



**SMART
ATTICA** European
Digital
Innovation
Hub

ΕΚΠΑΙΔΕΥΤΙΚΟ ΥΛΙΚΟ
ΓΙΑ ΤΟ ΔΙΑΔΙΚΤΥΑΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ

12.02.2025 ΑΘΗΝΑ



ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ & ΦΙΛΟΣΟΦΙΑ:

Από την ανθρώπινη σκέψη
στη μηχανική μάθηση



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ



ΕΚΠΑΙΔΕΥΤΙΚΟ ΥΛΙΚΟ
ΓΙΑ ΤΟ ΔΙΑΔΙΚΤΥΑΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ

12.02.2025 ΑΘΗΝΑ



ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ & ΦΙΛΟΣΟΦΙΑ:

Από την ανθρώπινη σκέψη
στη μηχανική μάθηση



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ

1. Εισαγωγή στο εκπαιδευτικό πρόγραμμα

Επισκόπηση του θέματος

Η σχέση μεταξύ τεχνητής νοημοσύνης (TN, AI) και φιλοσοφίας αποτελεί ένα από τα πιο ενδιαφέροντα και αμφιλεγόμενα θέματα της σύγχρονης επιστήμης και τεχνολογίας. Στην καρδιά αυτής της συζήτησης βρίσκονται ερωτήματα συναφή με το νου, τη σκέψη, τη συνείδηση. Η φιλοσοφία του νου εξετάζει την φύση της ανθρώπινης σκέψης, προσπαθώντας να κατανοήσει τη σχέση του νου με το σώμα, την αντίληψη και την συνείδηση. Από την άλλη, η TN αφορά στις προσπάθειες να αναπαραχθούν αυτές οι νοητικές διαδικασίες μέσω υπολογιστικών συστημάτων.

Αρχικά, οι φιλοσοφικές συζητήσεις για τη νοημοσύνη και τη συνείδηση επικεντρώνονταν γύρω από την ικανότητα του ανθρώπου να σκέφτεται, να αντιλαμβάνεται και να συνειδητοποιεί την ύπαρξή του. Σήμερα, η ανάπτυξη της TN προσθέτει μια νέα διάσταση σε αυτά τα φιλοσοφικά ερωτήματα διότι η δυνατότητα των μηχανών να "σκέφτονται", να μαθαίνουν από δεδομένα και να λαμβάνουν αποφάσεις θέτει νέες προκλήσεις για την κατανόηση της νοημοσύνης.

Η εξέλιξη της TN έχει οδηγήσει στην ανάπτυξη της μηχανικής μάθησης και των νευρωνικών δικτύων, που είναι ικανά να επεξεργάζονται μεγάλες ποσότητες δεδομένων και να προσαρμόζονται σε νέα δεδομένα χωρίς ανθρώπινη παρέμβαση. Αυτές οι τεχνολογίες έχουν φέρει επανάσταση σε τομείς όπως η υγειονομική περίθαλψη, οι μεταφορές, η χρηματοδότηση και η αναγνώριση εικόνας.

Ωστόσο, η ταχεία εξέλιξη της TN εγείρει και σημαντικά φιλοσοφικά και ηθικά ερωτήματα: *Μπορεί μια μηχανή να αναπτύξει συνείδηση ή να έχει ηθικές αξίες; Ποιες είναι οι επιπτώσεις της TN στην ανθρώπινη νοημοσύνη και ηθική; Μπορεί η μηχανική μάθηση να αναπαράγει την ανθρώπινη σκέψη, ή πρόκειται απλώς για έναν μηχανισμό που εκτελεί προκαθορισμένες εντολές χωρίς πραγματική κατανόηση;*

