

Τεχνητή νοημοσύνη: μύθοι και πραγματικότητα

Παραφέστας Νίκος

02-06-2024

Αγόρι: Μην προσπαθείς να λυγίσεις το κουτάλι. Αυτό είναι αδύνατο. Αντίθετα ... προσπάθησε μόνο να κατανοήσεις την αλήθεια.

Νίο: Ποια αλήθεια;

Αγόρι: Ότι δεν υπάρχει κουτάλι. Τότε θα καταλάβεις ότι δεν είναι το κουτάλι που λυγίζει, αλλά εσύ.

Απόσπασμα από την ταινία "The Matrix" (1999)

Εισαγωγή

Τα ΜΜΕ καθημερινά αναγγέλλουν την έλευση μιας νέας εποχής για την ανθρωπότητα. Η «τεχνητή νοημοσύνη» («ΑΙ») ήρθε για να αλλάξει τα πάντα. Η ακαδημαϊκή κοινότητα (σε μεγάλο μέρος) το αποδέχτηκε, οι επιχειρήσεις το χρηματοδότησαν και το κράτος το νομοθέτησε. Είναι όμως τα πράγματα τόσο απλά; Για αρχή ας ξεκαθαρίσουμε τις έννοιες.

1. Νοημοσύνη

Για να εξετάσουμε την νοημοσύνη των μηχανών πρέπει πρώτα να αναζητήσουμε πως εμείς την κατανοούμε σαν έννοια. Για παράδειγμα, η εγκυκλοπαίδεια "Britannica" δίνει τον ακόλουθο ορισμό¹ για την ανθρώπινη νοημοσύνη:

Νοητική ποιότητα που αποτελείται από την ικανότητα να μαθαίνει κανείς από την εμπειρία, να προσαρμόζεται σε νέες καταστάσεις, να κατανοεί και να χειρίζεται αφηρημένες έννοιες και να χρησιμοποιεί τη γνώση για να χειραγωγήει το περιβάλλον του.

Αυτός ο ορισμός είναι και αρκετά αφηρημένος. Δε θα μπορούσε όμως να είναι διαφορετικά. Η νοημοσύνη είναι μια εξαιρετικά πλατιά έννοια και, σε φιλοσοφικό επίπεδο, υπάρχει πληθώρα προσεγγίσεων για τον ορισμό της.

Μία από αυτές είναι η **θεωρία της πολλαπλής νοημοσύνης**. Σύμφωνα με αυτή, η νοημοσύνη πρέπει να εξετάζεται σε διαφορετικά διακριτά επίπεδα, αντί ενός γενικού. Κατά τον Gardner, προτείνονται τα ακόλουθα 8 είδη (Bacon, 2008) :

• Μουσική νοημοσύνη

Αφορά την ικανότητα στην αντίληψη των ήχων, των τόνων και του ρυθμού της μουσικής.

• Χωροταξική νοημοσύνη

Σχετίζεται με την ικανότητα του μυαλού να αντιλαμβάνεται τον χώρο.

• Γλωσσική νοημοσύνη

Αφορά την ικανότητα αναπαραγωγής γραπτού και προφορικού λόγου.

• Λογικομαθηματική νοημοσύνη

Σχετίζεται με αριθμούς, με τη λογική, την αφαιρετική ικανότητα και την ικανότητα επαγωγικής και απαγωγικής λογικής

• Κινησθητική νοημοσύνη

Καθορίζει την αδρή και λεπτή κινητικότητα και την ικανότητα του σώματος να κινείται με συντονισμό.

• Διαπροσωπική νοημοσύνη

Η ικανότητα των ανθρώπων στην κατανόηση άλλων. Στο να δείχνουν ενσυναίσθηση και να ανταλλάσσουν απόψεις.

• Ενδοπροσωπική νοημοσύνη

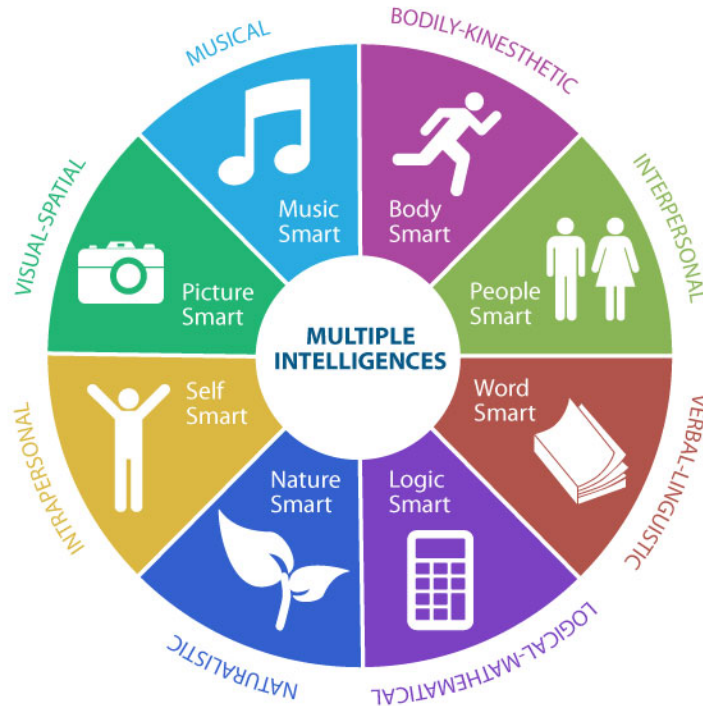
Η κατανόηση του εσωτερικού μας κόσμου.

¹ Sternberg, Robert J.. "human intelligence". Encyclopedia Britannica, <https://web.archive.org/web/20240528050757/https://www.britannica.com/science/human-intelligence-psychology>.

• **Νατουραλιστική νοημοσύνη**

Η σχέση μας με το φυσικό κόσμο.

Τα παραπάνω απικονίζονται σχηματικά και στην εικόνα 1.



Εικόνα 1: Θεωρία της πολλαπλής νοημοσύνης².

