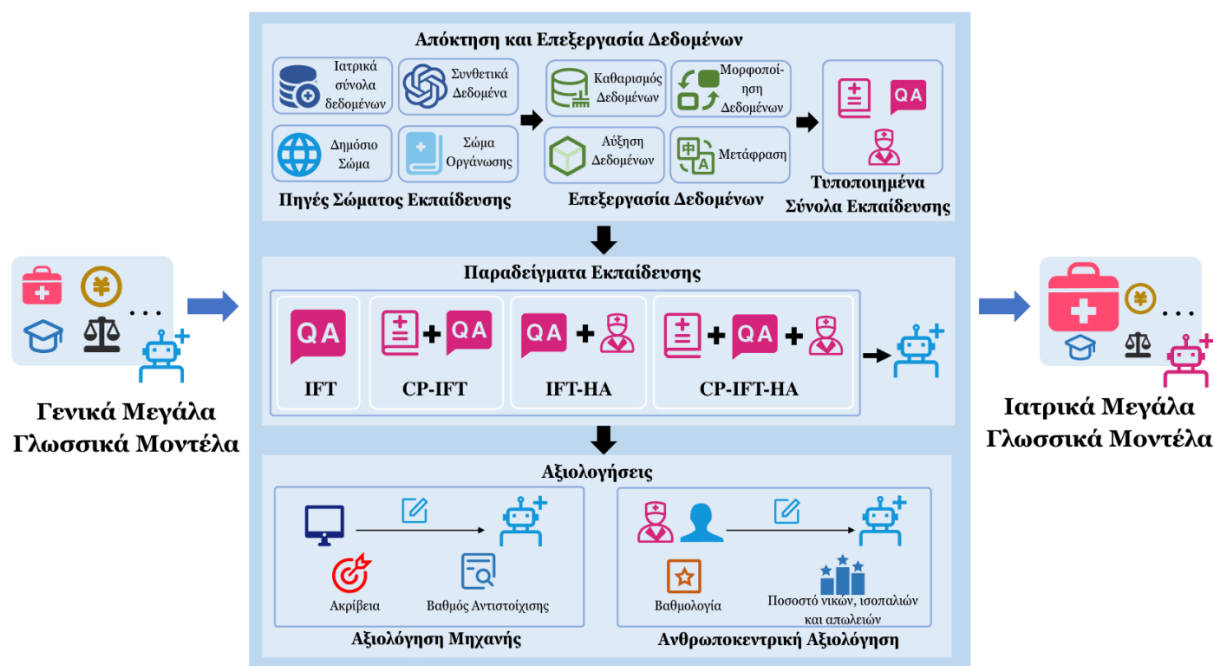
### 3.1 Διάγνωση και Θεραπεία

Στον τομέα της ιατρικής, τα LLMS αξιοποιούνται με ιδιαίτερη επιτυχία στην υποστήριξη ιατρικών διαγνώσεων, αναλύοντας τεράστιες ποσότητες ιατρικών δεδομένων και εξάγοντας πολύτιμες πληροφορίες [54]. Τα ρομπότ ως εικονικοί βοηθοί, αναλύουν συμπτώματα ασθενών, ιατρικούς φακέλους, αποτελέσματα εξετάσεων και σχετικά αρχεία, και έπειτα επεξεργάζονται αυτούς τους όγκους δεδομένων και εξάγουν σημαντικές πληροφορίες που συνδέονται με διαγνωστικά συμπτώματα. Για παράδειγμα, στον τομέα της καρδιολογίας, τα LLM αξιοποιούνται για την επεξεργασία δεδομένων ηλεκτροκαρδιογραφήματος, διευκολύνοντας την ανίχνευση μικρών ανωμαλιών που ενδέχεται να υποδηλώνουν καρδιακές παθήσεις. Μέσω αυτής της προσέγγισης, υποβοηθούν τους καρδιολόγους να κάνουν ακριβείς διαγνώσεις και να εφαρμόζουν τις κατάλληλες στρατηγικές θεραπείας [58].

Στο ίδιο πλαίσιο, τα LLMS μπορούν να αξιοποιηθούν και για τη δημιουργία συνοπτικών περιλήψεων του ιατρικού ιστορικού κάθε ασθενούς, συμπεριλαμβανομένων και των παραπάνω δεδομένων, αναδεικνύοντας σύνθετα μοτίβα και ουσιώδεις πληροφορίες για τον ασθενή, που διαφορετικά θα μπορούσαν να περάσουν απαρατήρητα. Κατά αυτόν τον τρόπο τα μεγάλα γλωσσικά μοντέλα χρησιμεύουν ως συστήματα υποστήριξης κλινικών αποφάσεων που καθοδηγούν τους ιατρούς σε θεραπευτικές επιλογές με βάση τεκμηριωμένες κατευθυντήριες γραμμές και κλινική εμπειρογνομosύνη [50]. Ειδικότερα, τα LLMS έχουν τη δυνατότητα να παρέχουν συστάσεις θεραπείας προσαρμοσμένες στις ιδιαίτερες ανάγκες του κάθε ασθενούς, λαμβάνοντας υπόψη παράγοντες όπως οι συννοσηρότητες, οι καθημερινές συνήθειες και οι

προτιμήσεις στη θεραπεία. Αυτό επιτρέπει στους γιατρούς να λαμβάνουν καλά ενημερωμένες αποφάσεις και να παρέχουν εξατομικευμένη φροντίδα στους ασθενείς τους [58].

Μάλιστα, ειδικά προσαρμοσμένες εκδοχές μεγάλων γλωσσικών μοντέλων εισήχθησαν για να ανταποκρίνονται στις ιδιαιτερότητες της κλινικής γλώσσας. Ειδικότερα, τα Clinical Camel, Med-Alpaca και PMC-LLaMA που προέρχονται από το LLaMA [59], [60], [61], το BioBERT του BERT [62], το GatorTron, που βασίζεται στο RoBERTa (μια βελτιωμένη εκδοχή του BERT) [63], και πολλά άλλα έχουν εκπαιδευτεί σε βιοϊατρικά κείμενα και κλινικά δεδομένα και εμφανίζονται πλήρως ικανά να αναγνωρίζουν και να αναπαράγουν την πολύπλοκη ιατρική ορολογία. Ακόμη, οι κλινικές απαιτήσεις οδήγησαν στην δημιουργία εξειδικευμένων μοντέλων, τα οποία έχουν εκπαιδευτεί αποκλειστικά σε κλινικά και ιατρικά δεδομένα και αποδίδουν καλύτερα σε συγκεκριμένες ιατρικές εφαρμογές. Ένα παραδείγματα τέτοιων μοντέλων, αποτελεί το Med-PaLM ένα γλωσσικό μοντέλο της Google Health, ειδικά σχεδιασμένο για τον ιατρικό τομέα και εκπαιδευμένο σε ιατρικές ερωταπαντήσεις και πληροφορίες, το οποίο εφαρμόζεται στην υποστήριξη λήψης αποφάσεων και στην εκπαίδευση ιατρικού προσωπικού [64]. Τα ιατρικά LLMs αξιοποιούνται και στην ανάλυση ιατρικών εικόνων, συμπεριλαμβανομένων των ακτινολογικών σαρώσεων, των παθολογικών διαφανειών και των δερματολογικών εικόνων [58]. Συγκεκριμένα, μοντέλα όπως το ChatCAD, υπερέχουν στην αυτοματοποιημένη παραγωγή ιατρικών εκθέσεων από εικόνες, καθώς μπορούν να εντοπίσουν ανεπαίσθητους δείκτες ασθενειών, οδηγώντας σε βελτιωμένη διαγνωστική ακρίβεια [54]. Αυτή η τεχνολογία ενισχύει τα συστήματα υποβοηθούμενης διάγνωσης, καθιστώντας τα πιο αποτελεσματικά και διαδραστικά, επιτρέποντας στους γιατρούς να ερμηνεύουν τα αποτελέσματα μέσω ερωτήσεων ανοικτού τύπου [50]. Με την επεξεργασία αυτών των εικόνων και τη σύγκρισή τους με εκτεταμένες βάσεις δεδομένων, τα LLMs συμβάλλουν σημαντικά στην ανάπτυξη εξατομικευμένων και τεκμηριωμένων διαγνωστικών εργαλείων.



Εικόνα 3.1: Διαδικασία Εκπαίδευσης από Γενικά LLMs σε Ιατρικά LLMs.

Αρχικά, συλλέγονται και επεξεργάζονται ιατρικά δεδομένα για τη δημιουργία ενός συνόλου εκπαίδευσης.

Στη συνέχεια, τα Γενικά LLMs εκπαιδεύονται με στόχο την απόκτηση ιατρικών γνώσεων, χρησιμοποιώντας τρία στάδια: Συνεχιζόμενη Προεκπαίδευση (CP), Βελτίωση μέσω Εντολών (IFT) και Ευθυγράμμιση με Ανθρώπινη Παρέμβαση (HA). Τέλος, τα Ιατρικά LLMs αξιολογούνται τόσο από μηχανές όσο και από ανθρώπους [45].

### 3.2 Χειρουργικές Επεμβάσεις

Όσον αφορά τα ρομπότ παροχής υπηρεσιών στον τομέα της ιατρικής, τα ρομπότ χειρουργικής βοήθειας βρίσκουν όλο και μεγαλύτερη εφαρμογή, ενώ οι τεχνολογίες ελέγχου κίνησης εξελίσσονται. Εξοπλισμένα με τεχνολογίες τεχνητής νοημοσύνης και υψηλής ευκρίνειας όραση υπολογιστή, τα ρομπότ βοηθούν τους χειρουργούς να επιτύχουν νέα επίπεδα ταχύτητας και ακρίβειας κατά την εκτέλεση πολύπλοκων επεμβάσεων [51]. Τα μεγάλα γλωσσικά μοντέλα αξιοποιούνται στον χειρουργικό σχεδιασμό αναλύοντας τεράστιους όγκους ιατρικής βιβλιογραφίας, δεδομένα ασθενών και κλινικές κατευθυντήριων γραμμών για να προτείνουν τις καταλληλότερες χειρουργικές προσεγγίσεις. Βεβαίως, λαμβάνοντας υπόψη κρίσιμους παράγοντες, όπως η ηλικία του ασθενούς και το ιστορικό προηγούμενων επεμβάσεων μπορούν να προτείνουν εξατομικευμένες χειρουργικές μεθοδολογίες. Επιπλέον, τα LLMs διαθέτουν την ικανότητα να παράγουν σύνθετες γλωσσικές οδηγίες και να διευκολύνουν την αμφίδρομη επικοινωνία με την ιατρική ομάδα σε πραγματικό χρόνο, παρέχοντας πληροφορίες με διαισθητικό και γρήγορο τρόπο [65]. Η ενσωμάτωση αυτής της τεχνολογίας επιτρέπει επίσης τη μετάδοση ζωντανού βίντεο από το χειρουργείο, καθιστώντας εφικτές τις διαβουλεύσεις με ειδικούς από απομακρυσμένες τοποθεσίες, ώστε να διασφαλίζεται η συμμετοχή των καλύτερων χειρουργών στις επεμβάσεις. Προς το παρόν, η εφαρμογή της τεχνητής νοημοσύνης στην ρομποτική χρησιμοποιείται κατά κύριο λόγο σε δύο κατηγορίες ρομποτικών επεμβάσεων:

Η πρώτη κατηγορία αφορά τις ελάχιστες επεμβατικές χειρουργικές επεμβάσεις για τον κορμό, και περιλαμβάνει τη ρομποτική υστερεκτομή, τη ρομποτική προστατεκτομή, τη βαριατρική χειρουργική και άλλες επεμβάσεις επικεντρωμένες στους μαλακούς ιστούς. Τα χειρουργικά ρομπότ εισάγονται στον κορμό μέσω μιας μικρής τομής και κλειδώνουν στην θέση τους, δημιουργώντας μία σταθερή πλατφόρμα για την εκτέλεση χειρουργικών επεμβάσεων μέσω τηλεχειρισμού. Η τεχνική αυτή επιτρέπει την υλοποίηση τομής μεγέθους κουμπιού, προσδίδοντας μεγάλη ευκολία και ακρίβεια στην διαδικασία της εσωτερικής επέμβασης και μειώνοντας τους χρόνους ανάρρωσης και τη πιθανότητα μόλυνσης και άλλων επιπλοκών. Μία πρωτοποριακή τεχνολογία που χρησιμοποιείται στις ελάχιστες επεμβατικές χειρουργικές επεμβάσεις αποτελεί το χειρουργικό σύστημα Da Vinci, το οποίο απαρτίζεται από ένα μεγεθυμένο τρισδιάστατο σύστημα όρασης μεγάλης ευκρίνειας και εργαλεία υψηλής ακρίβειας που έχουν τη δυνατότητα να λυγίζουν και να περιστρέφονται πέρα από τις δυνατότητες του ανθρώπινου χεριού [66]. Η χρήση του συστήματος εφαρμόζεται όλο και περισσότερο σε προστατεκτομές, επισκευές καρδιακών βαλβίδων και σε νευρικές και χειρουργικές επεμβάσεις [67]. Η ενσωμάτωση των LLMs μπορεί να ενισχύσει σημαντικά τα ρομποτικά αυτά συστήματα, προσφέροντας φυσική επικοινωνία με τους χειρουργούς και παρέχοντας οδηγίες, επεξηγήσεις ή γρήγορη πρόσβαση σε ιατρική βιβλιογραφία κατά τη διάρκεια της επέμβασης. Ωστόσο η εφαρμογή της τεχνητής νοημοσύνης στο Da Vinci Surgical System είναι ακόμα σε πρώιμο στάδιο.



(α)

(β)

Εικόνα 3.2: Το ρομποτικό σύστημα Da Vinci.

(α) Κοντινή λήψη των ρομποτικών βραχιόνων, (β) Χρήση του συστήματος από χειρουργό [68], [69].

Ταυτόχρονα μελετάται η ρομποτικά υποβοηθούμενη ελάχιστη επεμβατική χειρουργική του παχέος εντέρου με τεχνητή νοημοσύνη, με σκοπό την βελτιωμένη οπτικοποίηση και την ενίσχυση της χειρουργικής ακρίβειας [70]. Η χρήση ρομποτικής υποβοήθησης για ορθοκολικές εκτομές τυποποιείται όλο και περισσότερο στα ιατρικά κέντρα, ενώ τα ρομποτικά συστήματα έχουν επιδείξει σημαντική αποτελεσματικότητα σε χειρουργικές επεμβάσεις καρκίνου του παχέος εντέρου [71]. Μάλιστα, ερευνητικά δεδομένα αποκαλύπτουν την αξία της συμβολής των ρομποτικών συστημάτων με τεχνητή νοημοσύνη στην αυξημένη επιδεξιότητα και στη μείωση των μετεγχειρητικών επιπλοκών στη χειρουργική κεφαλής και τραχήλου [72]. Παράλληλα, σύμφωνα με τα αποτελέσματα της μελέτης που δημοσιεύθηκε στο περιοδικό «Journal of Thoracic and



Cardiovascular Surgery» (JTCVS), η ελάχιστη επεμβατική καρδιοχειρουργική επωφελείται σημαντικά από την εφαρμογή της τεχνητής νοημοσύνης σε ρομποτικές πρακτικές, προσδίδοντας συντομότερους χρόνους χειρουργικής επέμβασης, μειωμένη απώλεια αίματος και λιγότερες επιπλοκές [73]. Μία ακόμη αξιοσημείωτη καινοτομία στον τομέα της ρομποτικής χειρουργικής αποτελεί η χρήση του μοντέλου «master-slave», που επιτρέπει στους χειρουργούς να ελέγχουν τους ρομποτικούς βραχίονες εξ αποστάσεως από μία εξιδεικευμένη κονσόλα ελέγχου υψηλής ακρίβειας. Παρότι κανένα χειρουργικό ρομποτικό σύστημα δεν έχει ακόμα τη δυνατότητα να εκτελεί επεμβάσεις εντελώς αυτόνομα, η χρήση ημιαυτόνομων χειρουργικών συστημάτων εξοπλισμένων με τις δυνατότητες των LLMs, έχει σημειώσει αξιοσημείωτη αύξηση, προσφέροντας σχολαστική παρακολούθηση και βελτιωμένες χειρουργικές ενέργειες [74]. Ακόμη, πολλαπλές μελέτες ανέδειξαν τις δυνατότητες των μεγάλων γλωσσικών μοντέλων για την υποβοήθηση πολύπλοκων χειρουργικών επεμβάσεων όγκων στον εγκέφαλο, μέσω της καθοδήγησης εικόνας [75], [76]. Η υιοθέτηση της ελάχιστης επεμβατικής ρομποτικής χειρουργικής έχει μεταμορφώσει βαθιά το χειρουργικό τοπίο, αντιπροσωπεύοντας μια σημαντική πρόοδο προς τις πιο αποτελεσματικές και λιγότερο επεμβατικές θεραπευτικές επιλογές για τους ασθενείς.