Ένα ακόμη βασικό ζήτημα που εγείρεται σχετίζεται με την ηθική της TN. Καθώς οι μηχανές αποκτούν ολοένα και μεγαλύτερη αυτονομία, το ερώτημα για την υπευθυνότητα και την ηθική συμπεριφορά των συστημάτων TN γίνεται ολοένα πιο σημαντικό. Για παράδειγμα, *μπορεί ένα μηχανήμα να λάβει ηθικές αποφάσεις σε καταστάσεις που απαιτούν ηθική κρίση, όπως σε στρατιωτικές εφαρμογές και, κατά δεύτερον, ποιος είναι υπεύθυνος όταν μια μηχανή προκαλεί ζημιά ή παραβιάζει ηθικούς κανόνες;*

Η φιλοσοφία της TN είναι, επίσης, στενά συνδεδεμένη με την αναζήτηση της αλήθειας σχετικά με την ανθρώπινη φύση. Οι υποστηρικτές της TN προτείνουν ότι, καθώς οι μηχανές προχωρούν στην ικανότητά τους να προσομοιώνουν ανθρώπινες διαδικασίες σκέψης, μπορούν να προσφέρουν νέες κατανοήσεις για τη νοημοσύνη, την ηθική και την ανθρώπινη συνείδηση. Ωστόσο, άλλοι υποστηρίζουν ότι η TN είναι περιορισμένη στο να μιμείται μόνο την ανθρώπινη σκέψη και δεν μπορεί να κατανοήσει την πλήρη εμπειρία ή τη συνείδηση όπως οι άνθρωποι.

Προς τούτο, αυτό το εκπαιδευτικό υλικό αποσκοπεί στην παρουσίαση των αλληλεπιδράσεων μεταξύ της φιλοσοφίας της ανθρώπινης σκέψης και της TN, προσδιορίζοντας τα βασικά

ερωτήματα που προκύπτουν και τις μελλοντικές προκλήσεις που σχετίζονται με την ανάπτυξη και τη χρήση αυτών των τεχνολογιών.

Μαθησιακοί στόχοι

Με την ολοκλήρωση του εκπαιδευτικού προγράμματος, οι συμμετέχοντες θα είναι σε θέση:

- Να γνωρίζουν τις βασικές φιλοσοφικές θεωρίες για τη φύση της ανθρώπινης σκέψης, του νου και της συνείδησης.
- Να αντιλαμβάνονται την ανάπτυξη της ΤΝ, από τα πρώτα βήματα της θεωρητικής έρευνας μέχρι τις σύγχρονες εφαρμογές στον τομέα της μηχανικής μάθησης.
- Να αναλύουν τις ομοιότητες και διαφορές μεταξύ της ανθρώπινης νοημοσύνης και της ΤΝ, εξετάζοντας τα όρια και τις δυνατότητες της δεύτερης.
- Να εντοπίζουν τα φιλοσοφικά και ηθικά διλήμματα που σχετίζονται με την ανάπτυξη και χρήση της ΤΝ, όπως η ηθική της ΤΝ και οι κίνδυνοι της αυτονομίας των μηχανών.
- Να αναγνωρίζουν τις μελλοντικές προοπτικές της ΤΝ, τις πιθανές επιπτώσεις της στην κοινωνία και την ανθρώπινη νοημοσύνη, καθώς και τις φιλοσοφικές συνέπειες της ανάπτυξης των τεχνολογιών αυτών.

2.Εισαγωγή στη φιλοσοφία της τεχνητής νοημοσύνης

Τεχνητή νοημοσύνη & Φιλοσοφία

Η ΤΝ αναφέρεται στη δυνατότητα των μηχανών να εκτελούν εργασίες που απαιτούν ανθρώπινη νοημοσύνη, όπως η αναγνώριση εικόνας, η γλώσσα, η μάθηση και η λήψη αποφάσεων. Η ΤΝ στοχεύει στη δημιουργία υπολογιστικών συστημάτων που προσομοιώνουν τον ανθρώπινο νου και προσπαθούν να αναπαραστήσουν τις λειτουργίες του.

