

ΑΡΧΙΜ. ΙΓΝΑΤΙΟΣ ΜΟΥΡΤΖΑΝΟΣ
Δρ ΘΕΟΛΟΓΙΑΣ ΚΑΙ ΒΥΖΑΝΤΙΝΗΣ ΦΙΛΟΛΟΓΙΑΣ
ΠΡΩΤΟΣΥΓΚΕΛΛΟΣ Ι. Μ. ΛΑΡΙΣΗΣ ΚΑΙ ΤΥΡΝΑΒΟΥ

ΑΠΟ ΤΗΝ ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΣΤΗΝ ΥΠΕΡΝΟΗΜΟΣΥΝΗ

Η καθόλα δημιουργική φαντασία των έρευνητών των σημερινών κοινο-
νιών, συνεπικουρούμενη βεβαίως πάντοτε καί από τίν ιλιγγιώδη τα-
χύτητα των εξελίξεων, παρέχει ένδειξεις εκ των προτέρων για τά έπερχόμενα
βήματα του ανθρώπου στην Τεχνητή Νοημοσύνη, συχνά συγχέοντας τίν έπι-
στήμη μέ τίν έπιστημονική φαντασία.

Ό προβληματισμός δέν είναι καινούργιος, αλλά τόν έχει ξεκινήσει τό 1637
στό βιβλίο V «Περί τής μεθόδου» του ό Rene Descartes, ό όποιος έρεύνησε εάν
θά μπορούσε νά κατασκευαστεί κάποια μηχανή μέ πλήρη ψυχική ζωή. Χαρα-
κτηριστικά αναφέρει: «Ένώ ή νοημοσύνη είναι ένα καθολικό μέσο πού μπορεί
νά χρησιμεύσει για όλες τīs απρόβλεπτες καταστάσεις, οι μηχανές χρειάζονται
κάποια ειδική προσαρμογή για κάθε συγκεκριμένη ενέργεια. Από αυτό συνε-
πάγεται ότι είναι αδύνατον νά υπάρχει επαρκής ποικιλομορφία σε όποιοδήποτε
μηχάνημα πού επιτρέπει νά ενεργεί σε όλα τά γεγονότα τής ζωής μέ τόν ίδιο
τρόπο πού ή νοημοσύνη μας μās κάνει νά ενεργούμε»¹. Τό αν αυτή ή θεώρηση
του Descartes σήμερα έχει διαψευσθεί ή όχι, μένει νά αποδειχθεί στην πράξη.

Τό «ιερό δισκοπότηρο» τής έρευνας για τίν Τεχνητή Νοημοσύνη είναι ή
κατασκευή «γενικής» Τεχνητής Νοημοσύνης ή Τεχνητής Νοημοσύνης ανθρω-
πινου επιπέδου, ευρύτερα γνωστής ως «Τεχνητής Γενικής Νοημοσύνης», ή
όποια είναι ευρεία στον μέγιστο βαθμό: ικανή, δηλαδή, νά πετύχει ουσιαστικά
όποιοδήποτε στόχο, συμπεριλαμβανομένης τής μάθησης. Τό στάδιο αυτό, στό

¹ Rene Descartes, *Discours de la Methode Pour bien conduire sa raison, et chercher la verité dans les sciences*, 1637 έλλ. μεταφρ. *Λόγος περί τής Μεθόδου. Για τίν καλή καθοδήγηση του λογικού μας και τίν αναζήτηση τής αλήθειας στις επιστήμες*. Έκδ. Παπαζήση, 1976.

όποιο ή ίδια ή νοημοσύνη μπορεί να σχεδιάσει τό λογισμικό της, φαίνεται να είναι πολύ κοντά εξαιτίας των εξελίξεων των νευρωνικών δικτύων και της μηχανικής μάθησης. Ο όρος «Τεχνητή Γενική Νοημοσύνη» έγινε ευρύτερα γνωστός από τους έρευντές Shane Legg, Mark Gubrud και Ben Goerzel για να δηλώσει πιά συγκεκριμένα τήν τεχνητή γενική νοημοσύνη ανθρώπινου επιπέδου: τήν ικανότητα να πετυχαίνεις όποιοδήποτε στόχο τουλάχιστον τόσο καλά όσο οί άνθρωποι, κάτι πού προϋποθέτει τήν τέλεια μίμηση του ανθρώπινου έγκεφάλου και ίσως και τήν ύπαρξη συνείδησης. Αύτές οί μορφές τεχνητής νοημοσύνης, άν και γνωστές από χαρακτήρες έπιστημονικής φαντασίας και από γνωστούς και μεγάλους άστéρες του Hollywood (όπως τό C3PO του Star Wars), βρίσκονται στήν καλύτερη περίπτωση στό πολύ μακρινό μέλλον.