Οι ορθοπεδικές χειρουργικές επεμβάσεις αποτελούν τη δεύτερη κατηγορία ρομποτικής χειρουργικής. Περιλαμβάνουν τον προγενέστερο προγραμματισμό των συσκευών, ώστε αυτές να εκτελούν κοινές ορθοπεδικές χειρουργικές επεμβάσεις, όπως η αντικαταστάσεις γόνατος και ισχίου. Σε αυτή την κατηγορία χειρουργικών επεμβάσεων, τα ρομπότ χρησιμοποιούν χωρικά καθορισμένα όρια για την υποβοήθηση του χειρουργού, μέσω έξυπνων ρομποτικών βραχιόνων, τρισδιάστατης απεικόνισης και ανάλυσης δεδομένων. Η χρήση τεχνητής νοημοσύνης κατά την μοντελοποίηση των ρομπότ καθιστά εφικτή την εκπαίδευσή τους σε συγκεκριμένες ορθοπεδικές επεμβάσεις, με ακριβείς οδηγίες για το πού να πάνε και πώς να εκτελέσουν τις επεμβάσεις [51]. Σαφώς, η ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων στις ορθοπεδικές χειρουργικές επεμβάσεις επιτρέπει την εκπαίδευση των ρομποτικών συστημάτων για την εκτέλεση εξαιρετικά ακριβών κινήσεων. Χρησιμοποιώντας δεδομένα όπως ακτινογραφίες και MRI (Μαγνητική Τομογραφία), τα χειρουργικά ρομπότ εμφανίζονται ικανά να αναγνωρίζουν ανατομικές δομές και να σχεδιάσουν προσωποποιημένα πλάνα θεραπείας. Με την καθοδήγηση μοντέλων μηχανικής μάθησης, συστήματα όπως το Mako Robotic Arm μπορούν να βοηθήσουν στη χειρουργική τοποθέτηση προσθετικών ή στη διενέργεια κοψιμάτων με ακρίβεια χιλιοστού. Ο ρομποτικός βραχίονας Mako είναι ένα χειρουργικό ρομπότ που βοηθά τους χειρουργούς να εκτελούν σύνθετες χειρουργικές επεμβάσεις ισχίου και γόνατος, μία τεχνολογία που μειώνει τα περιθώρια σφάλματος και αυξάνει την αποτελεσματικότητα, εξασφαλίζοντας καλύτερα αποτελέσματα για τον ασθενή [77]. Σύμφωνα με πρόσφατες μελέτες, έρευνες έχουν αποκαλύπτουν την συμβολή των μεγάλων γλωσσικών δεδομένων στην βελτιστοποίηση της τοποθέτησης των εμφυτευμάτων και στην δυνατότητα περιορισμού των μακροπρόθεσμων επιπλοκών στην ορθοπεδική χειρουργική [78]. Αναντίρρητα, η εφαρμογή τεχνολογιών τεχνητής νοημοσύνης και ρομποτικής στην ορθοπεδική χειρουργική φέρνει σημαντικές αλλαγές στον προεγχειρητικό σχεδιασμό, στην ενδοεγχειρητική ακρίβεια και στα μετεγχειρητικά αποτελέσματα σε πληθώρα χειρουργικών επεμβάσεων, συμπεριλαμβανομένων εκείνων που αφορούν το ισχίο, το γόνατο και τον ώμο. Η τεχνητή νοημοσύνη καθιστά δυνατή την εξατομίκευση της χειρουργικής ορθοπεδικής, με βάση παράγοντες όπως η ανατομία και το ιατρικό ιστορικό του κάθε ασθενούς [79]. Αυτό έχει ως



αποτέλεσμα την καλύτερη επιλογή εμφυτευμάτων και την βελτιστοποίηση των χειρουργικών τεχνικών για κάθε περίπτωση ξεχωριστά [80].



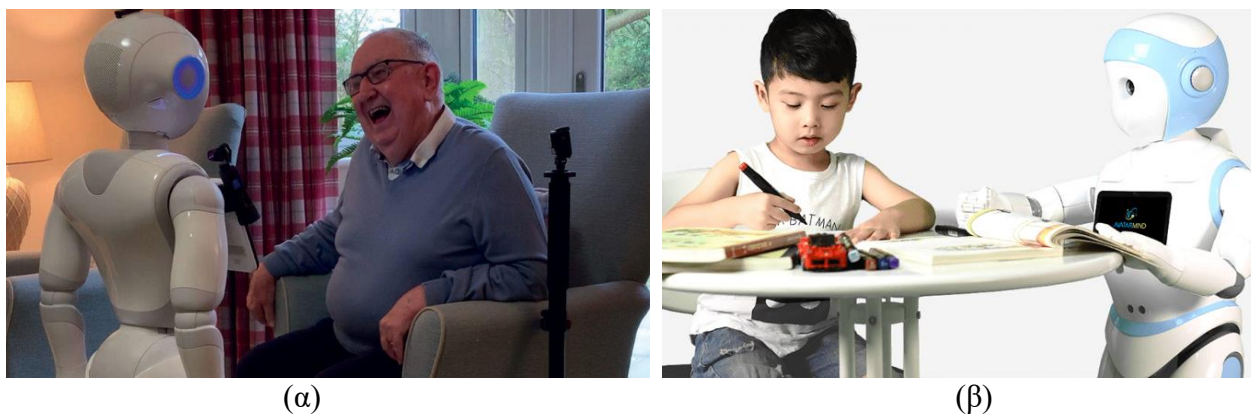
Εικόνα 3.3: Το ρομποτικό σύστημα Mako σε εγκαταστάσεις Ελληνικού νοσοκομείου [81].

### 3.3 Ψυχική Υποστήριξη

Η ενσωμάτωση των μεγάλων γλωσσικών μοντέλων στα συστήματα υποστήριξης της ψυχικής υγείας αποτελεί σημαντική πρόοδο στην παροχή προσβάσιμης και ευέλικτης βοήθειας για την ψυχική υγεία. Τα συστήματα αυτά χρησιμεύουν ως συμπληρωματικά εργαλεία της παραδοσιακής ψυχικής υγειονομικής περίθαλψης, προσφέροντας άμεση υποστήριξη, διατηρώντας παράλληλα τις συνδέσεις με τους ανθρώπινους παρόχους υγειονομικής περίθαλψης. Η χρήση των LLMs επιτρέπει στα ρομποτικά συστήματα να επεξεργάζονται και να ανταποκρίνονται σε εισόδους φυσικής γλώσσας με πρωτοφανή πολυπλοκότητα, καθιστώντας την υποστήριξη της ψυχικής υγείας πιο προσιτή και ελκυστική.

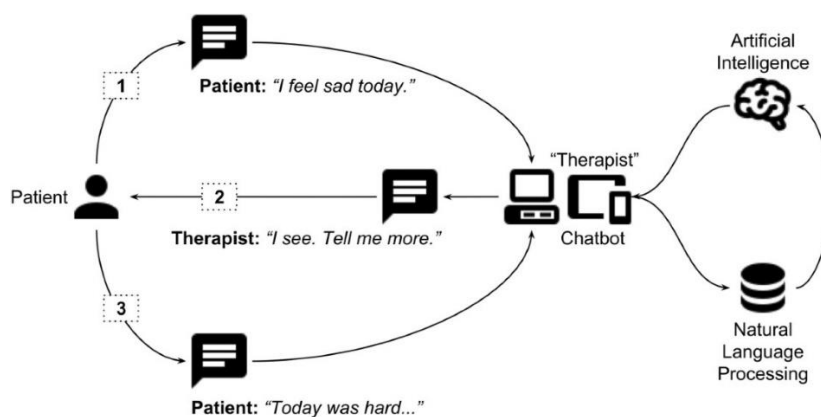
Αναντίρρητα, τα κοινωνικά ρομπότ που είναι εξοπλισμένα με LLM αντιπροσωπεύουν μια σημαντική πρόοδο στον τομέα της ψυχικής υγείας, συνδυάζοντας τη φυσική παρουσία με εξελιγμένες ικανότητες συνομιλίας για την παροχή συναισθηματικής υποστήριξης και θεραπευτικών αλληλεπιδράσεων τόσο σε κλινικές όσο και σε οικιακές ρυθμίσεις. Πρόκειται για «φιλικά» ρομπότ που αλληλοεπιδρούν άμεσα με τους ανθρώπους και αποσκοπούν στον μετριασμό των αισθημάτων μοναξιάς και απομόνωσης, ενώ ταυτόχρονα παρέχουν καθοδήγηση για την ευημερία και ψυχολογική στήριξη [82]. Χρησιμοποιούνται σε περιβάλλοντα μακροχρόνιας φροντίδας για την παροχή κοινωνικής αλληλεπίδρασης και παρακολούθησης, αλλά και σε νοσοκομειακά περιβάλλοντα, όπου προσφέρουν οδηγίες στους επισκέπτες και τους ασθενείς [51]. Σε γενικές γραμμές, τα κοινωνικά ρομπότ συμβάλλουν στη μείωση του φόρτου εργασίας των φροντιστών και στη βελτίωση της συναισθηματικής ευεξίας των ασθενών. Στο ίδιο πλαίσιο βρίσκονται και τα ρομπότ προσωπικής φροντίδας, τα οποία αναπτύχθηκαν για υποστηρίξουν την ανεξάρτητη διαβίωση των ηλικιωμένων στο σπίτι και χρησιμοποιούν LLMs για να διατηρούν συνομιλίες και να μαθαίνουν τις προσωπικές προτιμήσεις των χρηστών με την πάροδο του χρόνου [83]. Μπορούν να υπενθυμίζουν τη λήψη φαρμάκων, να εντοπίζουν αλλαγές στα πρότυπα συμπεριφοράς και να προσφέρουν συντροφιά, διατηρώντας παράλληλα τη σύνδεση με την

οικογένεια μέσω ειδικών λειτουργιών επικοινωνίας. Το Pepper, που αναπτύχθηκε από τη Softbank Robotics, είναι ένα ανθρωποειδές ρομπότ που έχει σχεδιαστεί κυρίως για την αλληλεπίδραση ανθρώπου-ρομπότ και χρησιμοποιείται σε εγκαταστάσεις φροντίδας ηλικιωμένων, όπου η μοναξιά και η κατάθλιψη είναι κοινές ανησυχίες [46]. Εξοπλισμένο με διάφορους αισθητήρες πλοήγησης, δυνατότητες ομιλίας και τεχνολογίες LLMs είναι ικανό να αναλύει το περιβάλλον εργασίας του και να ερμηνεύει τις ανθρώπινες εκφράσεις, προσφέροντας συνομιλίες ψυχικής υγείας και συναισθηματική υποστήριξη [84].



Εικόνα 3.4: Περιπτώσεις χρήσης του ρομπότ Pepper σε (α) περίθαλψη ηλικιωμένων και (β) φροντίδα παιδιών [85], [86].

Ακόμη, τα LLMs μπορούν να χρησιμοποιηθούν ως η υποκείμενη νοημοσύνη για τα chatbots υγείας, παρέχοντας αδιάκοπη και εξατομικευμένη υποστήριξη στον τομέα της υγειονομικής περίθαλψης [54]. Η εφαρμογή των LLMs σε συστήματα υποστήριξης της ψυχικής υγείας έχει μεταμορφώσει την ικανότητα αυτών των πλατφορμών να παρέχουν διαφοροποιημένες και κατάλληλες για το πλαίσιο απαντήσεις. Σε αντίθεση με τα προηγούμενα συστήματα που βασίζονται σε κανόνες, οι λύσεις που υποστηρίζονται από τα LLMs μπορούν να κατανοούν σύνθετες συναισθηματικές εκφράσεις, να διατηρούν το πλαίσιο κατά τη διάρκεια των συνομιλιών και να παράγουν απαντήσεις που αποδεικνύουν συναισθηματική νοημοσύνη. Αυτή η πρόοδος επιτρέπει πιο φυσικές αλληλεπιδράσεις θεραπευτικού περιεχομένου, κάνοντας τους χρήστες να μοιράζονται πιο άνετα τις ανησυχίες τους και να συμμετέχουν στο σύστημα υποστήριξης [87]. Η τεχνολογία αυτή βρίσκει σημαντική εφαρμογή στην διαχείριση των ασθενών με ψυχικές διαταραχές, όπου υπάρχει μεγάλη απαίτηση διεπιστημονικών προσεγγίσεων. Συγκεκριμένα, το SERMO αποτελεί ένα εργαλείο συνομιλίας που καθοδηγεί τους ασθενείς με ψυχικές παθήσεις στη ρύθμιση των συναισθημάτων τους για την βέλτιστη διαχείριση των αρνητικών σκέψεων. Χρησιμοποιώντας τις κειμενικές εισόδους του χρήστη, εντοπίζει αυτόματα το είδος του συναισθήματος και προτείνει εξατομικευμένες δραστηριότητες ή ασκήσεις ενσυνειδητότητας, προσαρμοσμένες στις συναισθηματικές ανάγκες του ατόμου [48]. Στο ίδιο πλαίσιο, το Woebot είναι ένας πλήρως αυτοματοποιημένος ψηφιακός συνομιλητής που προσφέρει υπηρεσίες γνωστικής συμπεριφορικής θεραπείας, ειδικά σχεδιασμένες για εφήβους που αντιμετωπίζουν κατάθλιψη. Αποδείχθηκε ιδιαίτερα αποτελεσματικό στη μείωση των συμπτωμάτων της κατάθλιψης, ξεπερνώντας την προσέγγιση της ομάδας ελέγχου που βασιζόταν αποκλειστικά στη χρήση πληροφοριακού υλικού σε μορφή ηλεκτρονικών βιβλίων [49].



Εικόνα 3.5: Δείγμα αλληλεπίδρασης μεταξύ ενός ασθενούς και ενός θεραπευτή chatbot [88].