Η φιλοσοφία της ΤΝ, από την άλλη πλευρά, εξετάζει τα θεμελιώδη ερωτήματα που προκύπτουν από την ανάπτυξη αυτών των συστημάτων. Τα βασικά ζητήματα περιλαμβάνουν:

- Η φύση της νοημοσύνης: Τι είναι η νοημοσύνη; Είναι κάτι που μπορούν να κατανοήσουν οι μηχανές ή απλώς οι άνθρωποι;
- Η συνείδηση: Μπορεί μια μηχανή να έχει συνείδηση όπως οι άνθρωποι;
- Η ηθική και υπευθυνότητα: Εάν μια μηχανή μπορεί να μάθει και να αποφασίζει, ποιος είναι υπεύθυνος για τις πράξεις της;

Η σχέση της ΤΝ με τη φιλοσοφία είναι κρίσιμη, διότι οι φιλοσοφικές θεωρίες του νου και της συνείδησης έχουν άμεσες συνέπειες για τον τρόπο που αναπτύσσουμε και κατανοούμε τις μηχανές που προσπαθούν να προσομοιώσουν τις ανθρώπινες ικανότητες.

Από τη φιλοσοφία της νοημοσύνης στην τεχνητή νοημοσύνη

Η ιστορία της ΤΝ μπορεί να εντοπιστεί στους φιλοσοφικούς προβληματισμούς του αρχαίου κόσμου. Φιλόσοφοι όπως ο Αριστοτέλης προσδιόρισαν «την ανθρώπινη σκέψη ως μια ικανότητα λογικής ανάλυσης και κατηγοριοποίησης», ενώ άλλοι, όπως ο Ρενέ Ντεκάρτ θεωρούσαν «τον άνθρωπο ως το μόνο ον που κατέχει πλήρη συνείδηση και λογική σκέψη».



Εικόνα: Άλαν Τούρινγκ
(Wikipedia)

Παρά ταύτα, οι πρώτες σύγχρονες φιλοσοφικές έννοιες της νοημοσύνης εμφανίστηκαν με τη σκέψη των φιλοσόφων του 20ού αιώνα, όπως του Άγγλου μαθηματικού **Άλαν Τούρινγκ** (Alan Matheson Turing, 23 Ιουνίου 1912 – 7 Ιουνίου 1954). Στην εργασία του «Η Υπολογιστική Μηχανή και η Νοημοσύνη» (1950), ο Τούρινγκ πρότεινε το γνωστό «Τεστ Τούρινγκ» για να εξετάσει εάν οι μηχανές μπορούν να μιμηθούν την ανθρώπινη νοημοσύνη.

Από τις αρχές του 20ού αιώνα, οι επιστήμονες και οι φιλόσοφοι άρχισαν να αναπτύσσουν θεωρίες που προσπαθούσαν να εξηγήσουν πώς μπορούσαν οι μηχανές να μάθουν και να σκέφτονται. Η θεωρία των νευρωνικών δικτύων και οι πρώτες εφαρμογές υπολογιστών το 1956 (αρχή της ΤΝ) σηματοδότησαν την εποχή της ΤΝ. Οι επιστήμονες άρχισαν να κατανοούν ότι, εάν μπορούσαν να αναπαράγουν τον τρόπο που λειτουργεί ο ανθρώπινος εγκέφαλος, οι μηχανές ίσως μπορούσαν να εκτελούν «νοητικές» εργασίες.

Τα μαθηματικά μοντέλα, η μηχανική μάθηση και τα νευρωνικά δίκτυα αναπτύχθηκαν, επιτρέποντας στις μηχανές να μάθουν από δεδομένα και να βελτιώνονται με την εμπειρία, όπως ακριβώς κάνει και ο άνθρωπος.