2. Τεχνητή Νοημοσύνη

2.1. Εισαγωγή

Στο προηγούμενο κεφάλαιο καταλήξαμε στο ότι η ευφυΐα είναι ένα σύνθετο ζήτημα με πολλαπλές ερμηνείες για να χωρέσει σε έναν απλό ορισμό. Μια ολιστική προσέγγιση, όπως αυτή του Gardner, δίνει μια συγκεκριμένη κατεύθυνση για την κατανόηση της πολυεπίπεδης πολυπλοκότητας της ανθρώπινης υπόστασης. Αυτή η πολυπλοκότητα είναι αποτέλεσμα εκατοντάδων χιλιάδων χρόνων εξέλιξης ενός μονοκύτταρου οργανισμού σε αυτό που πλέον είναι ο σύγχρονος άνθρωπος. Η τεχνητή νοημοσύνη, από την άλλη έχει μόλις λίγα χρόνια ζωής.

2.2. Ιστορία (σύντομη) της τεχνητής νοημοσύνης

Από τα μέσα της προηγούμενης δεκαετίας, από τότε δηλαδή που εμφανίστηκαν οι πρώτοι ηλεκτρονικοί υπολογιστές, εμφανίστηκαν ερευνήτριες που ασχολούνταν με τον τομέα της τεχνητής νοημοσύνης. Τα πιο σημαντικά ιστορικά στάδια του προηγούμενου αιώνα είναι³:

[1927-1946] Η Χρυσή εποχή της επιστημονικής φαντασίας: Είναι η εποχή που ταινίες όπως η "Metropolis" (1927) και «ο Μάγος του ΟΖ» (1936) εμφάνιζαν ανθρωποειδείς μηχανές με συναισθήματα. Είναι οι ταινίες που ενέπνευσαν τους πρώτους επιστήμονες που ασχολήθηκαν με την τεχνητή νοημοσύνη.

² Από Sajaganesandip, <https://wikiless.org/wiki/File:Multiple-intelligence.jpg>, Άδεια: CC BY-SA 4.0

³ <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

[1949] Manchester Mark 1, κατασκευάζεται ο πρώτος υπολογιστής με δυνατότητα αποθήκευσης πληροφοριών: Έως τότε οι υπολογιστές μπορούσαν μόνο να εκτελέσουν εντολές αλλά όχι να αποθηκεύσουν δεδομένα.

[1950] «Μπορούν οι μηχανές να σκέφτονται;» από τον Άλαν Τούρινγκ: Στην μελέτη του ο Τούρινγκ εξετάζει την κατασκευή ευφυών μηχανών και για πρώτη φορά εμφανίζεται το «τεστ Τούρινγκ». Σύμφωνα με αυτό αν δεν μπορεί να γίνει αντιληπτό σε έναν άνθρωπο ότι μία συζήτησή του με μια μηχανή (μέσω τηλεγράφου) δεν είναι με πραγματικό άνθρωπο, τότε η μηχανή «σκέφτεται».

[1955] Κατασκευάζεται το πρώτο πρόγραμμα τεχνητής νοημοσύνης, «Logic Theorists»: Για πολλές το πρώτο πρόγραμμα τεχνητής νοημοσύνης. Πραγματοποιούσε επίλυση μαθηματικών προβλημάτων με σχετική επιτυχία.

[1956] Το καλοκαιρινό ερευνητικό πρόγραμμα Dartmouth: Διάφοροι επιστήμονες συναντήθηκαν για να συζητήσουν για την τεχνητή νοημοσύνη. Εκτός από την καθιέρωση του ονόματος, θεωρείται ότι εκεί γεννήθηκε η τεχνητή νοημοσύνη από τον ιδρυτή του συνεδρίου τον Μαρβιν Μίσκι.

[1963] Η DARPA χρηματοδοτεί το MIT: Οι επιτυχίες στον τομέα της τεχνητής νοημοσύνης, οδηγούν στη χρηματοδότηση πολλών ινστιτούτων από κρατικές υπηρεσίες των ΗΠΑ για έρευνα. Ένα παράδειγμα είναι η Defense Advanced Research Projects Agency.

[1965] Ο νόμος του Μουρ: Πρόκειται για μια ιστορική εμπειρική παρατήρηση, ότι ο αριθμός των τρανζίστορ ενός ολοκληρωμένου κυκλώματος διπλασιάζεται κάθε δύο χρόνια. Αυτή η πρόβλεψη περιγράφει την τάση στην κατασκευή επεξεργαστών ηλεκτρονικών υπολογιστών να αυξάνουν σε ισχύ και να μειώνονται σε μέγεθος. Γενικά, έως σήμερα, έχει επιβεβαιωθεί με μικρές σχετικά αποκλίσεις.

[1968 - 1970] Εποχή της ευφορίας: Καθώς η ευφορία μεγάλωνε (μαζί με την χρηματοδότηση), υπήρχε μεγάλη (τελικά ανεδαφική) αισιοδοξία. Μερικές από τις δηλώσεις της εποχής ήταν:

«Έως το 2001 θα έχουμε μηχανές με νοημοσύνη η οποία θα ξεπερνάει αυτή των ανθρώπων»⁴

«Σε 3 με 8 χρόνια θα έχουμε μηχανή με γενική εφύια όση ενός ανθρώπου»⁵

[1980] Ο Εντουαρντ Φειγκενμπάουερ παρουσιάζει τα «έμπειρα συστήματα»: Το πρόγραμμα ρωτούσε έναν εμπειρογνώμονα σε έναν τομέα πώς να ανταποκριθεί σε μια δεδομένη κατάσταση και, μόλις αυτό γινόταν πρακτικά για κάθε περίπτωση, οι μη ειδικοί θα μπορούσαν να λάβουν συμβουλές από αυτό. Τα έμπειρα συστήματα χρησιμοποιήθηκαν ευρέως στις βιομηχανίες.

[1980] Το «πρόγραμμα 5ης γενιάς υπολογιστικών συστημάτων» από την Ιαπωνία: Η Ιαπωνία έως το 1990 διέθεσε 400 εκατομμύρια δολάρια στην έρευνα σχετικά με την τεχνητή νοημοσύνη. Οι περισσότεροι στόχοι του προγράμματος απέτυχαν και η χρηματοδότηση σταμάτησε.