Η τρίτη και τελευταία έκδοχή τής τεχνητής νοημοσύνης στηρίζεται σέ μία μελλοντολογική θεωρία, ή όποία είναι αρκετά εϋφάνταστη, αλλά όχι και παντελώς άσπίρηκτη: Η Superintelligent AI ή Artificial General Intelligence (AGI), ή υπερευφυής TN ή υπερνοημοσύνη, σχεδιάζει ή ίδια πέραν του λογισμικού και τό υλισμικό της, πού συμβάλλει μ' αυτό τόν τρόπο στή δημιουργία αυτόνομων και αυτοσυνειδητών όντοτήτων, οί όποιες και μπορούν να ανταγωνιστούν τήν ανθρώπινη νοημοσύνη και ίσως άκόμη και να τήν υπερβούν. Η διαδικασία επαναλαμβανόμενης αυτοβελτίωσης όδηγεί σέ ένα κρίσιμο σημείο καμπής, τομής στήν ανθρώπινη ιστορία λόγω τής έκθετικής αύξησης τής τεχνικής προόδου, τό όποιο θά πυροδοτήσει όπωσδήποτε δραματικές άλλαγές. Η πορεία αυτή μπορεί να όνομαστéι «λάβρα νοημοσύνης». Θά συντελούνται μέ ταχύτατους ρυθμούς συνεχόμενες, ριζικές τεχνικές άλλαγές, οί όποιες θά έχουν ως άποτέλεσμα τήν άλλαγή τής ανθρώπινης συνθήκης και θά καθιστούν τό ανθρώπινο είδος ύποδεέστερο ως πρός όλες τές ικανότητες του από τήν νέα αυτή τεχνητή νοημοσύνη². Αυτό τό κρίσιμο σημείο άντιστοιχεί άκριβώς, σέ εκείνο μετά τό όποιο οί μηχανές γίνονται ικανές να σχεδιάζουν οί ίδιες τήν Τεχνητή Νοημοσύνη. Μέχρι να φτάσουμε εκεί, ή άνοδος τής στάθμης θά προκαλείται από τούς άνθρωπους πού βελτιώνουν τές μηχανές: μετά ίσως να προκαλούν τήν άνοδο μηχανές πού θά βελτιώνουν άλλες μηχανές, πιθανώς πολύ πιά γρήγορα απ' όσο θά μπορούσαν οί άνθρωποι. Μολονότι δέν μπορούμε να προβλέψουμε τί άκριβώς θά συμβεί μετά από μία τέτοια έκρηξη νοημοσύνης – γι' αυτό ο Vernor Vinge³ τήν άποκάλεσε [Τεχνολογική] «Μοναδικότητα» (Singularity).

² Τάσις, Θεοφάνης, *Φιλοσοφία τής Ανθρώπινης Αναβάθμισης*, έκδ. Άρμός 2021, σ. 269.

³ Vernor Vinge, *The Coming Technological Singularity*, 1993.

larity)⁴ – ό φυσικός και έρευνητής τής Τεχνητής Νοημοσύνης Steve Omohudro⁵ υποστήριξε σέ ένα βαρυσίμαντο δοκίμιο του, τό 2008, ότι παρ' όλα αυτά μπορούμε νά προβλέψουμε όρισμένες πτυχές από τήν συμπεριφορά τής υπερνοήμονος Τεχνητής Νοημοσύνης, σχεδόν ανεξάρτητα από τούς όποιους στόχους μπορεί νά διαθέτει. Γιά τόν Ray Kurzweil ή μοναδικότητα θά έρθει αναγκαία καί αναπόφευκτα λόγω τής έκθετικής αύξησης τής τεχνικής εξέλιξης σύμφωνα μέ τόν νόμο τής επιταχυνόμενης παραγωγής αποτελεσμάτων⁶. Μετά τήν έλευσί της τό μόνο πού θά άπομένει στήν ανθρωπότητα είναι νά θαυμάζει εύγνωμονούσα τό δημιούργημά της, τό όποιο θά πρέπει νά τήν αντιμετώπιζει μέ σεβασμό καί καλοσύνη.