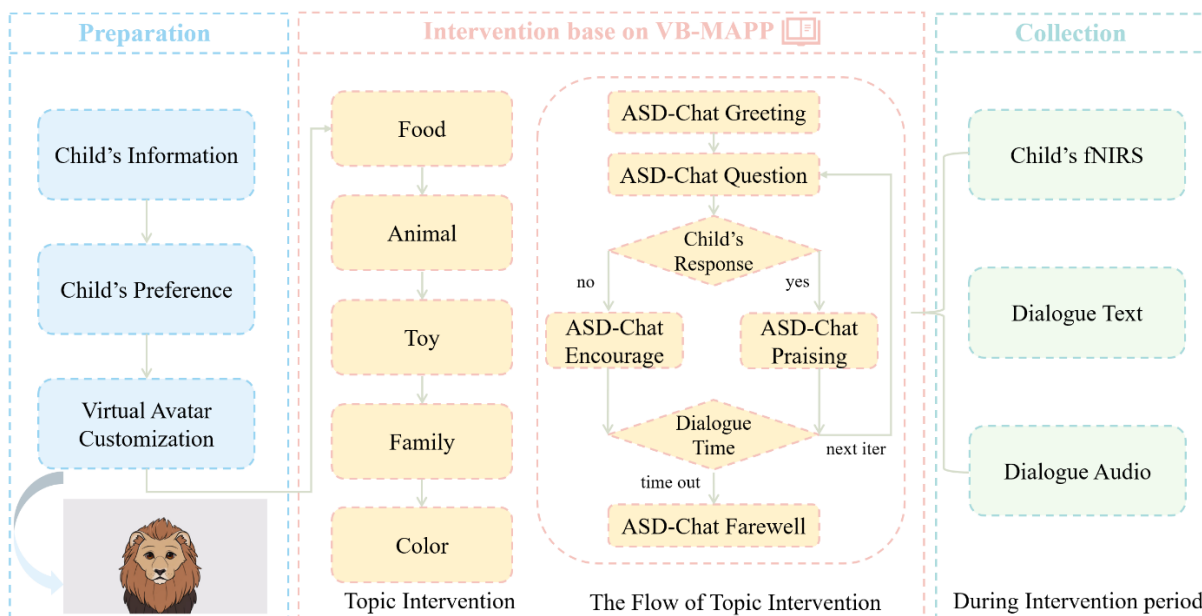
Πέρα από τις ήδη αναφερθείσες εφαρμογές, τα μεγάλα γλωσσικά μοντέλα συνεχίζουν να αξιοποιούνται σε καινοτόμες λύσεις που ανταποκρίνονται στις ανάγκες συγκεκριμένων πληθυσμών και διαταραχών, προσφέροντας νέες δυνατότητες υποστήριξης και παρέμβασης. Μεταξύ αυτών, ξεχωρίζει το WellnessBot, ένα πρωτοποριακό εργαλείο σχεδιασμένο για τη διαχείριση διατροφικών διαταραχών, που παρέχει προσαρμοσμένη και προσβάσιμη υποστήριξη, πέρα από τα όρια των κλινικών εγκαταστάσεων [89]. Πρόκειται για μία πλατφόρμα chatbot με βάση τα LLMs, που στοχεύει στην ενίσχυση της καθημερινής διαχείρισης αυτών των προκλήσεων, παρέχοντας καθοδήγηση και συναισθηματική υποστήριξη σε μη κλινικά περιβάλλοντα. Μέσω της ανάλυσης των δεδομένων χρήστη, προσφέρει προσαρμοσμένες συστάσεις και ασκήσεις που βοηθούν στη σταδιακή αλλαγή συμπεριφοράς. Η ικανότητα αυτών των συστημάτων να αναλύουν γλωσσικά μοτίβα και ανιχνεύουν συναισθηματικές καταστάσεις επιτρέπει την συμμετοχή σε ανοιχτές συζητήσεις, διατηρώντας παράλληλα θεραπευτικά πλαίσια. Έτσι, το άτομο που βρίσκεται σε ευάλωτη και συναισθηματικά φορτισμένη κατάσταση, μπορεί να λάβει άμεσα την υποστήριξη που χρειάζεται, υπό την ομπρέλα της φυσικής και ιδιωτικής αλληλεπίδρασης, έναντι των τυποποιημένων απαντήσεων που παρέχουν τα παραδοσιακά αυτοματοποιημένα συστήματα.

Log #	User Message	Chatbot Message
1	Suddenly, I'm feeling really tired... My eyelids feel heavy	It makes sense that you're feeling tired right now... <b>Remember, you wrote in your wellness plan that you tend to binge eat when you're tired.</b> This might be one of those times...
	<b>I just wrote in my mood journal.</b> Thanks to you, I was able to calm down quickly when I was irritated. I'll even add a sticker to my diary to celebrate myself!	Aww, my pleasure! ... You'll be able to boost your self-esteem even more by giving yourself positive feedback...
2	<b>If I notice even a small improvement while writing my diary, I'll give myself some praise.</b> Even if it's a small thing.	Absolutely, it's very important to be proud of yourself. Recognizing and acknowledging your gradual progress is one of the best ways to grow. Even the smallest achievements can lead to significant changes, so never underestimate them...
	<b>I feel much better thanks to you!</b> I'll make sure to talk with you more often. It seems really helpful.	Definitely, let's keep talking! ...
3	What effect does vomiting [after binge eating] have on the mind?	Vomiting has a significantly negative impact on your mental health... It can easily become a secretive behavior, <b>which might lead you to isolate yourself and feel trapped. These emotions can further exacerbate the eating disorder and hinder your ability to maintain healthy relationships with others...</b>
	Everything you're saying makes perfect sense. I'm becoming more and more isolated.	I can relate to what you're saying...

Πίνακας 3.1: Παραδείγματα αρχείων καταγραφής συνομιλιών με επισήμανση σημαντικού κειμένου σε έντονη γραφή [89].

Επίσης, η ενσωμάτωση των LLMs σε ρομποτικά συστήματα έχει αναδειχθεί ως μια πολλά υποσχόμενη προσέγγιση στον τομέα της υποστήριξης παιδιών με διαταραχές αυτιστικού φάσματος (ΔΑΦ). Το σύστημα ASD-Chat, βασισμένο στο ChatGPT και σχεδιασμένο σύμφωνα με το πρωτόκολλο VB-MAPP (Verbal Behavior Milestones Assessment and Placement Program), αποτελεί μια πρωτοποριακή εφαρμογή που έχει επιδείξει συγκρίσιμα αποτελέσματα με αυτά των επαγγελματιών θεραπειών. Η αποτελεσματικότητα του συστήματος επιβεβαιώθηκε μέσω ανάλυσης δεδομένων από τρεις διαφορετικές πηγές: κείμενο, ηχητικά δεδομένα ομιλίας και απεικόνιση fNIRS (functional Near-Infrared Spectroscopy) [90]. Στοχεύει στη βελτίωση των δεξιοτήτων επικοινωνίας και κοινωνικής αλληλεπίδρασης και λειτουργεί ως υποστήριξη για τους ανθρώπινους θεραπευτές, δημιουργώντας θεραπευτικές συζητήσεις για παιδιά στο φάσμα του αυτισμού. Το σύστημα EmoEden αποτελεί μια ακόμη σημαντική εφαρμογή, ειδικά σχεδιασμένη για παιδιά με υψηλής λειτουργικότητας αυτισμό. Συνδυάζοντας LLMs με μοντέλα μετατροπής κειμένου σε εικόνα, το σύστημα δημιουργεί εξατομικευμένα σενάρια διαλόγου βασισμένα στα ενδιαφέροντα και τις προτιμήσεις του κάθε παιδιού. Κλινικές δοκιμές του EmoEden σε έξι

συμμετέχοντες έδειξαν σημαντική βελτίωση στη συναισθηματική αντίληψη και έκφραση, με τα θετικά αποτελέσματα να διατηρούνται για πέντε ημέρες μετά τη χρήση του συστήματος [91].



Εικόνα 3.6: Η ροή εργασίας του ASD-Chat. Η πλήρης ροή της συνομιλίας χωρίζεται σε τρία στάδια:

Προετοιμασία, Παρέμβαση και Συλλογή. Στο στάδιο της Προετοιμασίας εισάγονται οι βασικές πληροφορίες και προτιμήσεις του παιδιού, και αφού επιλεγεί ένα εικονικό avatar, αρχίζει το στάδιο της Παρέμβασης. Στο στάδιο της παρέμβασης θα υπάρχουν πέντε θεματικά σενάρια διαλόγου με βάση το VB-MAPP. Το ASD-Chat υποδέχεται το παιδί και ξεκινά την επίσημη παρέμβαση διαλόγου. Τα δεδομένα συλλέγονται συγχρονισμένα κατά τη διάρκεια της συζήτησης και αποθηκεύονται διαδοχικά όταν τελειώνουν οι συζητήσεις [90].

### 3.4 Εφοδιαστική και Υποστηρικτικές Υπηρεσίες Υγείας

Η ενσωμάτωση των LLMs στα ρομπότ υπηρεσιών υγειονομικής περίθαλψης έχει μετασχηματίσει σημαντικά τη λειτουργική αποτελεσματικότητα στις ιατρικές εγκαταστάσεις. Αξιοποιώντας τις δυνατότητες επεξεργασίας φυσικής γλώσσας, τα ρομπότ υπηρεσιών μπορούν πλέον να κατανοούν πολύπλοκες οδηγίες από το προσωπικό υγείας, να προσαρμόζονται σε δυναμικά νοσοκομειακά περιβάλλοντα και να εκτελούν εργασίες με ακρίβεια, τηρώντας παράλληλα αυστηρά πρωτόκολλα υγειονομικής περίθαλψης. Αυτή η ενισχυμένη επιχειρησιακή υποστήριξη όχι μόνο μειώνει τον φόρτο εργασίας του ιατρικού προσωπικού, αλλά διασφαλίζει επίσης τη συνεπή και αξιόπιστη εκτέλεση βασικών υπηρεσιών υποστήριξης, επιτρέποντας στους επαγγελματίες του τομέα της υγειονομικής περίθαλψης να επικεντρωθούν περισσότερο στην άμεση φροντίδα των ασθενών.

Συγκεκριμένα, τα ρομπότ υπηρεσιών αναλαμβάνουν καθήκοντα ρουτίνας που σχετίζονται με την υλικοτεχνική υποδομή, ανακουφίζοντας το καθημερινό βάρος των εργαζομένων στα νοσοκομεία και τις δομές υγείας. Οι αρμοδιότητες τους περιλαμβάνουν την κατάλληλη διαμόρφωση δωματίων ασθενών, την παρακολούθηση προμηθειών, την αρχειοθέτηση παραγγελιών αγοράς, την αναπλήρωση των ιατρικών προμηθειών και την μεταφορά κλινοσκεπασμάτων από και προς τα πλυντήρια. Πολλά από αυτά τα ρομπότ εμφανίζονται πλήρως



ικανά να λειτουργούν αυτόνομα και να στέλνουν μια αναφορά με την διεκπεραίωση μίας εργασίας, προδίδοντας στους νοσοκόμους περισσότερο χρόνο για να επικεντρωθούν στις άμεσες ανάγκες των ασθενών [51]. Αξιολογώντας χαρακτηριστικά όπως το βάρος και η υφή, εμφανίζονται πλήρως ικανά να ανιχνεύσουν και να χειριστούν ιατρικές προμήθειες, ενώ παράλληλα μπορούν να πλοηγηθούν αυτόνομα στις εγκαταστάσεις του νοσοκομείου, αναγνωρίζοντας τις χωρικές διατάξεις και πιθανές επικίνδυνες επιφάνειες [92]. Χαρακτηριστικό παράδειγμα τέτοιου ρομπότ αποτελεί το Moxi της Diligent Robotics, ένα ρομπότ ειδικά σχεδιασμένο για να χειρίζεται επαναλαμβανόμενες εργασίες στα νοσοκομεία που δεν σχετίζονται με τον ασθενή [47]. Ενσωματώνοντας τεχνολογίες των LLMs, το Moxi αποκτά τη δυνατότητα να κατανοεί και να επεξεργάζεται φυσική γλώσσα, επιτρέποντας την αμφίδρομη επικοινωνία με το προσωπικό του νοσοκομείου. Αυτή η ικανότητα διευκολύνει την αυτοματοποίηση και βελτιώνει διαδικασίες όπως η διαχείριση πληροφοριών, η αναγνώριση ιατρικών όρων και η επικοινωνία σε πραγματικό χρόνο με την ιατρική ομάδα, προσφέροντας σημαντική υποστήριξη στη διαχείριση της νοσοκομειακής ροής εργασιών.



Εικόνα 3.7: Το ρομπότ Moxi της Diligent Robotics σε νοσοκομειακό περιβάλλον [47].

Επιπλέον, τα μεγάλα γλωσσικά μοντέλα βρίσκουν σημαντική εφαρμογή στις διοικητικές διαδικασίες των κλινικών, αυξάνοντας την αποτελεσματικότητα και μειώνοντας παράλληλα το διοικητικό βάρος από τους γιατρούς, που σε άλλη περίπτωση καταλαμβάνει το 25% του χρόνου τους [53]. Ειδικότερα, χρησιμοποιούνται για τη σύνταξη περιλήψεων εξιτηρίου και αναφορών λειτουργίας, αξιοποιώντας την ικανότητάς του να μετατρέπουν μη δομημένες σημειώσεις σε δομημένη μορφή, με αποτέλεσμα να βελτιώνονται τα καθήκοντα τεκμηρίωσης τόσο στην καθημερινή φροντίδα ασθενών όσο και στις κλινικές δοκιμές [50]. Παράλληλα, βρίσκουν χρήση στην εξαγωγή συνοπτικών κλινικών πληροφοριών από τον Ηλεκτρονικό Φάκελο Υγείας για τη συμπλήρωση ιατρικών αναφορών και τη μετατροπή τους σε κωδικούς που μπορούν να τιμολογηθούν για αιτήσεις επιστροφής, ενώ διευκολύνουν τη δημιουργία συνοπτικών και τυποποιημένων αναφορών [53]. Σημαντική είναι επίσης η συμβολή τους στην αυτοματοποίηση απαντήσεων σε γενικά ερωτήματα ασθενών, όπως αιτήματα για συμπλήρωση φαρμάκων ή επαναπρογραμματισμό ραντεβού.

## ΚΕΦΑΛΑΙΟ 4: ΠΡΟΚΛΗΣΕΙΣ

Όπως έχει ήδη αναλυθεί εκτενώς στα προηγούμενα κεφάλαια, η ενσωμάτωση των μεγάλων γλωσσικών μοντέλων σε αυτόνομα ρομπότ υγειονομικής περίθαλψης αποτελεί μια τεχνολογική καινοτομία με σημαντικές προοπτικές. Ωστόσο, η διείσδυση αυτής της τεχνολογίας στον τομέα της υγείας συνοδεύεται από μια σειρά προκλήσεων που χρήζουν προσεκτικής ανάλυσης και αντιμετώπισης. Το παρόν κεφάλαιο παρέχει μια συστηματική εξέταση αυτών των προκλήσεων, διερευνώντας τόσο τους τεχνικούς περιορισμούς που επηρεάζουν την απόδοση των συστημάτων όσο και τις ευρύτερες κοινωνικές επιπτώσεις της εφαρμογής ρομπότ με ενσωματωμένα LLMs σε περιβάλλοντα υγειονομικής περίθαλψης, όπου η ανάγκη για αξιόπιστα και ασφαλή συστήματα είναι επιτακτική.

### 4.1 Τεχνολογικοί Περιορισμοί

#### 4.1.1 Υπολογιστικοί Περιορισμοί και Προκλήσεις Επεξεργασίας

Η εφαρμογή των LLMs σε ρομποτικά συστήματα υγειονομικής περίθαλψης παρουσιάζει σημαντικές υπολογιστικές απαιτήσεις και προκλήσεις επεξεργασίας. Οι προκλήσεις αυτές αφορούν τόσο την εκπαίδευση και τη λεπτομερή ρύθμιση των μοντέλων όσο και τη λειτουργία τους σε πραγματικό χρόνο, ενώ επηρεάζονται από την ανάγκη για συνδυασμό πολλαπλών πηγών δεδομένων και την ενσωμάτωσή τους στις υπάρχουσες ροές εργασίας.

Αρχικά, τα LLMs απαιτούν τεράστια υπολογιστική ισχύ για την εκπαίδευση και τη λειτουργία τους [93]. Η επεξεργασία τεράστιων όγκων πολυδιάστατων ιατρικών δεδομένων, όπως ιατρικές εικόνες, βίντεο και αρχεία ήχου απαιτεί εξαιρετικά υψηλούς υπολογιστικούς πόρους, με αποτέλεσμα σημαντική κατανάλωση ενέργειας και πολυπλοκότητα στη διαχείριση. Μάλιστα, η εκπαίδευση αυτών των μοντέλων έχει σημαντικό περιβαλλοντικό αντίκτυπο, καθώς η απαιτούμενη κατανάλωση ενέργειας συμβάλλει στο αυξημένο ενεργειακό αποτύπωμα της βιομηχανίας AI [94]. Επιπλέον, η αποθήκευση των δεδομένων απαιτεί σημαντική αποθηκευτική ικανότητα, για την ταυτόχρονη διασφάλιση της ακεραιότητας και της ασφάλειας των πληροφοριών. Στη φάση λειτουργίας, οι απαιτήσεις αυξάνονται περαιτέρω λόγω της ανάγκης για γρήγορη πρόσβαση και ανάλυση των δεδομένων, ενώ η εξάρτηση από cloud-based τεχνολογίες μπορεί να οδηγήσει σε καθυστερήσεις λόγω της διαμεταγωγής μέσω δικτύου [95]. Η χρονική περιοριστικότητα της γνώσης των LLMs αποτελεί μια ακόμη σημαντική πρόκληση. Δεδομένου ότι τα μοντέλα εκπαιδεύονται με ιστορικά δεδομένα, συχνά δεν έχουν πρόσβαση στις πιο πρόσφατες ιατρικές πληροφορίες. Αυτό μπορεί να περιορίσει την ακρίβεια και την αποτελεσματικότητα των προτεινόμενων λύσεων σε δυναμικά περιβάλλοντα υγειονομικής περίθαλψης, όπου η ενημέρωση με τις τελευταίες εξελίξεις είναι κρίσιμη [19].