[1986] Κατασκευάζεται το Navlab, το πρώτο αυτόνομο όχημα

[1997] Ο υπερυπολογιστής «Deep Blue», νικάει τον Κασπάροφ στο σκάκι: Αποτέλεσε πρώτη είδηση σε όλα τα δελτία ειδήσεων της εποχής. Ο υπολογιστής ανέλυε όλες τις πιθανές κινήσεις (brute force) του αντιπάλου για να επιλέξει την επόμενη του κίνηση.

[1997] Το πρώτο ευρείας κυκλοφορίας πρόγραμμα μετατροπής γλώσσας σε φωνή από την Dragon Systems.

⁴ Άρθουρ Κλαρκ και Στιβ Κουμπρικ

⁵ Δήλωση του Μαρβιν Μίσκι στο περιοδικό Time

2.3. Ορισμοί Τεχνητής Νοημοσύνης

Ο στόχος λοιπόν, για όσες ασχολήθηκαν με την τεχνητή νοημοσύνη, από τη σύλληψη της ως έννοια, ήταν:

Τεχνητή Νοημοσύνη \geq Ανθρώπινη Νοημοσύνη

Αυτό το θεωρητικό μοντέλο τεχνητής νοημοσύνης, που απαιτεί μια μηχανή να έχει νοημοσύνη ίση ή μεγαλύτερη από τον άνθρωπο, αν υπήρχε, θα είχε συνειδητή αυτογνωσία, την ικανότητα να λύνει προβλήματα και να μαθαίνει και να σχεδιάζει το μέλλον. Αυτή ονομάζεται ισχυρή ή γενική τεχνητή νοημοσύνη. Παρά τις, με βεβαιότητα, δηλώσεις επιστημόνων (μόλις 50 χρόνια πριν) ότι απέχουμε από λίγα χρόνια έως λίγες δεκαετίες από αυτόν τον στόχο, τελικά ως σήμερα δεν έχει επιτευχθεί⁶. Ειδικές περιπτώσεις, όπως αυτές της ήττας του Κασπάροφ στο σκάκι αποδίδονται όχι στην αυξημένη νοημοσύνη των υπολογιστών, αλλά στην μεγαλύτερη επεξεργαστική τους δυνατότητα να εκτελούν πράξεις, σε σχέση με εμάς. Δεν μπορούμε να ισχυριστούμε, για παράδειγμα, ότι ένα απλό αριθμητικό κομπιουτεράκι μας ξεπερνά σε νοημοσύνη. Όταν μια μηχανή μπορεί να κάνει περιορισμένες ενέργειες που προσομοιάζουν τμήματα της ανθρώπινης νοημοσύνης, τότε ονομάζεται αδύναμη ή στενή τεχνητή νοημοσύνη.

Σήμερα η ορολογία έχει αλλάξει, χωρίς, όμως, να υπάρχει ένας αμοιβαία συμφωνημένος ορισμός. Για παράδειγμα, η εταιρεία OPENAI η οποία έχει μονοπωλήσει τα φώτα της δημοσιότητας με το πρόγραμμα chatGPT (στο οποίο θα αναφερθούμε σε επόμενη παράγραφο), ορίζει την γενικευμένη τεχνητή νοημοσύνη ως⁷:

«...εξαιρετικά αυτόνομα συστήματα που ξεπερνούν τους ανθρώπους στην πιο οικονομικά επικερδή εργασία»

Πρόκειται για μια εξαιρετικά αφηρημένη και βολική ερμηνεία. Είναι όμως ταυτόχρονα ξεκάθαρη για τον βασικό στόχο της εταιρείας σε σχέση με την τεχνητή νοημοσύνη. Η ανάπτυξή της σε δραστηριότητες που αποσκοπούν σε κερδοφορία.

Σε τελείως διαφορετική κατεύθυνση σε μία μελέτη του ο Yoshihiro Maruyama, από το Εθνικό Πανεπιστήμιο της Αυστραλίας, ορίζει τα παρακάτω 8 χαρακτηριστικά ενός συστήματος ως απαραίτητα για να θεωρηθεί γενικευμένη τεχνητή νοημοσύνη (Maruyama, 2020):

- Λογική
- Αυτονομία
- Ανθεκτικότητα
- Ακεραιότητα
- Ηθική
- Συναισθημα
- Ενσάρκωση
- Ενσωμάτωση

Παρατηρούμε ότι η παραπάνω κατηγοριοποίηση, προσεγγίζει σημαντικά τον χαρακτηρισμό της ανάλυσης για την ανθρώπινη ευφυΐα του Gardner.

⁶ <https://web.archive.org/web/20240508212358/https://www.brookings.edu/articles/how-close-are-we-to-ai-that-surpasses-human-intelligence/>

⁷ <https://web.archive.org/web/20240508231921/https://openai.com/charter/>

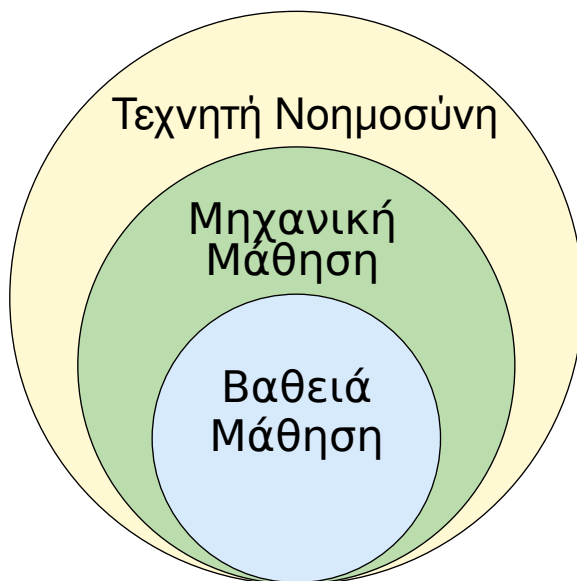
2.4. Βασικές έννοιες

2.4.1. Τεχνικές τεχνητής νοημοσύνης

Οι έννοιες τεχνητή νοημοσύνη, μηχανική μάθηση (machine learning), βαθιά μάθηση (deep learning) και νευρωνικά δίκτυα (neural networks) ακούγονται και συγχέονται συχνά. Παρακάτω γίνεται προσπάθεια να ξεδιαλύνουμε αυτές τις ορολογίες.