Η φιλοσοφία της ΤΝ δεν περιορίζεται στην τεχνική ανάπτυξή της, αλλά αναδεικνύει τα εξής θεμελιώδη φιλοσοφικά ερωτήματα:

Τι είναι η συνείδηση; Η συνείδηση είναι μία από τις πιο περίπλοκες έννοιες της φιλοσοφίας. Οι φιλόσοφοι έχουν προσπαθήσει να την κατανοήσουν για αιώνες. Μπορούν οι μηχανές να αποκτήσουν συνείδηση; Αν μια μηχανή μπορεί να μιμείται τις ανθρώπινες αντιδράσεις, μπορεί να είναι «συνειδητή»; Αυτό το ερώτημα τίθεται συνεχώς στο πλαίσιο της ΤΝ και είναι κεντρικό στην ανάπτυξη ηθικών και φιλοσοφικών θεωριών γύρω από την ΤΝ.

Μπορούν οι μηχανές να σκέφτονται; Η ιδέα ότι μια μηχανή μπορεί να σκέφτεται αμφισβητεί την ίδια την έννοια της σκέψης. Μπορεί μια υπολογιστική μηχανή, μέσω αλγορίθμων και δεδομένων, να λάβει αποφάσεις ή να επιλύσει προβλήματα όπως οι άνθρωποι; Η σκέψη είναι συνυφασμένη με τη συνείδηση, και πολλοί φιλόσοφοι υποστηρίζουν ότι οι μηχανές μπορεί να εκτελούν εργασίες που μοιάζουν με σκέψη, αλλά δεν έχουν την πραγματική «κατανόηση» που έχει ο ανθρώπινος νους.

Η ηθική και η νοημοσύνη των μηχανών: Το ερώτημα για τη νοημοσύνη των μηχανών συνδέεται και με την ηθική της ΤΝ. Αν μια μηχανή μπορεί να «σκεφτεί» και να αποφασίσει, ποιος είναι υπεύθυνος για τις πράξεις της; Πώς μπορούμε να διασφαλίσουμε ότι οι μηχανές θα χρησιμοποιούνται για το καλό της ανθρωπότητας, χωρίς να προκαλούν βλάβη ή να αναπαράγουν κοινωνικές αδικίες;

Ο ρόλος του νου στην ΤΝ: Τέλος, είναι σημαντικό να ξέρουμε το βαθμό στον οποίο η ΤΝ αναπαράγει τον ανθρώπινο νου. Οι μηχανές, ακόμα και αν επιτυγχάνουν να μιμηθούν κάποιες νοητικές διαδικασίες, μπορεί να μην είναι σε θέση να βιώσουν τον κόσμο ή να έχουν εμπειρίες όπως οι άνθρωποι. Η σύγκριση του ανθρώπινου νου με την ΤΝ θέτει τα όρια της τεχνολογίας και την ανάγκη για κριτική σκέψη.

3. Ανθρώπινη σκέψη & Φιλοσοφικές θεωρίες

Φιλοσοφία του νου: Διαφορετικές θεωρίες για τη φύση της ανθρώπινης σκέψης

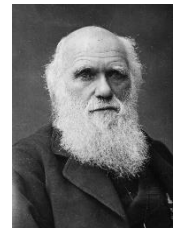
Η ανθρώπινη σκέψη και η συνείδηση υπήρξαν κεντρικά θέματα στη φιλοσοφία του νου, με διαφορετικές θεωρίες να επιχειρούν να εξηγήσουν τη φύση και τη λειτουργία τους. Από τη διπλότητα νου και σώματος του **Ρενέ Ντεκάρτ** έως τις εξελικτικές και ψυχολογικές προσεγγίσεις της συνείδησης από τον **Κάρολο Δαρβίνο** και τον **Ρόμπερτ Σόλομον**, κάθε προσέγγιση φωτίζει διαφορετικές πτυχές αυτού του πολύπλοκου ζητήματος.