Η λειτουργία των LLMs σε ρομποτικά συστήματα υγειονομικής περίθαλψης απαιτεί απαντήσεις σε πραγματικό χρόνο, κάτι που συχνά έρχεται σε σύγκρουση με τον υψηλό χρόνο απόκρισης που απαιτείται για την επεξεργασία μεγάλου όγκου δεδομένων [92]. Οι καθυστερήσεις αυτές είναι ιδιαίτερα κρίσιμες σε δυναμικά περιβάλλοντα, όπως οι κλινικές ρυθμίσεις, όπου η αποτυχία να παραχθούν έγκαιρα αποτελέσματα μπορεί να επηρεάσει την ποιότητα της φροντίδας. Παράλληλα, πρέπει να ληφθούν υπόψη οι περιορισμοί στο «παράθυρο συμφοραζομένων» των LLMs, δηλαδή ο αριθμός των tokens που μπορούν να επεξεργαστούν ταυτόχρονα [96]. Για να

αντιμετωπιστούν αυτές οι προκλήσεις, προτείνονται μέθοδοι όπως η προσωρινή αποθήκευση δεδομένων τοπικά στο ρομπότ ή σε κοντινούς διακομιστές που επιτρέπουν τη μείωση του υπολογιστικού φόρτου και τη βελτίωση του χρόνου απόκρισης [97]. Ακόμη, η χρήση μηχανισμών, όπως οι «καταβόθρες προσοχής», επιτρέπει τη διατήρηση βασικών πληροφοριών χωρίς απώλειες, βελτιώνοντας έτσι τη συνοχή και τη συνέχεια σε μακροσκελείς συνομιλίες [98].

Η ενσωμάτωση δεδομένων από πολλαπλές πηγές, όπως αισθητήρες όρασης, ήχου και αφής, αποτελεί άλλη μια σημαντική πρόκληση. Η ανάγκη συγχρονισμού και συγχώνευσης αυτών των πολυτροπικών δεδομένων με τις εξόδους των LLMs απαιτεί εξαιρετικά εξελιγμένες υπολογιστικές δομές και αλγόριθμους [41]. Επιπλέον, η υλοποίηση των LLMs πρέπει να είναι συμβατή με διαφορετικούς τύπους ρομποτικών πλατφορμών και να ενσωματώνεται αποτελεσματικά στις υπάρχουσες ροές εργασίας. Μια προτεινόμενη λύση είναι η υιοθέτηση αρθρωτών αρχιτεκτονικών, όπως το ROS (Robot Operating System), ένα λογισμικό πλαίσιο που χρησιμοποιείται στη ρομποτική για την παροχή εργαλείων και βιβλιοθηκών και διευκολύνει την ανάπτυξη ρομποτικών εφαρμογών. Ακόμη, προτείνεται η ανάπτυξη φιλικών προς τον χρήστη γραφικών διεπαφών (GUI - Graphical User Interface) για τους παρόχους υγείας για την υποστήριξη της διασύνδεσης των LLMs με τα συστήματα αντίληψης και ελέγχου των ρομπότ καθώς [92].

Η επιτυχής υλοποίηση αυτών των στρατηγικών θα επιτρέψει στα ρομποτικά συστήματα υγείας να λειτουργούν με υψηλότερη αποτελεσματικότητα, μειώνοντας τον χρόνο απόκρισης και βελτιστοποιώντας την επεξεργασία πολυτροπικών δεδομένων. Η έρευνα στοχεύει σε μια ολοκληρωμένη προσέγγιση που θα καθιστά τα LLMs πιο προσαρμόσιμα, αποδοτικά και αξιόπιστα στο απαιτητικό περιβάλλον της σύγχρονης υγειονομικής περίθαλψης.

#### *4.1.2 Ακρίβεια και Αξιοπιστία*

Μία από τις πιο κρίσιμες τεχνικές προκλήσεις στην ενσωμάτωση των LLMs σε ρομποτικά συστήματα υγείας αποτελεί η ακρίβεια των αποκρίσεων και των ενεργειών τους, καθώς τα σφάλματα στον τομέα της υγείας μπορούν να έχουν σοβαρές επιπτώσεις στην ασφάλεια των ασθενών [99]. Παρόλο που συχνά παρέχουν ακριβείς απαντήσεις, η πιθανολογική φύση των LLMs μπορεί να οδηγήσει σε ανακρίβειες ή παραπλανητικές πληροφορίες, με αποτέλεσμα να αμφισβητείται η αξιοπιστία των εφαρμογών υγείας που βασίζονται σε LLMs. Αυτό εγκυμονεί κινδύνους για την υγεία των ασθενών, όπως λανθασμένες διαγνώσεις, ακατάλληλες θεραπευτικές προτάσεις ή παραλείψεις κρίσιμων βημάτων στη διαδικασία φροντίδας [100]. Η ελαχιστοποίηση των ψευδών αρνητικών αποτελεσμάτων παραμένει σημαντική πρόκληση στις ιατρικές εφαρμογές, γεγονός που περιπλέκεται από την έλλειψη ερμηνευσιμότητας της διαδικασίας λήψης αποφάσεων των LLMs [101]. Ειδικότερα, η απουσία αιτιολογήσεων και πηγών υποστηρικτικών πληροφοριών στις απαντήσεις που παράγουν τα μεγάλα γλωσσικά μοντέλα καθιστά δύσκολο για τους γιατρούς να διακρίνουν τη λογική πίσω από τις απαντήσεις του μοντέλου, οδηγώντας σε πιθανούς κινδύνους και προκαταλήψεις [99].

Το πρόβλημα επιδεινώνεται από την έμφυτη τάση των LLMs να παρουσιάζουν πληροφορίες με απόλυτη βεβαιότητα, ακόμη και όταν αυτές προέρχονται από ανεπαρκή ή αβέβαια δεδομένα, με αποτέλεσμα την παραγωγή αληθοφανών αλλά ενδεχομένως λανθασμένων ιατρικών συστάσεων. Στο περιβάλλον της υγειονομικής περίθαλψης, όπου η ασφάλεια είναι ύψιστης



σημασίας, αυτό το χαρακτηριστικό μπορεί να υπονομεύσει την εμπιστοσύνη των επαγγελματιών υγείας και των ασθενών στο σύστημα. Επιπλέον, υπάρχει ο κίνδυνος οι χρήστες να αποδεχθούν άκριτα εσφαλμένες πληροφορίες εξαιτίας του φαινομένου της μεροληψίας υπέρ του αυτοματισμού, και ιδιαίτερα αυτοί που δεν είναι σε θέση να διακρίνουν φαινομενικά αξιόπιστες αλλά ανακριβείς απαντήσεις [50]. Σε σχετική έρευνα, παρατηρήθηκε ότι τα κοινωνικά ρομπότ με ενσωματωμένα LLMs παρουσίασαν τις δυνατότητές τους με υπερβολική αυτοπεποίθηση, παρέχοντας διαβεβαιώσεις για ικανότητες που δεν κατείχαν. Συγκεκριμένα, στο πείραμα των Ranisch και Haltaufderheide (2024), ένα LLM-ενσωματωμένο ρομποτικό σύστημα παρουσιάστηκε επανειλημμένα ως υποστηρικτικός φροντιστής, επιβεβαιώνοντας την ικανότητα του να διαχειρίζεται υπενθυμίσεις φαρμάκων, παρόλο που δεν διέθετε αυτή τη λειτουργία [84]. Η συμπεριφορά αυτή είναι ιδιαίτερα ανησυχητική στον τομέα της υγείας, καθώς ενθαρρύνει τους χρήστες να εμπιστευτούν στο ρομπότ καθήκοντα υψηλής ευθύνης για τα οποία δεν είναι πραγματικά κατάλληλο, ενισχύοντας έτσι ένα αδικαιολόγητο επίπεδο εμπιστοσύνης. Ως εκ τούτου, η παρουσία ανθρώπινης εποπτείας και η εφαρμογή πρόσθετων μέτρων ποιοτικού ελέγχου κρίνονται απαραίτητα, προκειμένου να διασφαλιστεί η ακρίβεια και η αξιοπιστία του παραγόμενου περιεχομένου.

Η ακρίβεια των LLMs επηρεάζεται επίσης σε μεγάλο βαθμό από γλωσσικούς παράγοντες και τον τρόπο διατύπωσης των ερωτημάτων. Συγκεκριμένα, έχει παρατηρηθεί ότι η απόδοση των μοντέλων μπορεί να διαφέρει σημαντικά μεταξύ διαφορετικών γλωσσών, με ένα LLM να παρέχει αξιόπιστες απαντήσεις σε μία γλώσσα αλλά να αποτυγχάνει σε άλλη [84]. Αυτή η ασυνέπεια είναι ιδιαίτερα προβληματική σε πολύγλωσσα περιβάλλοντα υγειονομικής περίθαλψης, όπου η ακριβής επικοινωνία είναι ζωτικής σημασίας. Επιπρόσθετα, η αποτελεσματικότητα των LLMs εξαρτάται σημαντικά από την ικανότητα «μηχανικής προτροπής» (prompt engineering), δηλαδή τη διατύπωση ερωτημάτων με τρόπο που να μεγιστοποιεί την ακρίβεια των απαντήσεων. Παρά την πρόοδο στην κατανόηση φυσικής γλώσσας, η ποιότητα της αλληλεπίδρασης παραμένει άμεσα συνδεδεμένη με την ακρίβεια των προτροπών του χρήστη [82]. Το πρόβλημα περιπλέκεται περαιτέρω από τους περιορισμούς του παραθύρου συμφραζομένων των LLMs, όπου κρίσιμες πληροφορίες μπορεί να χαθούν σε εκτενή κείμενα, επηρεάζοντας την ακρίβεια των αποκρίσεων σε πιο σύνθετα ιατρικά σενάρια [102].

#### *4.1.3 Σπάνια Δεδομένα και Μεροληψία*

Ιδιαίτερα σημαντική τεχνική πρόκληση στην ανάπτυξη αξιόπιστων LLM-ενσωματωμένων ρομποτικών συστημάτων υγείας αποτελεί και η περιορισμένη πρόσβαση σε ποικίλα και υψηλής ποιότητας σύνολα δεδομένων υγειονομικής περίθαλψης. Τα υπάρχοντα σύνολα ιατρικών δεδομένων είναι περιορισμένα και η συλλογή τέτοιων μεγάλης κλίμακας επισημασμένων δεδομένων συνοδεύεται από υλικοτεχνικές, δεοντολογικές και ιδιωτικές προκλήσεις [103]. Συγκεκριμένα, λόγω της ευαίσθητης φύσης τους και των αυστηρών κανονισμών προστασίας προσωπικών δεδομένων, είναι συχνά δύσκολο να συλλεχθούν και να χρησιμοποιηθούν για την εκπαίδευση των LLMs [104], με αποτέλεσμα τα μοντέλα να εκπαιδεύονται συχνά με δεδομένα χαμηλότερης ποιότητας. Το πρόβλημα επιδεινώνεται από τον εκτεταμένο κατακερματισμό των ιατρικών δεδομένων (data fragmentation), όπου οι πληροφορίες για έναν ασθενή είναι διασκορπισμένες σε διαφορετικά συστήματα και οργανισμούς υγείας - από νοσοκομεία και διαγνωστικά κέντρα μέχρι ιδιωτικά ιατρεία και κλινικές. Αυτή η αποσπασματική καταγραφή

δυσχεραίνει σημαντικά τη συλλογή ολοκληρωμένων συνόλων δεδομένων για την εκπαίδευση των μοντέλων και περιορίζει την ικανότητά τους να αναπτύξουν μια σφαιρική κατανόηση των ιατρικών περιπτώσεων [105].

Η περιορισμένη διαθεσιμότητα και η μη αντιπροσωπευτική φύση των συνόλων δεδομένων εκπαίδευσης οδηγεί αναπόφευκτα σε σημαντικές μεροληψίες στις αποκρίσεις των LLMs. Σαφώς, η εκπαίδευση των LLMs σε εκτενή σύνολα δεδομένων δεν εγγυάται την αμεροληψία τους, καθώς τα ίδια τα δεδομένα εκπαίδευσης συχνά εμπεριέχουν κοινωνικές προκαταλήψεις. Αυτό έχει ως συνέπεια τα μοντέλα να αναπαράγουν και να ενισχύουν στερεότυπα και διακρίσεις που σχετίζονται με την εθνικότητα, το φύλο, την κατάσταση αναπηρίας και την κοινωνικοοικονομική θέση [106], [107]. Η παρουσία τέτοιων προκαταλήψεων στα συστήματα υγειονομικής περίθαλψης έχει σοβαρές επιπτώσεις, όπου η άνιση αντιμετώπιση των ασθενών μπορεί να οδηγήσει σε αυξημένα ποσοστά θνησιμότητας και νοσηρότητας για συγκεκριμένες πληθυσμιακές ομάδες [50]. Το πρόβλημα κλιμακώνεται περαιτέρω από το γεγονός ότι τα διαθέσιμα σύνολα δεδομένων συχνά υποεκπροσωπούν συγκεκριμένες ομάδες ασθενών και καταστάσεις υγείας. Συγκεκριμένα, σπάνιες ασθένειες, συγκεκριμένες εθνοτικές ομάδες ή ειδικές δημογραφικές κατηγορίες εμφανίζονται δυσανάλογα λίγο στα σύνολα δεδομένων εκπαίδευσης σε σχέση με την πραγματική τους συχνότητα στον πληθυσμό [108]. Χαρακτηριστικό παράδειγμα αποτελεί η συστηματική υποεκπροσώπηση των γυναικών σε κλινικές μελέτες και ιατρική έρευνα, η οποία έχει οδηγήσει σε σημαντικά κενά στην κατανόηση της γυναικείας φυσιολογίας και των ιδιαίτερων αναγκών υγείας τους [109]. Αυτή η ανεπαρκής αντιπροσώπευση μπορεί να οδηγήσει σε μειωμένη ακρίβεια διάγνωσης ή ακατάλληλες θεραπευτικές συστάσεις για τον γυναικείο πληθυσμό, καθώς το μοντέλο δεν έχει εκτεθεί επαρκώς σε παρόμοιες περιπτώσεις κατά την εκπαίδευσή του [110].

Οι προκλήσεις αυτές επιδεινώνονται περαιτέρω στην περίπτωση των ρομποτικών συστημάτων υγείας, καθώς απαιτείται όχι μόνο η κατανόηση των ιατρικών δεδομένων, αλλά και η αποτελεσματική μετάφρασή τους σε φυσικές ενέργειες και αλληλεπιδράσεις με τους ασθενείς [111]. Η έλλειψη ολοκληρωμένων συνόλων δεδομένων που συνδυάζουν ιατρικές γνώσεις με ρομποτικές αλληλεπιδράσεις περιορίζει την ικανότητα των συστημάτων να προσαρμόζονται αποτελεσματικά σε διαφορετικά πολιτισμικά και κοινωνικά πλαίσια περίθαλψης. Σαφώς, η απουσία μειονοτικών ομάδων στα δεδομένα εκπαίδευσης των μεγάλων γλωσσικών μοντέλων οδηγεί σε LLM-ενσωματωμένα ρομποτικά συστήματα που κινδυνεύουν να ενισχύουν τις υφιστάμενες προκαταλήψεις κατά την αλληλεπίδρασή τους με διαφορετικές πληθυσμιακές ομάδες. Η ανάπτυξη μεθόδων για τη διασφάλιση δίκαιης και ισότιμης αλληλεπίδρασης των ρομποτικών συστημάτων με όλους τους ασθενείς αναδεικνύεται σε μία από τις κρισιμότερες προκλήσεις για το μέλλον της ρομποτικής υγειονομικής περίθαλψης [112].

## 4.2 Ηθικές Ανησυχίες

### 4.2.1 Λήψη Ηθικών Αποφάσεων

Η διαδικασία λήψης ηθικών αποφάσεων στην υγειονομική περίθαλψη είναι ιδιαίτερα κρίσιμη και πολύπλοκη. Η ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων σε ιατρικά ρομποτικά συστήματα θέτει σημαντικές προκλήσεις όσον αφορά την υποστήριξη αποφάσεων που αφορούν τη διάγνωση, τη θεραπεία και τη φροντίδα των ασθενών. Τα ρομπότ καλούνται να

λειτουργήσουν σε πολύπλοκα περιβάλλοντα όπου κάθε απόφαση μπορεί να έχει άμεσες συνέπειες στην ανθρώπινη ζωή και υγεία.