Μηχανική μάθηση

Η βαθιά μάθηση είναι υποσύνολο της μηχανικής μάθησης η οποία με τη σειρά της είναι και αυτή υποσύνολο της τεχνητής νοημοσύνης (εικόνα 2).



Εικόνα 2: Ιεραρχία τεχνητής νοημοσύνης⁸.

Η μηχανική μάθηση, είναι μία μέθοδος εκπαίδευσης των μηχανών από τα δεδομένα ή την εμπειρία η οποία αυτοματοποιεί τη δημιουργία αναλυτικών μοντέλων. Ενώ η μηχανική μάθηση έχει μερικά εξελιγμένα μαθηματικά «κόλπα στο μανίκι της», είναι σημαντικό να κατανοήσουμε ότι είναι μια μαθηματική διαδικασία ωμής βίας. Όταν παρέχονται αρκετά δεδομένα εκπαίδευσης και όταν εκτελούνται σε αρκετά ισχυρούς υπολογιστές, μπορεί να αξιοποιήσει τις αριθμητικές λειτουργίες σε μια ασυνήθιστη εξομίωση διαφόρων ανθρώπινων ικανοτήτων, όπως η ικανότητα αναγνώρισης προσώπων ή η αναπαραγωγή επιτραπέζιων παιχνιδιών στρατηγικής όπως το Go⁹.

Βαθιά μάθηση

Η «βαθιά μάθηση» αναφέρεται σε προσεγγίσεις μάθησης που βασίζονται σε δεδομένα που χρησιμοποιούν πολυεπίπεδα νευρωνικά δίκτυα και επεξεργασία για υπολογισμό. Ο όρος «βαθιά» αναφέρεται στην έννοια πολλών επιπέδων ή σταδίων μέσω των οποίων τα δεδομένα υποβάλλονται σε επεξεργασία για την ανάπτυξη ενός μοντέλου που βασίζεται σε δεδομένα. (Sarker, 2022)

⁸ Παράγωγο από Lollixzc, <https://wikiless.org/wiki/File:Multiple-intelligence.jpg>, διατίθεται με όμοια άδεια: CC BY-SA 4.0

⁹ Στρατηγικό επιτραπέζιο παιχνίδι που ξεκίνησε από την Κίνα. Οι δυο παίχτριες, άσπρη και μαύρη, προσθέτουν εναλλάξ πέτρες στο ταμπλό, περικυκλώνοντας περιοχή, φυλακίζοντας πέτρες της αντιπάλου και προστατεύοντας τις δικές τους πέτρες.

Βελτιστοποίηση

Η μηχανική μάθηση ενσωματώνει την ιδέα ότι ο τρόπος επίλυσης ενός προβλήματος είναι να βρεθεί ένας στόχος για βελτιστοποίηση. Η βελτιστοποίηση είναι ένα συγκεκριμένο είδος ορθολογισμού, που υποστηρίζει ότι η συμπύκνωση της πολυπλοκότητάς (του στόχου) του σε έναν υπολογισμό παρέχει ένα ανώτερο είδος λύσης. Ο ρόλος του βελτιστοποιητή είναι να επαναλαμβάνει επανειλημμένα τα δεδομένα εκπαίδευσης μέχρι να ελαχιστοποιήσει τη συνάρτηση απώλειας, δηλαδή την απόκλιση ανάμεσα στα δεδομένα που εισήχθησαν και την πρόγνωση που πρέπει να κάνει η μηχανή για να προσφέρει λύση ή, πιο απλά, το ενδεχόμενο η μηχανή να κάνει λάθος. Όταν η απόκλιση αποφασιστεί (από άτομο ή μηχανή) ότι είναι αποδεκτή, το σύστημα μηχανικής μάθησης θεωρείται ότι έχει εκπαιδευτεί (Sarker, 2022).

Δεδομένα

Για να υπάρξει, όμως, βελτιστοποίηση, πρέπει να υπάρχουν δεδομένα. Όσα περισσότερα, μάλιστα, τόσο μικρότερη η συνάρτηση απώλειας. Για αυτό, υπάρχει αυτή η μανία για τα πολλά μαζικά δεδομένα (big data) από τις εταιρείες. Είναι μάλιστα τόσο πολύτιμα τα δεδομένα, που οι εταιρείες-γίγαντες της τεχνολογίας (big-tech) πλέον στοχεύουν στην αχόρταγη απόκτηση των δεδομένων που προκύπτουν από τις χρήστριες των «υψηλών τους».

Νευρωνικά δίκτυα

Τα νευρωνικά δίκτυα από την άλλη, αποτελούν τη ραχοκοκαλιά των αλγορίθμων βαθιάς μάθησης. Είναι ο αριθμός των επιπέδων κόμβου ή το βάθος (τα επίπεδα) των νευρωνικών δικτύων που διακρίνει ένα μεμονωμένο νευρωνικό δίκτυο από έναν αλγόριθμο βαθιάς μάθησης. Με την σημερινή τεχνολογία, η μηχανική μάθηση είναι ουσιαστικά ισοδύναμη έννοια με τα συστήματα νευρωνικών δικτύων¹⁰

2.4.2. Το chatGPT

Το πιο διαδεδομένο πρόγραμμα σήμερα, όπως προωθείται κάθε φορά που ΜΜΕ, στελέχη επιχειρήσεων αλλά ακόμα και ακαδημαϊκών που θέλουν να αναφερθούν στην τεχνητή νοημοσύνη είναι το chatGPT. Το ενδιαφέρον της κοινής γνώμης είναι υψηλό, ενώ, ταυτόχρονα, είναι το πιο γρήγορα διαδεδομένο πρόγραμμα έως σήμερα, φθάνοντας τα 100 εκατομμύρια χρήστριες, μόλις σε 2 μήνες μετά την διάθεσή του¹¹.