Εικόνα: Ρενέ Ντεκάρτ (Wikipedia)

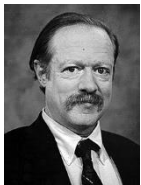
Ο Ρενέ Ντεκάρτ διατύπωσε τη θεωρία της διπλότητας νου και σώματος, σύμφωνα με την οποία ο νους και το σώμα είναι δύο διακριτές και ανεξάρτητες ουσίες. Ο νους, ως άυλη ουσία, είναι υπεύθυνος για τη σκέψη, τη συνείδηση και τις πνευματικές λειτουργίες, ενώ το σώμα, ως υλική ουσία, ακολουθεί τους νόμους της φυσικής. Παρ' όλο που αυτές οι δύο ουσίες είναι διαφορετικές, ο Ντεκάρτ πίστευε ότι αλληλεπιδρούν μέσω της επίφυσης, ενός μικρού αδένου στον εγκέφαλο. Η θεωρία του είχε βαθιά επίδραση στη φιλοσοφική και επιστημονική σκέψη, ενώ παράλληλα προκάλεσε συζητήσεις σχετικά με το πώς μια άυλη και μια υλική ουσία μπορούν να αλληλεπιδρούν.

Από την άλλη πλευρά, η θεωρία της συνείδησης, όπως διαμορφώθηκε από τον Κάρολο Δαρβίνο και τον Ρόμπερτ Σόλομον, εστιάζει στη φυσική και ψυχολογική βάση της ανθρώπινης εμπειρίας. Ο Δαρβίνος, μέσα από τη θεωρία της εξέλιξης, συνέδεσε τη συνείδηση με τη φυσική επιλογή. Υποστήριξε ότι η συνείδηση αναπτύχθηκε εξελικτικά ως μέσο προσαρμογής, επιτρέποντας στους ανθρώπους να μαθαίνουν, να αλληλεπιδρούν κοινωνικά και να επιβιώνουν σε σύνθετα περιβάλλοντα. Επιπλέον, ο Δαρβίνος πρότεινε ότι οι πνευματικές και συναισθηματικές λειτουργίες δεν είναι μοναδικές στον άνθρωπο, αλλά εμφανίζονται και σε άλλα ζώα, σε διαφορετικούς βαθμούς.



Εικόνα: Κάρολος Δαρβίνος (Wikipedia)

Ο Ρόμπερτ Σόλομον, στο πλαίσιο της φιλοσοφίας των συναισθημάτων, προσέγγισε τη συνείδηση ως εργαλείο διαχείρισης των συναισθημάτων και λήψης αποφάσεων. Εξήγησε ότι η συνείδηση συνδέεται στενά με την ελευθερία της βούλησης, επιτρέποντας στον άνθρωπο να ερμηνεύει, να επιλέγει και να ανταποκρίνεται στις καταστάσεις πέρα από τα απλά βιολογικά ένστικτα. Για τον Σόλομον, η συνείδηση είναι το θεμέλιο της προσωπικής ευθύνης και της ηθικής δράσης.



Εικόνα: Ρόμπερτ Σόλομον (Wikipedia)

Αυτές οι θεωρίες αντικατοπτρίζουν τις διαφορετικές προσεγγίσεις στη μελέτη της ανθρώπινης σκέψης και συνείδησης, συνδυάζοντας φιλοσοφικές, βιολογικές και ψυχολογικές οπτικές για να απαντήσουν στα μεγάλα ερωτήματα για τη φύση του νου.

Η ηθική της συνείδησης: Η σχέση ανθρώπινης συνείδησης και ηθικής συμπεριφοράς

Η συνείδηση αποτελεί τον πυρήνα της ανθρώπινης ικανότητας να αντιλαμβάνεται τον εαυτό του και τον κόσμο, ενώ παράλληλα διαδραματίζει κεντρικό ρόλο στη διαμόρφωση της ηθικής συμπεριφοράς. Μέσα από τη συνείδηση, το άτομο αναγνωρίζει τις συνέπειες των πράξεών του και αξιολογεί αν αυτές είναι σωστές ή λανθασμένες. Οι ηθικές αρχές δεν είναι απλώς αποτέλεσμα κοινωνικής επιβολής ή εξωτερικών κανόνων, αλλά πηγάζουν από την εσωτερική συνείδηση που καλλιεργείται μέσω της εμπειρίας, της σκέψης και της ενσυναίσθησης.