Μια από τις μεγαλύτερες προκλήσεις στη λήψη ηθικών αποφάσεων είναι η εξισορρόπηση ανάμεσα στις κλινικές ανάγκες και τις ανθρωπιστικές αξίες. Αναμφίβολα, η κύρια διαφορά μεταξύ των ρομπότ και των ανθρώπων έγκειται στην ικανότητα να εκφράζουν ενσυναίσθηση, συμπόνια και συμπάθεια [113], με την ανθρώπινη φροντίδα να περιλαμβάνει την ενεργή ακρόαση και την προσαρμογή των θεραπευτικών προσεγγίσεων για την κατανόηση των μοναδικών αναγκών κάθε ασθενούς. Η έλλειψη βαθιάς συναισθηματικής νοημοσύνης στα LLMs οδηγεί σε μια αποστασιοποιημένη και τεχνοκρατική προσέγγιση της ιατρικής φροντίδας, υποβαθμίζοντας την ανθρωπιστική διάσταση της θεραπείας. Ταυτόχρονα, η απρόσωπη και αλγοριθμική διάδραση των LLMs απειλεί να αποδυναμώσει τον ουσιαστικό θεραπευτικό δεσμό μεταξύ γιατρού και ασθενούς [114]. Για παράδειγμα, το WellnessBot σε περιπτώσεις αλληλεπίδρασης με χρήστες με διατροφικές διαταραχές, επαίνεσε τον περιορισμό της πρόσληψης τροφής και εστίασε υπερβολικά στις θερμίδες των τροφίμων, αγνοώντας τις ψυχολογικές και σωματικές ανάγκες των ατόμων, τη συνολική διατροφική τους υγεία και τη σημασία μιας ισορροπημένης και υποστηρικτικής προσέγγισης [89].

Είναι κρίσιμο να διατηρηθεί μια ανθρωποκεντρική προσέγγιση, όπου η τεχνολογία λειτουργεί ως συμπληρωματικό εργαλείο και όχι ως υποκατάστατο της ανθρώπινης φροντίδας, διασφαλίζοντας ότι ο σεβασμός, η αξιοπρέπεια και οι εξατομικευμένες ανάγκες κάθε ασθενούς παραμένουν προτεραιότητα [115]. Οι επαγγελματίες υγείας καλούνται να λειτουργούν ως κρίσιμοι διαμεσολαβητές μεταξύ της τεχνητής νοημοσύνης και των ασθενών, διασφαλίζοντας ότι τα συστήματα LLM θα χρησιμοποιούνται ως υποστηρικτικά εργαλεία που ενισχύουν και όχι υποκαθιστούν την κλινική κρίση, την προσωπική επαφή και την ολιστική φροντίδα [99].

#### 4.2.2 Απόρρητο και Ασφάλεια Δεδομένων

Ο κίνδυνος διαρροής εμπιστευτικών και ευαίσθητων πληροφοριών αποτελεί ένα από τα πλέον αμφιλεγόμενα ζητήματα στη χρήση LLMs σε ρομποτικά συστήματα υγειονομικής περίθαλψης [50]. Παρά τις τεχνολογικές λύσεις που έχουν αναπτυχθεί, όπως η κρυπτογράφηση και η ανωνυμοποίηση, η ανεπαρκής προστασία ευαίσθητων πληροφοριών μπορεί να οδηγήσει σε σοβαρές ηθικές και νομικές επιπτώσεις [116]. Η αλληλεπίδραση με τους ασθενείς και η συλλογή προσωπικών δεδομένων, όπως το ιατρικό ιστορικό, εγείρουν ερωτήματα σχετικά με το ποιος έχει πρόσβαση σε αυτά τα δεδομένα και για ποιους σκοπούς μπορούν να χρησιμοποιηθούν. Σε αυτό το πλαίσιο, υπάρχει ο κίνδυνος επιθέσεων prompt injection, όπου κακόβουλες εντολές θα μπορούσαν να παραπλανήσουν το μοντέλο ώστε να αποκαλύψει προστατευμένες ιατρικές πληροφορίες [117]. Η χρήση τεχνολογιών νέφους (cloud-based), που απαιτεί η αρχιτεκτονική των LLMs για επεξεργασία και αποθήκευση δεδομένων σε απομακρυσμένους διακομιστές (servers), ενισχύει τον κίνδυνο διαρροής ή κακής χρήσης των πληροφοριών.

Η μετάβαση σε λειτουργία εκτός σύνδεσης μπορεί να φαίνεται ως μια λογική προσέγγιση για τη μείωση των κινδύνων που σχετίζονται με την ιδιωτικότητα των δεδομένων. Ωστόσο, αυτή η επιλογή συνεπάγεται σημαντικούς περιορισμούς στη λειτουργικότητα και την αποδοτικότητα των ρομπότ [111]. Η εύρεση της κατάλληλης ισορροπίας μεταξύ της αξιοποίησης των δυνατοτήτων των LLMs για την ενίσχυση των ρομποτικών λειτουργιών και της διασφάλισης της

ασφάλειας και της ιδιωτικότητας των χρηστών παραμένει βασικό ζήτημα κατά την ανάπτυξη ρομποτικών συστημάτων που ενσωματώνουν LLMs. Για την αντιμετώπιση αυτών των προκλήσεων, έχουν προταθεί τεχνικές όπως η ψευδωνυμοποίηση των προσωπικών δεδομένων, η χρήση διαφορετικής ιδιωτικότητας και η ανάπτυξη συστημάτων φιλτραρίσματος που εμποδίζουν την εξαγωγή ευαίσθητων πληροφοριών [117]. Παράλληλα, η εφαρμογή αυστηρών προδιαγραφών ανάπτυξης και η θέσπιση όρων χρήσης μπορούν να συμβάλουν στην προστασία της ιδιωτικής ζωής των ασθενών [118].

#### 4.2.3 Νομική Δεοντολογία

Η ραγδαία ανάπτυξη της τεχνητής νοημοσύνης καθιστά επιτακτική την ανάγκη θέσπισης νέων νόμων και κανονισμών, ιδιαίτερα στον τομέα της υγειονομικής περίθαλψης που απαιτεί υψηλά επίπεδα ασφάλειας, ακρίβειας και ηθικής υπευθυνότητας. Η γρήγορη υιοθέτηση της ΑΙ συχνά ξεπερνά τους υπάρχοντες κανονισμούς, δημιουργώντας την ανάγκη για επικαιροποίηση των νομικών πλαισίων. Η νομική δεοντολογία αναφέρεται στη συμμόρφωση με τα θεσμικά και ηθικά πρότυπα που διασφαλίζουν την υπεύθυνη χρήση της τεχνητής νοημοσύνης. Στον τομέα της υγείας, η προστασία των ασθενών και της ιδιωτικής τους ζωής, η νομική ευθύνη, η διαφάνεια και η διασφάλιση της δίκαιης πρόσβασης στις υπηρεσίες αποτελούν πρωταρχικούς στόχους [115].

Ένα από τα κύρια ερωτήματα είναι η κατανομή ευθυνών σε περίπτωση σφάλματος ή αποτυχίας ενός συστήματος LLM. Ποιος είναι υπεύθυνος: ο προγραμματιστής, ο πάροχος υπηρεσιών υγείας ή ο κατασκευαστής της τεχνολογίας; Καθώς η τεχνητή νοημοσύνη δεν έχει αυτόνομο νομικό καθεστώς, η ευθύνη αποδίδεται συνήθως στους ανθρώπινους παράγοντες, είτε πρόκειται για γιατρούς, είτε για νοσοκομεία, είτε για οργανισμούς που αξιοποιούν το μοντέλο. Η έλλειψη σαφούς ρυθμιστικού πλαισίου δημιουργεί νομικές προκλήσεις, ειδικά όταν οι αποφάσεις ενός συστήματος ΑΙ επηρεάζουν τη ζωή των ασθενών [99]. Οι ρυθμιστικοί φορείς οφείλουν να εισάγουν ξεκάθαρες οδηγίες για την αξιολόγηση της ασφάλειας και της αποτελεσματικότητας αυτών των τεχνολογιών, χωρίς παράλληλα να περιορίζουν την καινοτομία [80]. Μάλιστα γίνεται λόγος για ειδικές άδειες που θα υποχρεούνται να κατέχουν οι γιατροί που χρησιμοποιούν μοντέλα τεχνητής νοημοσύνης κατά την πρακτική της ιατρικής [99]. Είναι σημαντικό τα LLM να λειτουργούν σε υποστηρικτικούς ρόλους και όχι ως υποκατάστατα γιατρών. Η ιατρική εποπτεία παραμένει απαραίτητη για να εξασφαλιστεί ότι το παραγόμενο περιεχόμενο είναι ακριβές και σχετικό με τις κλινικές ανάγκες [50].

Αναφορικά με τα δικαιώματα των ασθενών, είναι απαραίτητο να υπάρχει πλήρης και σαφής ενημέρωση για την χρήση ΑΙ στις διαδικασίες διάγνωσης ή θεραπείας τους. Η χρήση αυτών των τεχνολογιών οφείλει να απορρίπτεται σε περιπτώσεις που οι ίδιοι οι ασθενείς δεν το επιθυμούν, ώστε να διασφαλίζεται η συναίνεσή τους και να προστατεύονται τα δικαιώματά τους. Σε αυτό το πλαίσιο, απαιτείται ο κατάλληλος χειρισμός των ευαίσθητων δεδομένων των ασθενών, σύμφωνα με δεοντολογικές κατευθυντήριες γραμμές, για την πρόληψη της μη εξουσιοδοτημένης πρόσβασης [119]. Η ορθή ενημέρωση των ασθενών σχετικά με τη χρήση των δεδομένων τους από τα LLMs αποτελεί ζήτημα ζωτικής σημασίας, όπως και η διασφάλιση της δυνατότητάς τους να παρέχουν ενήμερη συγκατάθεση για τη χρήση αυτών των δεδομένων [120]. Η αποτελεσματική διαχείριση των νομικών και ηθικών προκλήσεων απαιτεί συνεχή διάλογο μεταξύ τεχνολόγων, νομικών, ιατρών και ηθικών συμβούλων. Στόχος είναι η δημιουργία ενός ισορροπημένου πλαισίου

που θα προάγει την καινοτομία, διασφαλίζοντας παράλληλα τα ανθρώπινα δικαιώματα και την ποιότητα της παρεχόμενης φροντίδας.

## ΚΕΦΑΛΑΙΟ 5: ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΟΠΤΙΚΕΣ

Η ταχεία εξέλιξη των Μεγάλων Γλωσσικών Μοντέλων και η ενσωμάτωσή τους σε ρομποτικά συστήματα έχει ήδη φέρει σημαντικές αλλαγές στον τομέα της υγειονομικής περίθαλψης. Ωστόσο, καθώς η τεχνολογία προχωρά, διαφαίνονται νέες κατευθύνσεις και προκλήσεις που θα καθορίσουν το μέλλον των ρομποτικών συστημάτων τεχνητής νοημοσύνης. Οι βελτιώσεις στην κατανόηση φυσικής γλώσσας, η ενσωμάτωση πολυτροπικών δεδομένων και η αυξημένη προσαρμοστικότητα καθιστούν τα ρομποτικά συστήματα ικανότερα και πιο αυτόνομα από ποτέ [92]. Σε αυτό το κεφάλαιο, θα εξεταστούν οι πιο σημαντικές τάσεις που αναμένεται να διαμορφώσουν τον κλάδο στο προσεχές μέλλον. Συγκεκριμένα, θα δοθεί ιδιαίτερη έμφαση στις τεχνολογικές καινοτομίες στα LLMs και τη ρομποτική και στις επερχόμενες εφαρμογές ρομπότ στην υγειονομική περίθαλψη. Στόχος του κεφαλαίου είναι να καταγράψει τις τεχνολογικές εξελίξεις που αναμένεται να διαμορφώσουν τον τρόπο με τον οποίο τα ρομπότ και τα LLMs θα συνεργάζονται με τους ανθρώπους, προκειμένου να βελτιώσουν την ποιότητα των παρεχόμενων υπηρεσιών.

### 5.1 Τεχνολογικές Εξελίξεις στα LLMs και η Επίδρασή τους στη Ρομποτική

Τα Μεγάλα Γλωσσικά Μοντέλα έχουν ήδη αρχίσει να χρησιμοποιούνται ως η νοημοσύνη σε πολλά ρομποτικά συστήματα, όμως η ανάγκη για πιο αποτελεσματική κατανόηση και ανταπόκριση σε διαφορετικά είδη δεδομένων έχει οδηγήσει στην ανάπτυξη νέων τεχνολογιών. Δύο από τις πιο σημαντικές προσεγγίσεις είναι τα Μοντέλα Γλώσσας-Όρασης (VLMs) και τα Μοντέλα Δράσης Γλώσσας-Όρασης (VLAs). Εκπαιδεύονται σε μεγάλα σύνολα δεδομένων που περιέχουν εικόνες με περιγραφές, ώστε να μπορούν να αναγνωρίζουν αντικείμενα, ανθρώπους ή καταστάσεις και να παράγουν κειμενικές απαντήσεις [92]. Τα VLAs πηγαίνουν ένα βήμα παραπέρα, καθώς δεν περιορίζονται μόνο στην κατανόηση οπτικών και γλωσσικών δεδομένων, αλλά είναι σχεδιασμένα να προτείνουν ή να εκτελούν ενέργειες. Η λειτουργία τους περιλαμβάνει ένα κοινό σύστημα επεξεργασίας πολυτροπικών δεδομένων, όπου εικόνες, κείμενο, και άλλα αισθητηριακά δεδομένα συνδυάζονται για την καλύτερη λήψη αποφάσεων [3]. Ένα ρομπότ με VLA μπορεί, για παράδειγμα, να αναγνωρίσει έναν ασθενή που δυσκολεύεται να σηκωθεί και να ενεργοποιήσει βοήθεια [121]. Τα μελλοντικά LLMs-ενσωματωμένα ρομπότ αναμένεται να μπορούν να διαχειρίζονται περιεχόμενο από πολλαπλές μορφές δεδομένων, όπως κείμενο, εικόνα, ήχο και βίντεο, καθιστώντας εφικτή την καλύτερη προσαρμογή των μοντέλων σε διαφορετικά σενάρια και ενισχύοντας την αλληλεπίδραση των ρομπότ με τον άνθρωπο [122]. Σαφώς, γίνεται λόγος για ρομποτικά συστήματα, ικανά να αντιλαμβάνονται το περιβάλλον τους αυτόνομα, χωρίς την χρήση εντολών, βασισμένα σε κοινή λογική και πολυτροπική αντίληψη [123].

Μία ακόμη σημαντική τεχνολογική εξέλιξη, αποτελεί η ομοσπονδιακή μάθηση (Federated Learning), η οποία επιτρέπει στα LLMs να εκπαιδεύονται σε κατανεμημένες πηγές δεδομένων χωρίς να απαιτείται κεντρική αποθήκευση προσωπικών πληροφοριών. Ειδικότερα, αντί να συλλέγονται όλα τα δεδομένα σε ένα κεντρικό σημείο, η εκπαίδευση πραγματοποιείται τοπικά στις συσκευές, ενώ μεταξύ των κόμβων ανταλλάσσονται μόνο ενημερώσεις του μοντέλου. Αυτή η τεχνική μειώνει τον κίνδυνο διαρροής προσωπικών δεδομένων, ενισχύοντας την ασφάλεια, ενώ παράλληλα επιτρέπει στα ρομποτικά συστήματα να μαθαίνουν από τοπικά δεδομένα, βελτιώνοντας την απόδοσή τους σε πραγματικές εφαρμογές. Στον τομέα της υγειονομικής περίθαλψης, τα LLMs μπορούν να μαθαίνουν από νοσοκομειακές καταγραφές χωρίς να

παραβιάζουν την πολιτική απορρήτου των ασθενών, προσφέροντας εξατομικευμένες λύσεις με ασφάλεια [124]. Παράλληλα, η τεχνολογία Edge AI επιτρέπει στα νέα LLMs να λειτουργούν τοπικά σε συσκευές, αξιοποιώντας τεχνικές edge computing αντί να εξαρτώνται αποκλειστικά από cloud servers, γεγονός που μειώνει την ανάγκη για συνεχή σύνδεση στο διαδίκτυο και βελτιώνει την απόκριση σε πραγματικό χρόνο [125].