Το υψηλό ενδιαφέρον όμως δεν σημαίνει πάντα πραγματική καινοτομία. Υπάρχουν πολλοί ερευνητές οι οποίοι δε συμφωνούν με την άποψη ότι το chatGPT είναι πραγματικά καινοτόμο. Σε μια συνέντευξή του στην Deutsche Welle, ο Dr Michael Cook, λέκτορας στο King's College του Λονδίνου, σημείωσε ότι προγράμματα σαν και αυτό δημιουργούν μια ψευδαίσθηση για το πόσο προηγμένα είναι. Ταυτόχρονα πρόσθεσε:

Το ότι ένας σκύλος μπορεί να κάνει χειραψία, όπως και ο Αϊνστάιν, δεν μπορεί να σημαίνει ότι το σκυλί είναι το ίδιο έξυπνο με τον Αϊνστάιν.

Ακόμα πιο αποφασιστικά, ο Νόαμ Τσόμσκι αναφέρει¹²:

Εν ολίγοις, το ChatGPT και τα αδέρφια του δεν είναι σε θέση να εξισορροπήσουν τη δημιουργικότητα με τους περιορισμούς. Είτε υπεργεννούν, παράγοντας και αλήθειες και ψεύδη, υποστηρίζοντας ηθικές και ανήθικες αποφάσεις εξίσου, είτε υπογεννούν, δείχνοντας μη δέσμευση σε ποιοσδήποτε αποφάσεις και αδιαφορία για τις συνέπειες. Δεδομένης της ηθικής, της ψεύτικης επιστήμης και της γλωσσικής ανικανότητας αυτών των συστημάτων, δεν μπορούμε παρά να γελάσουμε ή να κλάψουμε με τη δημοτικότητά τους.

Ενώ άλλοι, όπως ο Dr Dan McQuillan και ο πατέρας του ελεύθερου λογισμικού Dr Ritchard Stallman, χαρακτηρίζουν το chatGPT ως «μηχανή παραγωγής σαχλαμάτων»¹³¹⁴.

¹⁰ <https://www.gnu.org/philosophy/words-to-avoid.html#ArtificialIntelligence>

¹¹ <https://neuters.de/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

¹² <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

¹³ <https://www.vice.com/en/article/akex34/chatgpt-is-a-bullshit-generator-waging-class-war>

¹⁴ <https://www.gnu.org/philosophy/words-to-avoid.html#ArtificialIntelligence>

Προγράμματα επομένως σαν το chatGPT, έχουν αμφίβολα αποτελέσματα και, όπως θα δούμε παρακάτω, είναι αδύνατο να ελεγχθούν. Συνεπώς είναι απαραίτητο να εξετάζονται κριτικά από κράτη/επιχειρήσεις/επιστήμονες και να μην παρασύρονται από την «ευκολία» και τις πομπώδεις επικεφαλίδες. Όπως θα δούμε παρακάτω, υπάρχουν σημαντικά ηθικά ζητήματα που προκύπτουν τα οποία δεν πρέπει να παραβλέπονται.

3. Ηθικά ζητήματα γύρω από την τεχνητή νοημοσύνη

3.1. Ποιος ελέγχει τις μηχανές

Θα αποφασίζουν για εμάς οι μηχανές; Μήπως θα χάσουμε την θέση μας ως το πιο εξελιγμένο είδος; Είναι τελικά η τεχνητή νοημοσύνη το κουτί της Πανδώρας;

Τα παραπάνω ερωτήματα είναι συνηθισμένα και, συχνά, προκαλούν πανικό. Αρχικά, όπως έχει αναφερθεί, είμαστε τεχνολογικά πολύ μακριά από το να κατασκευάσουμε μηχανές με αυτογνωσία που να έχουν την ικανότητα να αποφασίζουν για το πώς θα δράσουν ανεξάρτητα από ανθρώπινη παρέμβαση. Είναι ακόμα οι άνθρωποι αυτοί που αποφασίζουν το πώς θα δράσει ο αλγόριθμος των μηχανών, τόσο σε επίπεδο μηχανικής μάθησης όσο και παραγωγής αποτελέσματος. Μάλιστα στις περισσότερες περιπτώσεις, απαιτούνται εργάτριες για να «τακτοποιήσουν» τα απέραντα δεδομένα, τα οποία συλλέγονται και τροφοδοτηθούν τον αλγόριθμο (ώστε να εκπαιδευτεί), βάζοντας ετικέτες¹⁵. Παράδειγμα είναι η υπερεργολαβική συνεργασία του chatGPT με εταιρεία ενοικίασης 50.000 εργαζομένων στην Κένυα για επισήμανση, με μισθό λιγότερο από 2\$/ώρα¹⁶.

3.2. Τεχνολογία «μαύρο κουτί»

Οι εταιρείες που παράγουν λογισμικό σχετικό με την τεχνητή νοημοσύνη, αλλά και αυτές που παρέχουν σχετικές υπηρεσίες¹⁷, βολεύονται μέσω της αποποίησης ευθυνών που προσφέρει η διατήρηση του μύθου για ανεξάρτητα «σκεπτόμενα μηχανήματα». Είδαμε όμως ότι στην πραγματικότητα είναι ανθρώπινη επιλογή το τι δεδομένα εισάγονται, πώς αξιοποιούνται από τα συστήματα και τελικά πώς εξάγονται και παρουσιάζονται στους χρήστες. Η εικόνα 3 σχηματοποιεί τα τρία στάδια που αναφέρθηκαν.



Εικόνα 3: Λειτουργία "AI" ¹⁸.

Δεδομένα

Ο τρόπος συλλογής των δεδομένων, τα οποία, όπως έχει αναλυθεί, είναι απαραίτητα για την εκπαίδευση των μηχανών, αποτελεί μια από τις πιο προβληματικές παραμέτρους της τεχνητής νοημοσύνης. Εταιρείες σαν την openAI, βρίσκονται συχνά στο επίκεντρο με κατηγορία ότι χρησιμοποιούν online περιεχόμενο χωρίς άδεια.