Εικόνα:
Ιμμάνουελ Κάντ
(Wikipedia)

Η ηθική της συνείδησης συνδέεται άρρηκτα με την ικανότητα αυτογνωσίας και αναστοχασμού. Ο άνθρωπος, ως ηθικό ον, όχι μόνο κρίνει τις πράξεις του, αλλά συχνά οδηγείται σε εσωτερική σύγκρουση όταν οι επιλογές του αντιβαίνουν στις προσωπικές του αξίες. Η συνείδηση, επομένως, λειτουργεί ως μια εσωτερική φωνή που κατευθύνει τη συμπεριφορά, ενθαρρύνοντας την υπευθυνότητα και την αυτονομία. Επιπλέον, η συνείδηση καθοδηγείται από το αίσθημα του "καθήκοντος" που περιέγραψε ο Ιμμάνουελ Κάντ, καθιστώντας το άτομο ικανό να επιλέγει το καλό ανεξάρτητα από εξωτερικές πιέσεις.

Ηθική συμπεριφορά χωρίς συνείδηση καθίσταται απρόσωπη και μηχανιστική, ενώ η απουσία ηθικής από τη συνείδηση οδηγεί σε εγωκεντρικές ή αντικοινωνικές τάσεις. Αυτή η σχέση υπογραμμίζει τη βαθιά αλληλεξάρτηση των δύο εννοιών, καθιστώντας τη συνείδηση την κινητήρια δύναμη για μια πιο δίκαιη και υπεύθυνη κοινωνία.

Σκέψη & Συναισθήματα: Η αλληλεπίδραση σκέψης και συναισθημάτων στην ανθρώπινη συμπεριφορά

Η ανθρώπινη συμπεριφορά διαμορφώνεται μέσα από την αλληλεπίδραση σκέψης και συναισθημάτων, δύο φαινομενικά διαφορετικών αλλά ουσιαστικά αλληλένδετων στοιχείων της ανθρώπινης εμπειρίας. Η σκέψη παρέχει τη λογική επεξεργασία των δεδομένων, ενώ τα συναισθήματα προσδίδουν νόημα και αξία στις καταστάσεις, οδηγώντας σε πιο ολοκληρωμένες αποφάσεις και πράξεις.

Η σκέψη συχνά λειτουργεί ως ο μηχανισμός που οργανώνει τις πληροφορίες, αξιολογεί επιλογές και προβλέπει αποτελέσματα. Τα συναισθήματα, όμως, εμπλουτίζουν αυτή τη διαδικασία με εμπειρικά και ηθικά κριτήρια. Για παράδειγμα, μια απόφαση που βασίζεται αποκλειστικά στη λογική μπορεί να φαίνεται σωστή, αλλά αν δεν λαμβάνει υπόψη τις συναισθηματικές επιπτώσεις, μπορεί να είναι άκαμπτη ή απάνθρωπη.

Αντίστροφα, τα συναισθήματα συχνά καθοδηγούν τη σκέψη. Ο φόβος, η χαρά ή η λύπη μπορούν να επηρεάσουν την ερμηνεία μιας κατάστασης και να επηρεάσουν τις επιλογές μας. Ωστόσο, αν τα συναισθήματα κυριαρχούν χωρίς την εξισορρόπηση της σκέψης, υπάρχει κίνδυνος παρορμητικότητας ή κακής κρίσης. Η ιδανική ανθρώπινη συμπεριφορά προκύπτει όταν σκέψη και συναισθήματα βρίσκονται σε αρμονία, δημιουργώντας ισορροπημένες και ηθικές επιλογές.