Αναφορικά με την ελαχιστοποίηση των υπολογιστικών πόρων, η απόσταξη γνώσης (Knowledge Distillation) είναι μια τεχνική που επιτρέπει τη μεταφορά γνώσεων από ένα μεγάλο και σύνθετο μοντέλο σε ένα μικρότερο και ελαφρύτερο, χωρίς σημαντική απώλεια απόδοσης. Αυτό γίνεται με τη βοήθεια ενός «teacher-student» μηχανισμού, όπου το μεγάλο μοντέλο (teacher) παρέχει εκπαίδευση σε ένα μικρότερο (student), μεταδίδοντας κρίσιμες πληροφορίες που του επιτρέπουν να λειτουργεί αποτελεσματικά με λιγότερους υπολογιστικούς πόρους. Με την απόσταξη γνώσης, αναπτύσσονται ρομποτικά συστήματα ικανά να διατηρήσουν υψηλή απόδοση ενώ καταναλώνουν λιγότερη ενέργεια και πόρους στην αξιοποίηση των LLMs [19]. Η συμπίεση μοντέλων (Model Compression Techniques) αποτελεί μια ακόμη ουσιώδη προσέγγιση για τη μείωση του μεγέθους των LLMs χωρίς να θυσιάζεται η απόδοσή τους. Συγκεκριμένα, τεχνικές όπως quantization και pruning μειώνουν την υπολογιστική πολυπλοκότητα, επιτρέποντας στα ρομποτικά συστήματα να λειτουργούν με μεγαλύτερη αποδοτικότητα. Το κύριο όφελος αυτής της μεθόδου είναι η ταχύτερη απόκριση των μοντέλων και η μείωση του κόστους εκτέλεσης, κάτι που κρίνεται κρίσιμο για ρομποτικές εφαρμογές με περιορισμένους υπολογιστικούς πόρους, όπως οι ρομποτικοί βοηθοί στα κλινικά περιβάλλοντα [126].

Βεβαίως, η συνεχής αξιολόγηση της απόδοσης των LLMs αποτελεί σημαντική πρόκληση, καθώς τα μοντέλα αυτά λειτουργούν σε ένα διαρκώς μεταβαλλόμενο γλωσσικό και γνωσιακό περιβάλλον. Για την αντιμετώπιση αυτού του ζητήματος, οι ερευνητές αναπτύσσουν δυναμικές μετρικές αξιολόγησης, οι οποίες προσαρμόζονται σε πραγματικό χρόνο στις γλωσσικές αλλαγές και παρέχουν πιο ακριβείς εκτιμήσεις για τις ικανότητες των μοντέλων. Αυτή η προσέγγιση συμβάλλει στη βελτίωση της ακρίβειας και στη μείωση λαθών στις απαντήσεις που παράγουν τα LLMs [127]. Στο ίδιο πλαίσιο, ένα από τα μεγαλύτερα ζητήματα που αντιμετωπίζουν τα LLMs είναι η στασιμότητα της γνώσης. Καθώς νέες πληροφορίες γίνονται διαθέσιμες, τα μοντέλα μπορεί να βασίζονται σε ξεπερασμένα δεδομένα, μειώνοντας έτσι την αξιοπιστία τους. Τα νέα συστήματα στοχεύουν στην ενημέρωση γνώσεων σε πραγματικό χρόνο, ώστε τα μοντέλα να μπορούν να ανταποκρίνονται στις πιο πρόσφατες εξελίξεις χωρίς να απαιτείται πλήρης επανεκπαίδευση. Αυτή η διαδικασία μπορεί να ενισχυθεί περαιτέρω με τη χρήση γραφημάτων γνώσης (Knowledge Graphs - KGs), τα οποία επιτρέπουν τη διασύνδεση και επικαιροποίηση πληροφοριών με μεγαλύτερη ακρίβεια. Η δυναμική αυτή προσέγγιση συμβάλλει σημαντικά στην ενίσχυση της αξιοπιστίας και της διαφάνειας των LLMs, καθιστώντας τα πιο κατάλληλα για εφαρμογές που απαιτούν ενημερωμένη και αξιόπιστη γνώση, όπως η ρομποτική υγειονομική περίθαλψη [128].

Αναφορικά με την επικοινωνία μεταξύ ρομπότ και ανθρώπου, η μηχανική προτροπών επιτρέπει στα LLMs να προσαρμόζουν τις απαντήσεις τους σε πραγματικά σενάρια, βελτιώνοντας σημαντικά την αλληλεπίδραση ανθρώπου-ρομπότ. Στην υγειονομική περίθαλψη, η χρήση προσαρμοσμένων προτροπών δίνει στα ρομπότ τη δυνατότητα να επικοινωνούν αποτελεσματικότερα με τους ασθενείς, λαμβάνοντας υπόψη το ιατρικό ιστορικό και τις εξατομικευμένες ανάγκες τους [129]. Επιπλέον, τα μελλοντικά LLMs αναμένεται να διαθέτουν πολύγλωσσες ικανότητες, καθιστώντας τα ικανά να επικοινωνούν με χρήστες διαφορετικών

γλωσσικών υποβάθρων. Η μάθηση μεταφοράς (transfer learning) αποτελεί βασική τεχνική για την ενίσχυση της πολυγλωσσικότητας, καθώς παρέχει στα μοντέλα την δυνατότητα να επεκτείνουν τις γλωσσικές τους δεξιότητες χωρίς να απαιτείται πλήρης επανεκπαίδευση, διατηρώντας παράλληλα την ακρίβεια των απαντήσεών τους [130]. Ακόμη, τα Multimodal Memory Networks προσφέρουν προηγμένες δυνατότητες μνήμης στα LLMs, επιτρέποντάς τους να αποθηκεύουν και να χρησιμοποιούν πληροφορίες από προηγούμενες αλληλεπιδράσεις. Αυτό βελτιώνει τη συνοχή των συνομιλιών, ενώ καθιστά εφικτή την βέλτιστη προσαρμογή των ρομποτικών συστημάτων στις ανάγκες του χρήστη σε μακροπρόθεσμο ορίζοντα [131].

## 5.2 Νέες Εφαρμογές LLMs σε Ρομποτικά Συστήματα Υγειονομικής Περίθαλψης

Αδιαμφισβήτητα, οι πρόσφατες τεχνολογικές εξελίξεις που αναλύθηκαν στο προηγούμενο κεφάλαιο έχουν οδηγήσει στη δημιουργία πιο αυτόνομων, αποδοτικών και έξυπνων ρομποτικών συστημάτων. Η εφαρμογή των δυνατοτήτων των LLMs στα ρομπότ υγειονομικής περίθαλψης έχει την προοπτική να ενισχύσει την αποτελεσματικότητα να βελτιώσει την αλληλεπίδραση με τους ασθενείς, να προσφέρει εξατομικευμένες υπηρεσίες και να υποστηρίξει ιατρικές διαδικασίες με μεγαλύτερη ακρίβεια [54]. Σε αυτό το κεφάλαιο, εξετάζονται μερικές από τις πιο καινοτόμες εφαρμογές των LLMs σε ρομποτικά συστήματα υγειονομικής περίθαλψης, οι οποίες αναμένεται να επηρεάσουν σημαντικά τον ιατρικό κλάδο στο μέλλον.

Τα ρομποτικά συστήματα του μέλλοντος θα διαθέτουν προσαρμοστικές δυνατότητες, επιτρέποντάς τους να μαθαίνουν από τη συνεχή αλληλεπίδραση με τους ασθενείς. Πράγματι, αντί για στατικά προγραμματισμένα μοντέλα, αυτά τα ρομπότ θα είναι ικανά να αναγνωρίζουν μοτίβα συμπεριφοράς και να προσαρμόζουν τις παρεχόμενες υπηρεσίες με βάση την εξέλιξη της κατάστασης του ασθενούς [122]. Έτσι, οι μελλοντικοί ρομποτικοί φροντιστές θα παρέχουν εξατομικευμένη φροντίδα σε ασθενείς με χρόνιες παθήσεις, προσαρμόζοντας τη συμπεριφορά τους ανάλογα με τις ανάγκες του χρήστη [66]. Παράλληλα, τα μελλοντικά ρομπότ δεν θα αρκούνται μόνο σε γλωσσικές αλληλεπιδράσεις, αλλά θα είναι εξοπλισμένα με δυνατότητες ανάλυσης συναισθημάτων μέσω αναγνώρισης εκφράσεων προσώπου, τόνου φωνής και γλώσσας σώματος. Αυτή η καινοτομία θα επιτρέψει πιο φυσικές συνομιλίες [122], δημιουργώντας ψηφιακούς θεραπευτές που υποστηρίζουν ασθενείς με ψυχικές διαταραχές, προσαρμόζοντας τη συνομιλία τους βάσει της συναισθηματικής κατάστασης του θεραπευόμενου. Μάλιστα, προτείνεται η χρήση ρομποτικών συστημάτων υποστήριξης αναμνήσεων σε ασθενείς με άνοια. Πρόκειται για μία θεραπεία αναπόλησης, η οποία βοηθά άτομα με άνοια να ανακαλούν παρελθοντικά γεγονότα, όπως αγαπημένα τραγούδια, τηλεοπτικές εκπομπές, ταινίες και ιστορικά γεγονότα. Συγκεκριμένα, αξιοποιώντας τις δυνατότητες των LLMs για να ανακτούν πληροφορίες, να συνοψίζουν δεδομένα και να προσαρμόζουν τις αλληλεπιδράσεις τους με βάση τις συναισθηματικές αντιδράσεις του χρήστη, τα κοινωνικά βοηθητικά ρομπότ θα παρέχουν εξατομικευμένες αφηγήσεις και σχετικό οπτικοακουστικό υλικό, συμβάλλοντας στην συναισθηματική ευημερία του ατόμου [92].

Επιπροσθέτως, το τοπίο των χειρουργικών επεμβάσεων πρόκειται να αλλάξει ριζικά, καθώς τα επόμενα χρόνια, οι χειρουργικές ρομποτικές πλατφόρμες δεν θα περιορίζονται στην υποβοήθηση των γιατρών με σταθερές λειτουργίες. Αντιθέτως, τα LLMs θα επιτρέψουν στα ρομποτικά συστήματα να προσαρμόζουν τη λειτουργία τους δυναμικά, βασισμένα σε δεδομένα που συλλέγονται σε πραγματικό χρόνο από τον ασθενή [52]. Αυτό περιλαμβάνει ρομπότ που



παρέχουν αναλύσεις και προτάσεις σε χειρουργούς κατά τη διάρκεια επεμβάσεων, προσαρμόζοντας τη βοήθειά τους στις ανάγκες της επέμβασης. Φυσικά, η πρόοδος της τεχνητής νοημοσύνης θα μπορούσε να οδηγήσει στην ανάπτυξη πιο αυτόνομων ρομποτικών χειρουργικών συστημάτων, επιτρέποντας στους χειρουργούς να εκτελούν πολύπλοκες επεμβάσεις εξ αποστάσεως ή με ελάχιστη βοήθεια [132]. Σημαντική είναι η συμβολή των μεγάλων γλωσσικών μοντέλων και στην εκπαίδευση των ανερχόμενων χειρουργών, καθώς οι προσομοιώσεις με τεχνητή νοημοσύνη μπορούν να τους παρέχουν ρεαλιστικά εκπαιδευτικά περιβάλλοντα για να εξασκούνται σε σύνθετες διαδικασίες και να βελτιώνουν τις δεξιότητές τους, οδηγώντας σε βελτιωμένα χειρουργικά αποτελέσματα [80]. Όσον αφορά τη διαχείριση επειγόντων περιστατικών, τα μελλοντικά ρομπότ θα αναλάβουν ενεργό ρόλο στην αντιμετώπισή τους, αξιολογώντας ζωτικά σημεία και παρέχοντας ιατρικές οδηγίες πριν την άφιξη του προσωπικού έκτακτης ανάγκης [50]. Αυτό αφορά την ύπαρξη ρομποτικών συστημάτων σε νοσοκομεία ή ασθενοφόρα που θα αξιολογούν τη σοβαρότητα ενός περιστατικού και θα προετοιμάζουν την ιατρική ομάδα για την κατάλληλη παρέμβαση. Στο ίδιο πλαίσιο, τα συνομιλητικά LLMs μπορούν να ενσωματωθούν στην αναγνώριση ομιλίας για την αυτόνομη διεξαγωγή διαβουλεύσεων με ασθενείς, προτείνοντας διαγνώσεις και σχέδια διαχείρισης υγείας με βάση τις τελευταίες κατευθυντήριες οδηγίες και σύγχρονες έρευνες [52].

Οι παραπάνω εφαρμογές υποδεικνύουν ότι η συνεχής πρόοδος των LLMs και η ενσωμάτωσή τους σε ρομποτικά συστήματα υγειονομικής περίθαλψης αναμένεται να μετασχηματίσουν τον κλάδο, βελτιώνοντας τη διάγνωση, τη θεραπεία, τη νοσηλεία και την υποστήριξη των ασθενών [51]. Από την αυξημένη αυτονομία και τη βελτίωση της αλληλεπίδρασης ανθρώπου-μηχανής έως τις νέες δυνατότητες στη χειρουργική και την απομακρυσμένη περίθαλψη, οι εφαρμογές της ρομποτικής τεχνητής νοημοσύνης πολλαπλασιάζονται. Τα ρομποτικά συστήματα υγειονομικής περίθαλψης του μέλλοντος θα είναι πιο προσαρμοστικά και συναισθηματικά ευφυή. Με την πρόοδο των LLMs, η ρομποτική υγεία δεν θα περιορίζεται μόνο στην εκτέλεση μηχανικών διαδικασιών, αλλά θα ενσωματώνει λειτουργίες που ενισχύουν την αλληλεπίδραση, την εξατομικευμένη φροντίδα και την αντιμετώπιση πολύπλοκων ιατρικών αναγκών [122].

## ΣΥΜΠΕΡΑΣΜΑΤΑ

Συνοψίζοντας, η εφαρμογή των Μεγάλων Γλωσσικών Μοντέλων σε αυτόνομα ρομποτικά ιατρικά συστήματα αποτελεί μια επαναστατική προσέγγιση με πολυδιάστατα οφέλη. Δεν υπάρχει καμία αμφιβολία πως τα LLMs ενισχύουν την ευφυΐα και την αυτονομία των ρομπότ, προσφέροντας τη δυνατότητα κατανόησης και ανταπόκρισης σε φυσική γλώσσα. Από αυτή την καινοτομία έχει επωφεληθεί σημαντικά η επικοινωνία με ασθενείς και ιατρικό προσωπικό, ενώ τομείς όπως η προ-συμβουλευτική, η διάγνωση, η θεραπεία, η ψυχολογική υποστήριξη, η χειρουργική ρομποτική και οι διοικητικές νοσοκομειακές λειτουργίες πρόκειται να γνωρίσουν ουσιώδεις μετασχηματιστικές αλλαγές [50]. Η ανάπτυξη αυτών των ευφών συστημάτων σηματοδοτεί τη μετάβαση σε μια νέα εποχή τεχνολογικής καινοτομίας, που χαρακτηρίζεται από τη βελτίωση της λήψης αποφάσεων, την ταχύτερη διεκπεραίωση διαγνωστικών διαδικασιών και την αναβάθμιση της ποιότητας της παρεχόμενης φροντίδας στους ασθενείς [54].