Η εφημερίδα New York times, για παράδειγμα, έκανε αγωγή σε Microsoft και OpenAI, ζητώντας αποζημίωση δισεκατομμυρίων δολλαρίων για τη χρήση περιεχομένου της εφημερίδας χωρίς να έχουν εξουσιοδότηση¹⁹. Φυσικά, δεν είναι η μόνη περίπτωση. Παρακάτω περιγράφονται μερικές μόνο από τις

¹⁵ Η επισήμανση δεδομένων συμβάλει στον προσδιορισμό ακατέργαστων δεδομένων (δηλαδή εικόνων, αρχείων κειμένου, βίντεο) και, στη συνέχεια, την προσθήκη μιας ή περισσότερων ετικετών σε αυτά ώστε ένα μοντέλο μηχανικής εκμάθησης να πραγματοποιήσει πιο ακριβείς προβλέψεις.

¹⁶ <https://time.com/6247678/openai-chatgpt-kenya-workers/>

¹⁷ Υπηρεσίες παρέχουν πλέον όχι μόνο μικρές και μεγάλες ιδιωτικές εταιρείες αλλά και κρατικές υπηρεσίες.

¹⁸ Από Παραφέστα Νίκο, διατίθεται με άδεια: CC BY-SA 4.0

¹⁹ <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html>

περιπτώσεις που οι "big tech" συλλέγουν αχόρταγα δεδομένα των χρηστών τους²⁰:

- Η Google χρησιμοποιεί τα δεδομένα που προκύπτουν κατά τη χρήση των υπηρεσιών Youtube, gmail, google drive κ.α.
- Η Microsoft πέρα από την εμπλοκή της στο chatGPT, χρησιμοποιεί ακόμα και δεδομένα χρήσης των windows για να συγκεντρώσει υλικό για μηχανική μάθηση. Το ίδιο ισχύει και για πλατφόρμες όπως το Outlook mail.
- Το X(itter) απέκλεισε την πρόσβαση προγραμμάτων στην πλατφόρμα για να συλλέγει έχει αποκλειστικότητα στο περιεχόμενο
- Η Meta έχει όρο στη χρήση των πλατφορμών της, facebook, instagram κτλ, ότι τα δεδομένα που διακινούνται περνάνε αυτόματα στην ιδιοκτησία της εταιρείας.

Με τον τρόπο όμως που πραγματοποιείται αυτή η συλλογή δεδομένων, προκύπτουν και ζητήματα υποεκπροσώπησης διαφόρων ομάδων. Για παράδειγμα έλεγχος ενός προγράμματος αναγνώρισης χαρακτηριστικών προσώπων είχε αρκετά καλά αποτελέσματα για λευκά άτομα, αλλά όχι για πιο σκουρόχρωμους²¹. Αυτό γιατί η «εκπαίδευση» του συστήματος, πραγματοποιήθηκε με βάση τα στάνταρ της «Silicon Valley», περιοχής που με συντριπτική πλειοψηφία λευκού πληθυσμού.

Για να υπάρξει έλεγχος στο πως τα δεδομένα μας χρησιμοποιούνται είναι απαραίτητο να υπάρχει απόλυτη διαφάνεια στο τι είναι αυτά τα δεδομένα, πως συλλέχθηκαν αλλά και αν είναι διαθέσιμα προς χρήση σε όλα. Έτσι μόνο θα έχουμε τη δυνατότητα να αξιολογούμε, να ελέγχουμε περιπτώσεις παραβιάσεων της ιδιωτικότητάς μας αλλά ταυτόχρονα και να βελτιώνουμε την αξιοπιστία τους.

Αλγόριθμος, αυτός ο άγνωστος

Παρόλο που το όνομα της εταιρείας openAI προδιαθέτει σε ανοιχτή (open) λειτουργία, η αλήθεια είναι πολύ διαφορετική. Στην πραγματικότητα, τα περισσότερα συστήματα τεχνητής νοημοσύνης, είναι εμπορικά προγράμματα, των οποίων ο κώδικας δεν είναι προσβάσιμος. Αυτό σημαίνει ότι δεν υπάρχει τρόπος να εξεταστεί από τρίτα μέρη το πώς επεξεργάζονται τα εισαγόμενα δεδομένα. Καθώς το πως λειτουργεί ένας αλγόριθμος εξαρτάται καθαρά από το πως θα ρυθμιστεί από τον χειριστή του, δεν μπορούμε να έχουμε γνώση για ζητήματα όπως, τον τρόπο «εκπαίδευσης» του αλγόριθμου και το κατά πόσο επιλέγονται ορισμένες παράμετροι να έχουν μεγαλύτερη βαρύτητα από άλλες.

Την ίδια ώρα, τα συστήματα αυτά, ήδη χρησιμοποιούνται σε πληθώρα κρίσιμων λειτουργιών, όπως στην υγεία, στην παιδεία (όλων των σχολικών βαθμίδων), στην λειτουργία των δικαστηρίων και σε άλλες κρίσιμες κρατικές υποδομές. Η χρήση του, μάλιστα, αναμένεται να αυξηθεί ακόμα περισσότερο το επόμενο διάστημα. Πώς όμως μπορούμε να εμπιστευτούμε, σε τόσο νευραλγικούς τομείς της κοινωνίας και της δημοκρατίας, αυτά τα «μαύρα κουτιά» την αξιοπιστία των οποίων οποίων στερούμαστε να ελέγξουμε; Πως μπορούμε να ελέγξουμε, για παράδειγμα, αν μία γιατρός εμπιστεύτηκε σωστά μια εκτίμηση θνησιμότητας από ένα πρόγραμμα, με αποτέλεσμα να στερήσει ένα φάρμακο από έναν ασθενή;

Στην εκπαίδευση (δευτεροβάθμια, τριτοβάθμια και πανεπιστημιακή), η χρήση του chatGPT έχει πλέον απογειωθεί. Το σύστημα ελέγχου λογοκλοπής Turnitin, ανακοίνωσε ότι το 2023 περίπου 14 εκατομμύρια φοιτητικές εργασίες είναι πιθανό να χρησιμοποιήσαν μορφές τεχνητής νοημοσύνης και ότι περίπου το 75% των μαθητριών στο Λύκειο χρησιμοποιεί συστήματα μηχανικής μάθησης ακόμα και αν αυτά έχουν απαγορευτεί/**. Παράλληλα, σε επίπεδο έρευνας, ακόμα και μέλη της ακαδημαϊκής κοινότητας χρησιμοποιούν την τεχνητή νοημοσύνη, τόσο για τη συγγραφή ερευνών όσο και για την ομότιμη αξιολόγηση για την δημοσίευσή τους σε επιστημονικά περιοδικά²². Αυτό συμβαίνει ενώ δεν υπάρχει καμία

²⁰ <https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html>

²¹ <https://www.turnitin.com/blog/one-year-of-chatgpt-where-are-we-now>

επιστημονική γνώση για την αξιοπιστία τους.