Η υπερνοημοσύνη είναι τό είδος αυτό μέ τό όποιο πολλοί άρέσκονται νά άσχολούνται, βεβαρημένο μέ μύθους, εικασίες καί τεράστιες παρανοήσεις, πού κυρίως άφορούν τόν χρονικό προσδιορισμό τής εμφάνισής της, αλλά δέν διστάζουν νά όραματίζονται μέχρι καί τήν δυνατότητα μεταφόρτωσης στίς μηχανές του ανθρώπινου έγκεφάλου, τή συνολική μετατροπή τών γνωστών μορφών ζωής, τήν προσδοκία μιάς «Αποκαλυπτικής» άθανασίας ή τήν έξαφάνιση του ανθρώπινου είδους καί τήν έποίκιση όλόκληρου του σύμπαντος από νέες υπερευφυείς όντότητες. Η πιθανολογούμενη έκρηξη νοημοσύνης στηρίζεται στήν υπόθεση ότι ή νοημοσύνη τών μηχανών θά συνεχίσει νά αυξάνεται μέ ρυθμό ανάλογο προς τήν τρέχουσα ίσχύ τους. Γιά νά καθησυχάσουμε αυτές τίς άνησυχίες, αλλά κυρίως γιά τήν ακρίβεια του επιστημονικού λόγου, πρέπει νά έπισημάνουμε ότι ή πλειονότητα τών ειδικών στήν Τεχνητή Νοημοσύνη σήμερα είναι πεπεισμένοι ότι ή εμφάνιση τής υπερνοημοσύνης δέν πρόκειται νά συμβεί, τουλάχιστον στον παρόντα αιώνα, μέ τίς λαμπρές έξαιρέσεις τών υπερανθρωπιστών καί μετανθρωπιστών όπως ό Ray Kurzweil, ό όποιος καί τήν προσδιορίζει χρονικά, περίπου στό 2045, αλλά καί ό Nick Bostrom⁷.

Η συζήτηση γιά τήν έπαυξημένη-μελλοντική τεχνητή νοημοσύνη θέτει ταυτόσημα έρωτήματα καί μιάς παραπέμπει στήν βιοηθική, πού ως κανονιστικός λόγος είναι στον πυρήνα της φιλοσοφική δραστηριότητα καί αυτή θά κρίνει άν οι ύποσχέσεις πού δίνει ή πρόοδος τών βιοϊατρικών έπιστημών καί τής βιο-

⁴ Max Tegmark, *Life 3.0*. σ. 396.

⁵ Steve Omohudro, "Basic AI Drives" and Catastrophic Risks <https://intelligence.org/files/BasicAI-Drives.pdf>

⁶ Kurzweil Ray, *The Singularity is Near: When Humans Transcend Biology*, Penguin Books, 2005, σ. 32.

⁷ Bostrom, Nick, *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

τεχνολογίας είναι πράγματι για τό καλό του ανθρώπου, δηλαδή αν όντως ώφε-
λουν και δέν βλάπτουν –για νά θυμηθοΰμε τόν Ίπποκράτη– όχι μόνο τόν συ-
γκεκριμένο άνθρωπο αλλά τόν Άνθρωπο οικουμενικά και διαχρονικά⁸. Η όμοι-
ότητα έγκειται στό ότι τά άμεσα άποτελέσματα τόσο τής βιοηθικής όσο και
των επαγγελιών τής αυξημένης τεχνητής νοημοσύνης επηρεάζουν τόν πυρήνα
τής ανθρώπινης ύπόστασης.

Άνάλογα μέ τό πώς αντικρίζει κανείς τό μέλλον, ή τεχνητή νοημοσύνη δια-
χωρίζει⁹ τούς ενδιαφερόμενους σέ ψηφιακούς ουτοπιστές, σέ όσους δηλαδή
προσδοκούν διαμέσου της ένα υπέρλαμπρο μέλλον, σέ τεχνοσκεπτικιστές, σέ
όσους άμφιβάλλουν για τίν βραχυπρόθεσμη επιτευξιμότητα των επαγγελο-
μένων ευεργετικών στόχων και σέ λουδίτες, σέ όσους δηλαδή διαφωνούν μέ
τίν εισαγωγή τής τεχνολογίας διαβλέποντας περισσότερους κινδύνους παρά
όφέλη, ενώ, φυσικά, δέν λείπουν και οι ενδιάμεσες κατηγορίες.

Άπέναντι σέ όλες αυτές τίς θεωρήσεις έχει αναπτυχθεί τό λεγόμενο «Κίνη-
μα τής ώφέλιμης Τεχνητής Νοημοσύνης»¹⁰, τό όποιο και πρεσβεΰει πώς τό
θετικό άποτύπωμα τής Τεχνητής Νοημοσύνης στή ζωή μας δέν είναι έγγυημέ-
νο, αλλά κάτι πού πρέπει νά διασφαλιστεί μέ σκληρή δουλειά πάνω στήν
έρευνα για τίν ασφάλειά της.

⁸ Δραγώνα–Μονάχου Μυρτώ, «Βιοηθική και Φιλοσοφία» στό (Συλλογικό), *Θέματα Βιοηθικής*, επι-
στημονική επιμέλεια Σταυρούλα Τσινόρεμα, Κίτσος Λούης, Πανεπιστημιακές Έκδόσεις Κρήτης,
2013, σ. 37.

⁹ Max Tegmark: *Life 3.0*. σ. 54 έπ.

¹⁰ Τό «Κίνημα τής ώφέλιμης ΤΝ» δημιουργήθηκε από τόν καθηγητή Φυσικής στό ΜΠΤ και πρόεδρο
του Ίνστιτούτου για τό Μέλλον τής Ζωής (<https://futureoflife.org>) Max Tegmark βλ. Max Tegmark
Life 3.0 σ. 59-64.