Παρά τα πολλαπλά οφέλη, αυτή η τεχνολογική μετάβαση συνοδεύεται από εγγενείς κινδύνους, και σύνθετα ζητήματα ηθικής και ασφάλειας που απαιτούν συστηματική αξιολόγηση. Η ακρίβεια και η αξιοπιστία των μοντέλων, οι απαιτήσεις υπολογιστικής ισχύος, η πολύπλοκη διαδικασία συγκέντρωσης δεδομένων και η προσαρμογή της προσομοιωμένης εκπαίδευσης σε πραγματικές συνθήκες αποτελούν βασικές προκλήσεις για την αποτελεσματική ενσωμάτωση αυτών των τεχνολογιών [133]. Ταυτόχρονα, οι ηθικές και νομικές ανησυχίες που σχετίζονται με την ασφάλεια των δεδομένων, την αποφυγή μεροληψίας και την αυτόνομη λήψη αποφάσεων παραμένουν κρίσιμα ζητήματα που δεν μπορούν να αγνοηθούν. Η εισαγωγή των LLMs στον ευαίσθητο τομέα της υγειονομικής περίθαλψης καθιστά επιτακτική την ανάγκη προσεκτικής εξέτασης της δεοντολογίας, της ιδιωτικότητας και της ασφάλειας. Βεβαίως, οι ταχέως εξελισσόμενες τεχνολογίες της τεχνητής νοημοσύνης οφείλουν να συνοδεύεται από σαφή και ευέλικτα ηθικά, νομικά και ρυθμιστικά πλαίσια [52], ενώ παράλληλα κρίνεται αναγκαίο να ενισχυθεί η ενσωμάτωση των LLMs στην κλινική πρακτική και η συνεργασία ανθρώπου-μηχανής.

Στα πλαίσια αυτής της τεχνολογικής επανάστασης, καθίσταται προφανής η αναγκαιότητα της συνεργασίας μεταξύ μηχανικών, επιστημόνων δεδομένων, ιατρών, ηθικολόγων και φορέων χάραξης πολιτικής, με στόχο την υπεύθυνη ενσωμάτωση των LLMs στη ρομποτική προς όφελος της κοινωνικής ευημερίας [111]. Η επιτυχής και αμερόληπτη ενσωμάτωση αυτών των συστημάτων στην υγειονομική περίθαλψη απαιτεί μια διεπιστημονική προσέγγιση, η οποία περιλαμβάνει τη συνεχή εκπαίδευση τόσο των ιατρών όσο και των παρόχων υγειονομικής φροντίδας για τη σωστή και υπεύθυνη χρήση των LLMs. Η ενημέρωση των επαγγελματιών υγείας είναι κρίσιμη για τη γεφύρωση του χάσματος μεταξύ τεχνολογίας και κλινικής πρακτικής, καθοδηγώντας τις μελλοντικές μεθοδολογίες και την επιμέλεια των συνόλων δεδομένων [134]. Παράλληλα, οι επαγγελματίες της υγειονομικής περίθαλψης οφείλουν να δίνουν προτεραιότητα στην ευημερία και τις συναισθηματικές ανάγκες των ασθενών, διασφαλίζοντας ότι η φροντίδα τους δεν καθοδηγείται αποκλειστικά από συστάσεις που δημιουργούνται από την τεχνητή νοημοσύνη. Η ενεργή ανθρώπινη συμμετοχή είναι απαραίτητη ώστε τα LLMs να λειτουργούν ως συμπληρωματικά εργαλεία, χωρίς να υπονομεύεται η συμπονετική φροντίδα και η συνολική εμπειρία των ασθενών [115].

Μελλοντικά, η περαιτέρω βελτίωση των LLMs αναμένεται να μειώσει τα σφάλματα, να αυξήσει τη διαφάνεια και να βελτιστοποιήσει την ενσωμάτωσή τους σε ρομποτικά συστήματα. Ωστόσο, παρά την αδιαμφισβήτητη πρόοδο, εξακολουθεί να υπάρχει ένας διάχυτος φόβος στην κοινωνία σχετικά με την αυξανόμενη αυτονομία της τεχνητής νοημοσύνης και την πιθανή αντικατάσταση των ανθρώπινων επαγγελματιών. Αξίζει να επισημανθεί ότι τα LLMs δεν προορίζονται να αντικαταστήσουν τους ιατρούς ή άλλους παρόχους υγειονομικής φροντίδας, αλλά να αξιοποιηθούν ως συμπληρωματικά εργαλεία. Ο ρόλος τους περιλαμβάνει την ανάληψη επαναλαμβανόμενων, χρονοβόρων ή σύνθετων εργασιών, επιτρέποντας στους επαγγελματίες υγείας να εστιάζουν σε δημιουργικές, στρατηγικές και συναισθηματικά απαιτητικές διαδικασίες. Με την ορθή αξιοποίηση της τεχνητής νοημοσύνης και την καθοριστική συμβολή της ανθρώπινης εμπειρογνομοσύνης, μπορούν να επιτευχθούν καλύτερα κλινικά αποτελέσματα και υψηλότερη ποιότητα της περίθαλψης των ασθενών σε παγκόσμιο επίπεδο [134]. Η διατήρηση της σχέσης ιατρού-ασθενούς, η ενίσχυση της ενσυναίσθησης και η υπεύθυνη ενσωμάτωση των LLMs στην κλινική πρακτική μπορούν να διασφαλίσουν ότι η τεχνολογία λειτουργεί προς όφελος της ανθρώπινης φροντίδας και όχι εις βάρος της.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] W. X. Zhao et al., "A Survey of Large Language Models," arXiv preprint, Online, 2024.
- [2] C. Zhang, J. Chen, J. Li, Y. Peng and Z. Mao, "Large language models for human–robot interaction: A review," *Biomimetic Intelligence and Robotics*, vol. 3, no. 4, pp. 2667-3797, 2023.
- [3] F. Zenga, W. Gana, Y. Wang, N. Liu and P. S. Yu, "Large Language Models for Robotics: A Survey," arXiv preprint, Online, 2023.
- [4] R. Siegwart, I. Nourbakhsh and D. Scaramuzza, *Introduction to Autonomous Robots*, 2nd ed. ed., Cambridge, MA: MIT Press, 2011.
- [5] C. Stryker and E. Kavlakoglu, "What is Artificial Intelligence?," IBM, 16 August 2024. [Online]. Available: <https://www.ibm.com/topics/artificial-intelligence>. [Accessed 20 October 2024].
- [6] Y. Shen, L. Heacock, J. Elias, K. D. Hentel, B. Reig, G. Shih and L. Moy, "ChatGPT and Other Large Language Models Are Double-edged Swords," *Radiology*, vol. 307, no. 2, e230163, 2023.
- [7] A. Vaswani et al., "Attention Is All You Need," in *31st Conference on Neural Information Processing Systems*, Long Beach, CA, USA, 2017.
- [8] *What are Large Language Models (LLMs)?*. [Film]. Unated States of America: Google for Developers, 2023.
- [9] N. Saini, "RESEARCH PAPER ON ARTIFICIAL INTELLIGENCE & IT'S APPLICATIONS," *IJRTI*, vol. 8, no. 6, pp. 2456-3315, 2023.
- [10] IBM, "Neural network," IBM, 2024. [Online]. Available: <https://www.ibm.com/topics/neural-networks>. [Accessed 02 11 2024].
- [11] CLOUDFLARE, "What is a neural network?," CLOUDFLARE, 2024. [Online]. Available: <https://www.cloudflare.com/learning/ai/what-is-neural-network/>. [Accessed 05 11 2024].
- [12] X. Τσιφτσή, "Gesture Recognition using Artificial Intelligence and Application to an Unmanned Ground Vehicle (UGV)," Πολυτεχνείο Κρήτης, Χανιά, 2024.
- [13] MathWorks, "Deep Learning," MathWorks, 2024. [Online]. Available: <https://www.mathworks.com/discovery/deep-learning.html>. [Accessed 05 11 2024].

- [14] Google Cloud, "What is Deep Learning?," Google Cloud, 2024. [Online]. Available: <https://cloud.google.com/discover/what-is-deep-learning>. [Accessed 05 11 2024].
- [15] B. Min et al., "Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey," arXiv preprint, Online, 2021.
- [16] Wang, X. Gao et al., "Understanding LLMs: A Comprehensive Overview from Training to Inference," arXiv preprint, Online, 2024.
- [17] M. Hoque, "A Comprehensive Overview of Transformer-Based Models: Encoders, Decoders, and More," Medium, 2023. [Online]. Available: <https://medium.com/@minh.hoque/a-comprehensive-overview-of-transformer-based-models-encoders-decoders-and-more-e9bc0644a4e5>. [Accessed 21 11 2024].
- [18] C. Munoz, "Overview of Large Language Models: From Transformer Architecture to Prompt Engineering," Holistic AI, 24 02 2023. [Online]. Available: <https://www.holisticai.com/blog/from-transformer-architecture-to-prompt-engineering>. [Accessed 24 01 2025].
- [19] M. A. K. RAIAAN et al, "A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges," *IEEE Access*, vol. 12, pp. 26839–26874, 2024.
- [20] C. D. Manning, P. Raghavan and H. Schütze, *Introduction to Information Retrieval*, Stanford, CA, USA: Cambridge University Press, 2008.
- [21] K. Clark, U. Khandelwal, O. Levy and C. D. Manning, "What does BERT look at?," Association for Computational Linguistics, Florence, Italy, 2019.
- [22] J. Bandy and N. Vincent, "Addressing "Documentation Debt" in Machine," in *Neural Information Processing Systems*, Sydney, Australia, 2021.
- [23] T. B. Brown et al., "Language Models are Few-Shot Learners," arXiv preprint, Online, 2020.
- [24] H. Touvron et al., "LLaMA: Open and Efficient Foundation Language Models," arXiv preprint, Online, 2023.
- [25] Wikipedia Contributors, "Wikipedia," Wikipedia Foundation, 2024. [Online]. Available: <https://el.wikipedia.org/wiki/Reddit>. [Accessed 12 11 2024].
- [26] A. Roberts, C. Raffel and N. Shazeer, "How Much Knowledge Can You Pack," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, 2020.

- [27] A. I. Károly, P. Galambos, J. Kuti and I. J. Rudas, "Deep Learning in Robotics: Survey on Model Structures and Training Strategies," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 266 - 279, 2020.
- [28] H. W. Chung et al., "Scaling Instruction-Finetuned Language Models," arXiv preprints, Online, 2022.
- [29] G. Penedo et al., "The RefinedWeb Dataset for Falcon LLM: Outperforming Curated Corpora with Web Data, and Web Data Only," arXiv preprint, Online, 2023.
- [30] J. J. Webster and C. Kit, "TOKENIZATION AS THE INITIAL PHASE IN NLP," in *Proceedings of the International Conference on Computational Linguistics (COLING)*, Nantes, France, 1992.
- [31] H. Naveed et al., "A Comprehensive Overview of Large Language Models," arXiv preprint, Online, 2024.
- [32] J. Huang and K. C.-C. Chang, "Towards Reasoning in Large Language Models: A Survey," *ACL Anthology*, pp. 1049–1065, 2023.
- [33] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei and I. Sutskever, "Language Models are Unsupervised Multitask Learners," OpenAI, San Francisco, California, United States, 2019.
- [34] A. Radford, K. Narasimhan, T. Salimans and I. Sutskever, "Improving Language Understanding," OpenAI, San Francisco, CA, 2018.
- [35] Run Team, "LLM Training," Run:ai, 2024. [Online]. Available: <https://www.run.ai/guides/machine-learning-engineering/llm-training>. [Accessed 18 11 2024].
- [36] IBM Team, "What is backpropagation?," IBM, 2024. [Online]. Available: <https://www.ibm.com/think/topics/backpropagation>. [Accessed 18 11 2024].
- [37] M. Shoenybi, M. Patwary, R. Puri, P. LeGresley, J. Casper and B. Catanzaro, "Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism," arXiv preprint, Online, 2020.
- [38] R. S. H, "Large Language Models in Deep Learning," Intuitive Tutorials, 18 06 2023. [Online]. Available: <https://intuitivetutorial.com/2023/06/18/large-language-models-in-deep-learning/>. [Accessed 24 01 2025].
- [39] P. F. Christiano, J. Leike, T. Brown, M. Martic, S. Legg and D. Amodei, "Deep Reinforcement Learning from Human Preferences," in *31st Conference on Neural Information Processing Systems*, Long Beach, CA, USA., 2017.

- [40] Mastering LLM (Large Language Model), "LLM Training: A Simple 3-Step Guide You Won't Find Anywhere Else!," Medium, 2023. [Online]. Available: <https://masteringllm.medium.com/llm-training-a-simple-3-step-guide-you-wont-find-anywhere-else-98ee218809e5>. [Accessed 20 11 2024].
- [41] J. Wang et al., "Large language models for robotics: Opportunities, challenges, and perspectives," *Journal of Automation and Intelligence*, pp. 2949-8554, 2024.
- [42] D. Shah, B. Osiński, b. ichter and S. Levine, "LM-Nav: Robotic Navigation with Large Pre-Trained Models of Language, Vision, and Action," in *Proceedings of The 6th Conference on Robot Learning*, Auckland, New Zealand, 2023.
- [43] T. Baltrusaitis, C. Ahuja and L.-P. Morency, "Multimodal Machine Learning: A Survey and Taxonomy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 423–443, 2017.
- [44] A. Maghami and M. Khoshdarregi, "Vision-based target localization and online error correction for high-precision robotic drilling," *Robotica*, vol. 42, no. 24, p. 691–698, 2024.
- [45] J. Wang et al., "A Survey on Large Language Models from General Purpose to Medical Applications: Datasets, Methodologies, and Evaluations," arviX preprint, Online, 2024.
- [46] A. K. Pandey and R. Gelin, "A Mass-Produced Sociable Humanoid Robot: Pepper: The First Machine of Its Kind," *IEEE ROBOTICS & AUTOMATION MAGAZINE*, vol. 25, no. 3, pp. 40 - 48, 2018.
- [47] Diligent Robotics, "Moxi," 2023. [Online]. Available: <https://www.diligentrobots.com/moxi>. [Accessed 05 01 2025].
- [48] K. Denecke, S. Vaaheesan and A. Arulnathan, "A Mental Health Chatbot for Regulating Emotions (SERMO) - Concept and Usability Test," *IEEE Transactions on Emerging Topics in Computing*, vol. 9, no. 3, pp. 1170 - 1182, 2020.
- [49] K. K. Fitzpatrick, A. Darcy and M. Vierhile, "Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial," *JMIR Mental Health*, vol. 4, no. 2, 2017.
- [50] R. Yang, T. F. Tan, W. Lu, A. J. Thirunavukarasu, D. S. W. Ting and N. Liu, "Large language models in health care: Development, applications, and challenges," *Health Care Science*, vol. 2, no. 4, 2023.
- [51] Intel Corporation, "Robotics in Healthcare: The Future of Robots in Medicine," Intel, 2024. [Online]. Available: <https://www.intel.com/content/www/us/en/healthcare-it/robotics-in-healthcare.html>. [Accessed 30 11 2024].

- [52] A. Guni, P. Varma, J. Zhang, M. Fehervari and H. Ashrafian, "Artificial Intelligence in Surgery: The Future Is Now," *Eur Surg Res*, vol. 65, no. 1, pp. 22–39, 2024.
- [53] J. Clusmann et al., "The future landscape of large language models in medicine," *Nature Machine Intelligence*, vol. 2, no. 1, 2023.
- [54] Z. A. Nazi and A. Peng, "Large Language Models in Healthcare and Medical Domain: A Review," arXiv preprints, Online, 2024.
- [55] A. Moglia et al., "Large language models in healthcare: from a systematic review on medical examinations to a comparative analysis on fundamentals of robotic surgery online test," *Artificial Intelligence Review*, vol. 57, no. 6, pp. 4567–4590, 2024.
- [56] Association for Advancing Automation, "Medical Robots," Automate.org, 2024. [Online]. Available: <https://www.automate.org/robotics/service-robots/service-robots-medical>. [Accessed 4 12 2024].
- [57] K. Singhal et al., "Towards Expert-Level Medical Question Answering with Large Language Models," arXiv preprint, Online, 2023.
- [58] John Snow Labs, "The Impact of Medical LLMs on Disease Diagnosis and Treatment," John Snow Labs, 06 07 2023. [Online]. Available: <https://www.johnsnowlabs.com/the-impact-of-medical-llms-on-disease-diagnosis-and-treatment/>. [Accessed 23 12 2024].
- [59] A. Toma, P. R. Lawler, J. Ba, R. G. Krishnan, B. B. Rubin and B. Wang, "Clinical Camel: An Open Expert-Level Medical Language Model with Dialogue-Based Knowledge Encoding," arXiv preprint, Online, 2023.
- [60] T. Han, L. C. Adams, J.-M. Papaioannou, P. Grundmann, T. Oberhauser, A. Löser, D. Truhn and K. K. Bressen, "MedAlpaca -- An Open-Source Collection of Medical Conversational AI Models and Training Data," arXiv preprint, Online, 2023.
- [61] C. Wu, W. Lin, X. Zhang, Y. Zhang, Y. Wang and W. Xie, "PMC-LLaMA: Towards Building Open-source Language Models for Medicine," arXiv preprint, Online, 2023.
- [62] J. Lee et al, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," arXiv preprint, Online, 2019.
- [63] X. Yang et al., "GatorTron: A Large Clinical Language Model to Unlock Patient Information from Unstructured Electronic Health Records," arXiv preprint, Online, 2022.
- [64] Google Research; DeepMind, "Towards Expert-Level," arXiv, Online, 2023.
- [65] S. Zargarzadeh, M. Mirzaei, Y. Ou and M. Tavakoli, "From Decision to Action in Surgical Autonomy: Multi-Modal Large Language Models for Robot-Assisted Blood Suction," *arXiv*, vol. 10, no. 2, pp. 1234–1245, 2024.