Μαθησιακά, καθώς η χρήση προγραμμάτων όπως το chatGPT στην εκπαίδευση είναι πρόσφατη, δεν υπάρχουν ακόμα ασφάλειες μελέτες για τα ζητήματα που ενδέχεται να προκύψουν. Μπορούμε όμως να κάνουμε μια εκτίμηση παρατηρώντας τα αποτελέσματα της μακροχρόνιας χρήσης, μια σαφώς πιο ήπιας τεχνολογίας, στην εκπαίδευση, της χρήσης ορθογράφου. Συγκεκριμένα, μια μελέτη του Χάρβαρντ του 2005 (Galleta, 2005) διαπίστωσε ότι 37 άτομα που παρήγαγαν μια μελέτη κάνοντας χρήση ορθογραφικού ελέγχου - έναντι 28 ατόμων χωρίς - βασίζονταν ουσιαστικά στην τεχνολογία για να κάνουν την ορθογραφία για αυτούς. Αυτό δεν είχε να κάνει με τις ικανότητές τους στην ορθογραφία, αλλά αντίθετα στην τεμπελιά και την αυτοπεποίθηση ότι δεν την χρειάζονταν. Αναρωτιόμαστε, λοιπόν: αν οι ορθογράφοι μας έκαναν ανορθόγραφους²³, τότε τι θα συμβεί αν μια, μη αξιολογημένη μηχανή, εκπονεί και διορθώνει τις εργασίες μαθητριών, φοιτητών και μελών της ακαδημαϊκής κοινότητας;

Τα ζητήματα, όμως, δε σταματούν στην παιδεία. Ήδη υπάρχουν σκέψεις για χρήση της τεχνητής νοημοσύνης ακόμα και στις δικαστικές αίθουσες στην Ελλάδα. Συγκεκριμένα στο τμήμα «Τεχνητή Νοημοσύνη στον τομέα της Δικαιοσύνης» στην παράγραφο 9.4.2 της «βίβλου ψηφιακού μηχανισμού 2020-2025»²⁴ (ΦΕΚ 2894/Β/5-7-2021) υπάρχει αναφορά για υποβοήθηση στην λήψη αποφάσεων των δικαστικών οργάνων από προγράμματα τεχνητής νοημοσύνης, ακόμα και για αυτόματη λήψη αποφάσεων για συγκεκριμένες υποθέσεις. Η εξοικονόμηση χρόνου και χρήματος που αναφέρεται ως αιτία, δεν δικαιολογεί την χρήση ενός μη διαπιστευμένου λογισμικού. Οι συνέπειες από μια πιθανή λάθος καταδίκη ανθρώπων από λανθασμένη εκτίμηση μπορεί να είναι ολέθριες για το δικαστικό σύστημα. Η περίπτωση του Compass, ενός προγράμματος εκτίμησης της πιθανότητας ένας ένοχος να τελέσει εκ νέου παραβάσεις, είναι χαρακτηριστική: το συγκεκριμένο λογισμικό, χρησιμοποιούνταν από τις δικαστικές των ΗΠΑ, ώστε να αποφασίσουν το ύψος της κάθειρξης. Μια έρευνα έδειξε όμως, ότι το συγκεκριμένο πρόγραμμα χρησιμοποιούσε αλγόριθμο ο οποίος ήταν εξαιρετικά μεροληπτικός απέναντι σε σκουρόχρωμες σε σχέση με λευκές²⁵. Αποτέλεσμα ήταν να υπάρξει παρέμβαση μεγάλης μερίδας δικαστικών και νομικών, ζητώντας να ανοιχτεί ο κώδικας τόσο του συγκεκριμένου προγράμματος όσο και των υπολοίπων που χρησιμοποιούν τη μέθοδο της πρόβλεψης με τεχνολογίες μηχανικής μάθησης, ώστε να ελεγχθούν²⁶.

3.3. AI Act

Στις 13 Μαρτίου 2024, η ευρωβουλή ενέκρινε το AI Act. Πρόκειται για έναν νόμο ο οποίος, μετά από 3 χρόνια διαβουλεύσεων, θα οριοθετούσε την τεχνητή νοημοσύνη. Παρόλο που όντως αποτελεί παγκόσμια το πρώτο μεγάλης κλίμακας νομοθέτημα, άφησε μια γλυκόπικρη γεύση²⁷.

Αρχικά, γιατί τόσο η δομή του, όσο και οι ασάφειες, ανοίγουν τον δρόμο στις γιγάντιες εταιρείες πληροφορικής να ενισχύσουν το μονοπώλιό τους²⁸. Αυτό συμβαίνει εν πρώτοις, γιατί υπάρχουν «θολές» διατάξεις τις οποίες μπορούν να εκμεταλλευτούν τα μεγάλα νομικά τους τμήματά. Επιπλέον, με δεδομένο ότι εταιρείες όπως Meta, Apple, Google, Microsoft ξοδεύουν πάνω από 100 εκατομμύρια ευρώ για να ασκήσουν πολιτική πίεση στην Ε.Ε./** και, συχνά όχι με διαφανείς μεθόδους²⁹, οι ασάφειες στο AI Act

²² <https://web.archive.org/web/20240120015516/https://onlinelibrary.wiley.com/doi/full/10.1002/leap.1570>

²³ <https://www.bbc.co.uk/news/education-18158665>

²⁴ https://digitalstrategy.gov.gr/website/static/website/assets/uploads/digital_strategy.pdf

²⁵ https://web.archive.org/web/20240612092819if_/https://www.propublica.org/article/how-we-analyzed-the-compass-recidivism-algorithm

²⁶ https://web.archive.org/web/20240329055803if_/https://www.theregister.com/2024/02/16/black_box_government/

²⁷ <https://www.euronews.com/next/2024/03/16/eu-ai-act-reaction-tech-experts-say-the-worlds-first-ai-law-is-historic-but-bittersweet>

²⁸ <https://www.aljazeera.com/news/2024/3/13/eu-parliament-greenlights-landmark-artificial-intelligence-regulations>

<https://www.euronews.com/my-europe/2023/09/11/tech-companies-spend-more-than-100-million-a-year-on-eu-digital-lobbying>

²⁹ <https://www.euronews.com/my-europe/2023/09/11/tech-companies-spend-more-than-100-million-a-year-on-eu-digital-lobbying>

ευνοούν την δυνατότητα να ασκηθεί επιπλέον πίεση και να υπάρχει ακόμα μεγαλύτερη ευνοϊκή μεταχείριση των εν λόγω εταιρειών.

Πέρα όμως από τον αθέμιτο ανταγωνισμό, ανακύπτουν και ζητήματα δημοκρατίας. Συγκεκριμένα, στους κανόνες γύρω από τη χρήση της τεχνητής νοημοσύνης, υπάρχουν σημαντικές εξαιρέσεις:

- Στα άρθρα 2.3 και 2.6, εξαιρούνται τα συστήματα τεχνητής νοημοσύνης για λόγους εθνικής ασφαλείας και
- στο άρθρο 5.2, η απαγόρευση της χρήσης συστημάτων τεχνητής νοημοσύνης σε πραγματικό χρόνο αίρεται για λόγους αστυνόμευσης.

Οι παραπάνω εξαιρέσεις, ειδικά μετά το σκάνδαλο παρακολουθήσεων με το Predator στην Ελλάδα³⁰ και το «ξεχειλίωμα» των ζητημάτων εθνικής ασφάλειας, δημιουργούν τεράστιες δημοκρατικές οπισθοχωρήσεις και μπορούμε να ισχυριστούμε ότι έρχονται ενάντια στο πνεύμα του άρθρου 19 του συντάγματος.

4. Συμπεράσματα και προτάσεις

Η κριτική προσέγγιση του συγγραφέα της παρούσας εργασίας στα συστήματα της τεχνητής νοημοσύνης, δεν έχει να κάνει με την ίδια την τεχνολογία αλλά με τις υπερβολές και την παραπληροφόρηση. Άλλωστε, καμία τεχνολογία δεν είναι από μόνη της ικανή για καλό ή για κακό. Η πυρηνική ενέργεια για παράδειγμα μπορεί να εξαυλώσει τον πλανήτη μπορεί, όμως, και υπό προϋποθέσεις να προσφέρει απεριόριστη καθαρή ενέργεια. Έτσι και η τεχνητή νοημοσύνη: παρόλο τους κινδύνους που παρουσιάστηκαν, μπορεί να αποτελέσει σημαντικό εργαλείο έρευνας και επιστήμης. Απαραίτητο είναι, όμως, να υπάρχει διαφάνεια ώστε να μπορούμε να έχουμε όλες αυτές τις αναγκαίες διαβεβαιώσεις ότι μπορούμε να την εμπιστευτούμε. Ήδη έχουμε την εμπειρία ότι αξιόπιστο λογισμικό είναι το λογισμικό στο οποίο μπορούμε να έχουμε πρόσβαση - το ελεύθερο λογισμικό.

Το Ελεύθερο λογισμικό παρέχει στους χρήστες την ελευθερία να εκτελούν, αντιγράφουν, διανέμουν, μελετούν, τροποποιούν και βελτιώνουν το Ελεύθερο λογισμικό. Για την ακρίβεια, αναφέρεται σε τέσσερις βασικές ελευθερίες³¹:

- Την ελευθερία να εκτελείτε το πρόγραμμα, για οποιονδήποτε σκοπό (ελευθερία 0).
- Την ελευθερία να μελετάτε τον τρόπο λειτουργίας του προγράμματος και να το προσαρμόζετε στις ανάγκες σας (ελευθερία 1). Η πρόσβαση στον πηγαίο κώδικα είναι προϋπόθεση για να ισχύει κάτι τέτοιο.
- Την ελευθερία να αναδιανέμετε αντίγραφα του προγράμματος ώστε να βοηθάτε το συνάνθρωπο σας (ελευθερία 2).
- Την ελευθερία να βελτιώνετε το πρόγραμμα και να δημοσιεύετε τις βελτιώσεις που έχετε κάνει στο ευρύ κοινό, ώστε να επωφεληθεί ολόκληρη η κοινότητα (ελευθερία 3). Η πρόσβαση στον πηγαίο κώδικα είναι προϋπόθεση για να ισχύει κάτι τέτοιο.

Αντί σαν είδος να ξοδεύουμε απεριόριστους πόρους, ανθρώπινους και χρηματικούς, σε εμπορικά κλειστά συστήματα με αμφίβολο αποτέλεσμα, ας επενδύσουμε σε ελεύθερα, με ανοιχτά πρότυπα και δεδομένα που θα ωφελήσουν την κοινωνική μας πρόοδο.

³⁰ http://web.archive.org/web/20231101155720if_/https://www.efsyn.gr/ellada/koinonia/207015_polites-parakoloythoyntai-i-dimokratia-apeileitai

Βιβλιογραφία

- Allyn & Bacon, "Educational Psychology: Theory and Practice" in *London* (2008). ISBN-10: 0-205-59200-7.
- Maruyama, Y., "The Conditions of Artificial General Intelligence: Logic, Autonomy, Resilience, Integrity, Morality, Emotion, Embodiment, and Embeddedness," R. (eds) *Artificial General Intelligence. AGI 2020. Lecture Notes in Computer Science()*, vol 12177. Springer, Cham. (2020). https://doi.org/10.1007/978-3-030-52152-3_25.
- Sarker, I.H., "AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems," *N COMPUT. SCI.* 3, 158 (2022). <https://doi.org/10.1007/s42979-022-01043-x>.
- Galleta,, "Does spell-checking software need a warning label?" *Communications of the ACM* Volume 48 Issue 7pp 82-86 (2005). <https://doi.org/10.1145/1070838.1070841>.