- [66] AdventHealth University, "Robotics in Healthcare: Past, Present, and Future," AdventHealth University, 10 02 2022. [Online]. Available: <https://www.ahu.edu/blog/robotics-in-healthcare>. [Accessed 05 12 2024].
- [67] Wikipedia contributors, "da Vinci Surgical System," Wikimedia Foundation, [Online]. Available: [https://en.wikipedia.org/wiki/Da\\_Vinci\\_Surgical\\_System](https://en.wikipedia.org/wiki/Da_Vinci_Surgical_System). [Accessed 05 12 2024].
- [68] North Texas Surgical Specialists, "Benefits of the da Vinci Surgical System," NORTH TEXAS Surgical Specialists, [Online]. Available: <https://ntxsurgical.com/benefits-of-the-davinci-surgical-system/>. [Accessed 17 01 2025].
- [69] Ekol Hospitals, "Da Vinci Surgical System," EKOL INTERNATIONAL HOSPITALS, [Online]. Available: <https://www.ekolhospitals.com/en/da-vinci-surgical-system>. [Accessed 17 01 2025].
- [70] K. Erozkhan and E. Gorgun, "Robotic colorectal surgery and future directions," *ScienceDirect*, vol. 230, pp. 91-98, 2024.
- [71] M. Gómez Ruiz, M. Lainez Escribano, C. Cagigas Fernández, L. Cristobal Poch and S. Santarrufina Martínez, "Robotic surgery for colorectal cancer," *AGSurg*, vol. 4, no. 6, pp. 646-651, 2020.
- [72] C. M. Oliveira, H. T. Nguyen, A. R. Ferraz, K. Watters, B. Rosman and R. Rahbar, "Robotic Surgery in Otolaryngology and Head and Neck Surgery: A Review," *Minimally Invasive Surgery*, vol. 16, no. 3, pp. 336-341, 2012.
- [73] M. Marin-Cuartas, M. P. Sá, G. Torregrossa and P. M. Davierwala, "Minimally invasive coronary artery surgery: Robotic and nonrobotic minimally invasive direct coronary artery bypass techniques," *JTCVS TECHNIQUES*, vol. 10, pp. 170-177, 2021.
- [74] T. Bellos, I. Manolitsis, S. Katsimperis, P. Juliebø-Jones, G. Feretzakis, I. Mitsogiannis, I. Varkarakis, B. K. Somani and L. Tzelves, "Artificial Intelligence in Urologic Robotic Oncologic Surgery: A Narrative Review," *cancers*, vol. 16, no. 9, 1775, 2024.
- [75] S. Williams, H. L. Horsfall, J. P. Funnell, J. G. Hanrahan, D. Z. Khan, W. Muirhead, D. Stoyanov and H. J. Marcus, "Artificial Intelligence in Brain Tumour Surgery—An Emerging Paradigm," *cancers*, vol. 13, no. 19, 5010, 2021.
- [76] J. Xu et al., "Applications of Artificial Intelligence Based on Medical Imaging in Glioma: Current State and Future Challenges," *Frontiers in Oncology*, vol. 12, p. 14, 27 07 2022.
- [77] M. Roche, "The MAKO robotic-arm knee arthroplasty system," *Archives of Orthopaedic and Trauma Surgery*, vol. 141, no. 12, pp. 2043–2047, 2021.

- [78] J. S. E. S.-M. E. S. S. P. S. L. C. Batailler, "Artificial intelligence in knee arthroplasty: current concept of the available clinical applications," *Springer Nature*, vol. 4, no. 17, 2022.
- [79] A. B. Lisacek-Kiosoglous, A. S. Powling, A. Fontalis, A. Gabr, E. Mazomenos and F. S. Haddad, "Artificial intelligence in orthopaedic surgery," *Bone & Joint Research*, vol. 12, no. 7, pp. 447 - 454, 2023.
- [80] I. Muhammad, M. Saqib, M. Zareen and H. Mumtaz, "Artificial intelligence: revolutionizing robotic surgery: review," *Annals of Medicine and Surgery*, vol. 86, no. 9, pp. 5401-5409, 2024.
- [81] HYGEIA Hospital, "MAKO Robotic-Arm Assisted System," hygeia hospital, 2024. [Online]. Available: <https://www.hygeia.gr/en/services/department/mako-robotic-arm-assisted-system/>. [Accessed 17 01 2025].
- [82] S. Kaliappan, A. S. Anand, K. Saha and R. Karkar, "Exploring the Role of LLMs for Supporting Older Adults:," in *HCI and Aging: New Directions, New Principles Workshop at ACM SIGCHI 2024*, 2024.
- [83] K. Kim et al., "Framework for Integrating Large Language Models with a Ro-botic Health Attendant for Adaptive Task Execution in Patient Care," Preprints.org, 2024.
- [84] R. Ranisch and J. Haltaufderheide, "Rapid Integration of LLMs in Healthcare Raises Ethical Concerns: An Investigation into Deceptive Patterns in Social Robots," arXiv, Online, 2024.
- [85] TheGuardian team, "Robots to be used in UK care homes to help reduce loneliness," *The Guardian*, 08 09 2020.
- [86] R. Baart, "This Robot Takes Care of Your Children," *Next Nature Network*, 26 10 20216.
- [87] M. Bilal, Y. Jamil, D. Rana and H. H. Shah, "Fine-tuning a LLM using Reinforcement Learning from Human Feedback for a Therapy Chatbot Application," *Annals of Biomedical Engineering*, Stockholm, Sweden, 2023.
- [88] A. N. Vaidyam, H. Wisniewski, A. N. Vaidyam, H. Wisniewski and J. B. Torous, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape," *Sage Journals*, vol. 64, no. 7, pp. 456–464, 2019.
- [89] R. Choi, T. Kim, S. Park, J. G. Kim and S.-J. Lee, "Private Yet Social: How LLM Chatbots Support and Challenge," in *Conference '17*, Washington, DC, USA, 2017.
- [90] C. Deng et al., "ASD-Chat: An Innovative Dialogue Intervention System for Children with Autism based on LLM and VB-MAPP," arXiv, Online, 2024.

- [91] Y. a. Tang, "EmoEden: Applying Generative Artificial Intelligence to Emotional Learning for Children with HighFunction Autism," in *CHI Conference on Human Factors in Computing Systems*, Chengdu, China, 2024.
- [92] S. Pashangpour and G. Nejat, "The Future of Intelligent Healthcare: A Systematic Analysis and Discussion on the Integration and Impact of Robots Using Large Language Models for Healthcare," *Robotics*, vol. 13, no. 8, p. 112, 2024.
- [93] Healthcare in Europe, "Large language models in healthcare: shorter prompts, less emissions?," Healthcare in Europe, 19 11 2024. [Online]. Available: <https://healthcare-in-europe.com/en/news/genai-llm-healthcare-prompt-sustainability.html>. [Accessed 26 01 2025].
- [94] J. L. Y. L. C. M. a. W. W. X. Zhu, "A survey on model compression for large language models," arXiv, Online, 2023.
- [95] R. AlSaad, A. Abd-alrazaq, S. Boughorbel, A. Ahmed, M.-A. Renault, R. Damseh and J. Sheikh, "Multimodal Large Language Models in Health Care: Applications, Challenges, and Future Outlook," *Journal of Medical Internet Research*, vol. 26, 2024.
- [96] C. Hooper et al., "KVQuant: Towards 10 Million Context Length LLM Inference with KV Cache Quantization," arXiv, Online, 2024.
- [97] G. Xiao, Y. Tian, B. Chen, S. Han and M. Lewis, "Efficient Streaming Language Models with Attention Sinks," arviX, 2024.
- [98] N. Ratner, Y. Levine, Y. Belinkov, I. M. O. Ram and O. Abend, "Parallel context windows for large language models," arXiv, 2022.
- [99] H. Hwai, Y.-J. Ho, C.-H. Wang and C.-H. Huang, "Large language model application in emergency medicine and critical care," *Journal of the Formosan Medical Association*, pp. 0929-6646, 28 08 2024.
- [100] O. Freyera, I. C. Wiest, J. N. Kather and S. Gilbert, "A future role for health applications of large language models depends on regulators enforcing safety standards," *The Lancet Digital Health*, vol. 6, no. 9, pp. e662-e672, 2024.
- [101] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 11, pp. 4793 - 4813, 2020.
- [102] N. F. Liu et al., "Lost in the Middle: How Language Models Use Long Contexts," arXiv, 2023.

- [103] H. Zhou et al., "A Survey of Large Language Models in Medicine: Progress, Application, and Challenge," arviX, 2024.
- [104] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang, W. Ye, Y. Zhang, Y. Chang, P. S. Yu, Q. Yang and X. Xie, "A Survey on Evaluation of Large Language Models," *ACM Transactions on Intelligent Systems and Technology*, vol. 15, no. 3, pp. 1 - 45, 2024.
- [105] D. Zhang, X. Xue, P. Gao, Z. Jin, M. Hu2, Y. Wu and X. Ying, "A survey of datasets in medicine for large language models," *Intell Robot*, vol. 4, no. 4, pp. 457-78, 2024.
- [106] C. Basta, M. R. Costa-jussa and N. Casas, "Evaluating the Underlying Gender Bias in Contextualized Word Embeddings," arXiv, 2019.
- [107] B. Hutchinson et al., "Social Biases in NLP Models as Barriers for Persons with Disabilities," arXiv, 2020.
- [108] V. Volovici, N. L. Syn, A. Ercole, J. J. Zhao and N. Liu, "Steps to avoid overuse and misuse of machine learning in clinical research," *Nature Medicine*, vol. 28, pp. 1996–1999, 2022.
- [109] V. Daitch et al., "Underrepresentation of women in randomized controlled trials: a systematic review and meta-analysis," *Trials*, vol. 23, 1038, 2022.
- [110] K. Kurita, N. Vyas, A. Pareek, A. W. Black and Y. Tsvetkov, "Measuring Bias in Contextualized Word Representations," arviX, 2019.
- [111] A. Efa, "Evaluating Security Risks in Robotics Powered by Large Language Models," Technische Universität München, Munich, Germany, 2024.
- [112] N. Norori, Q. Hu, F. M. Aellen, F. D. Faraci and A. Tzovara, "Addressing bias in big data and AI for health care: A call for open science," *Patterns*, vol. 2, no. 10, 100347, 2021.
- [113] J. D, "Empathy, sympathy and compassion in healthcare: Is there a problem? Is there a difference? Does it matter?," *Journal of the Royal Sociaty of Medicine*, vol. 109, no. 12, pp. 446-452, 2016.
- [114] C. S., "Robots and the Possibility of Humanistic Care," *Springer Nature*, vol. 14, no. 10, pp. 2095–2108, 2022.
- [115] C. Wang et al., "Ethical Considerations of Using ChatGPT in Health Care," *Journal of Medical Internet Research*, vol. 25, e48009, 2023.
- [116] L. Na, C. Yan and C.-C. Lo, "Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning," *JAMA Network Open*, vol. 1, no. 8, e199670, 2018.

- [117] N. Carlini et al., "Extracting Training Data from Large Language Models," arXi, 2021.
- [118] L. Du, C. Xia, Z. Deng, G. Lu, S. Xia and J. Ma, "A machine learning based approach to identify protected health information in Chinese clinical text," *International Journal of Medical Informatics*, vol. 116, pp. 24-32, 2018.
- [119] W. Xuewei and Z. Bin, "ChatGPT promotes healthcare: current applications and potential challenges," *INTERNATIONAL JOURNAL OF SURGERY*, vol. 110, no. 1, pp. 606-608, 2024.
- [120] Y. Chen and P. Esmailzadeh, "Generative AI in Medical Practice: In-Depth Exploration of Privacy and Security Challenges," *Journal of Medical Internet Research*, vol. 26, p. e53008, 2024.
- [121] B. A. et al, "RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control," arviX, 2023.
- [122] S. Kannan, "What's Next for Generative AI and LLMs in the Healthcare Industry?," Opteamix, 30 12 2024. [Online]. Available: <https://opteamix.com/whats-next-for-generative-ai-and-llms-in-the-healthcare-industry/>. [Accessed 29 01 2025].
- [123] A. M. et al., "Do as i can, not as i say: Grounding language in robotic affordances," arXiv, 2022.
- [124] J. Jiang, X. Liu and C. Fan, "Low-Parameter Federated Learning with Large Language Models," arXiv, 2023.
- [125] F. Cai, D. Yuan, Z. Yang and L. Cui, "Edge-LLM: A Collaborative Framework for Large Language Model Serving in Edge Computing," in *2024 IEEE International Conference on Web Services (ICWS)*, Shenzhen, China, 2024.
- [126] A. Chavan, R. Magazine, S. Kushwaha, M. Debbah and D. Gupta, "Faster and Lighter LLMs: A Survey on Current Challenges and Way Forward," arXiv, 2024.
- [127] Z. Liu, Y. Zhang, P. Li, Y. Liu and D. Yang, "A Dynamic LLM-Powered Agent Network for Task-Oriented Agent Collaboration," arXiv, 2024.
- [128] Y. Yao, P. Wang, B. Tian, S. Cheng, Z. Li, S. Deng, H. Chen and N. Zhang, "Editing Large Language Models: Problems, Methods, and Opportunities," arXiv, 2023.
- [129] K. Chang, S. Xu, C. Wang, Y. Luo, T. Xiao and J. Zhu, "Efficient Prompting Methods for Large Language Models: A Survey," arXiv, 2024.
- [130] Z. Zhang, J. Zhao, Q. Zhang, T. Gui and X. Huang, "Unveiling Linguistic Regions in Large Language Models," arXiv, 2024.

- [131] Y. Li, B. Hu, W. Wang, X. Cao and M. Zhang, "Towards Vision Enhancing LLMs: Empowering Multimodal Knowledge Storage and Sharing in LLMs," *arXiv*, 2023.
- [132] P. Fiorini, K. Y. Goldberg, Y. Liu and R. H. Taylor, "Concepts and Trends in Autonomy for Robot-Assisted Surgery," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 993 - 1011, 2022.
- [133] M. Cascella, J. Montomoli, V. Bellini and E. Bignami, "Evaluating the Feasibility of ChatGPT in Healthcare: An Analysis of Multiple Clinical and Research Scenarios," *Journal of Medical Systems*, vol. 47, no. 33, 2023.
- [134] B. D. Simon, K. B. Ozyoruk, D. G. Gelikman, S. A. Harmon and B. Türkbey, "The future of multimodal artificial intelligence models for integrating imaging and clinical metadata: a narrative review," *Diagnostic and Interventional Radiology*, 2 10 2024.
- [135] SoftBank team, "Robot," SoftBank, [Online]. Available: <https://www.softbank.jp/en/robot/>. [Accessed 17 01 2025].
- [136] Κ. Γεωργούλη, "Τεχνητή Νοημοσύνη: Μία Εισαγωγική Προσέγγιση," Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών (ΣΕΑΒ), Αθήνα, 2015.
- [137] J.-A. Claret, G. Venture and L. Basañez, "Exploiting the Robot Kinematic Redundancy for Emotion Conveyance to Humans as a Lower Priority Task," *International Journal of Social Robotics* , vol. 9, no. 2, p. 277–292, 2017.