Η αλληλεπίδραση αυτών των δύο παραμέτρων ενισχύεται μέσα από τη συνείδηση, η οποία δρα ως γέφυρα ανάμεσά τους. Καθώς ο άνθρωπος συνειδητοποιεί τις σκέψεις και τα συναισθήματά του, μπορεί να κατανοήσει καλύτερα τον εαυτό του και τους άλλους, οδηγώντας σε συμπεριφορές που είναι όχι μόνο αποτελεσματικές αλλά και ανθρώπινες.

4. Φιλοσοφικές σκέψεις γύρω από την τεχνητή νοημοσύνη

Ηθική & Υπεύθυνη τεχνητή νοημοσύνη

Η ταχεία ανάπτυξη της τεχνητής νοημοσύνης (TN, AI) έχει φέρει στο προσκήνιο ερωτήματα σχετικά με την ηθική των μηχανών και τον ρόλο τους στις ανθρώπινες κοινωνίες.

Μπορούν οι μηχανές να είναι ηθικές; Η ικανότητα των μηχανών να ενεργούν ηθικά εξαρτάται από το πώς ορίζουμε την ηθική. Αν η ηθική συνδέεται με την αυτονομία, την πρόθεση και την ενσυναίσθηση, τότε οι μηχανές, ως προϊόντα ανθρώπινου προγραμματισμού, δεν μπορούν να θεωρηθούν ηθικές. Ωστόσο, μπορούν να ακολουθούν ηθικές αρχές που έχουν ενσωματωθεί στους αλγόριθμους τους. Για παράδειγμα, μια μηχανή μπορεί να προγραμματιστεί να αποφεύγει πράξεις που βλάπτουν τον άνθρωπο, αλλά η συμμόρφωση αυτή δεν υποδηλώνει συνειδητή ηθική δράση.

Η βασική διαφορά μεταξύ ανθρώπων και μηχανών έγκειται στη συνείδηση και την πρόθεση. Ενώ οι μηχανές μπορούν να παίρνουν αποφάσεις βάσει κανόνων, στερούνται της συναισθηματικής νοημοσύνης και της δυνατότητας να αξιολογούν τις πράξεις τους με βάση την εσωτερική ηθική κρίση. Επομένως, οι μηχανές δεν είναι ηθικές με τον ίδιο τρόπο όπως οι άνθρωποι, αλλά μπορούν να θεωρηθούν "ηθικά εργαλεία" που υλοποιούν τις προθέσεις των δημιουργών τους.

Ποιες είναι οι προκλήσεις στην ηθική της τεχνητής νοημοσύνης; Η χρήση της TN ενέχει πολυάριθμες ηθικές προκλήσεις, με βασικότερη την προκατάληψη στους αλγόριθμους. Οι αλγόριθμοι της TN συχνά βασίζονται σε δεδομένα που περιέχουν υπάρχουσες κοινωνικές ανισότητες ή στερεότυπα, γεγονός που οδηγεί σε άνιση μεταχείριση. Για παράδειγμα, οι αλγόριθμοι που χρησιμοποιούνται για την πρόσληψη εργαζομένων μπορεί να ευνοούν συγκεκριμένες ομάδες με βάση το φύλο, τη φυλή ή την κοινωνική τάξη, εάν εκπαιδευτούν με προκατειλημμένα δεδομένα.

Επιπλέον, η έλλειψη διαφάνειας στις διαδικασίες λήψης αποφάσεων των μηχανών προκαλεί δυσπιστία. Τα λεγόμενα «μαύρα κουτιά» της TN, δηλαδή συστήματα όπου η λογική των αποφάσεων δεν είναι κατανοητή, καθιστούν δύσκολη την αναγνώριση λαθών ή αδικιών. Άλλες προκλήσεις περιλαμβάνουν την προστασία της ιδιωτικότητας, την ασφάλεια των δεδομένων και την πιθανή ανεξέλεγκτη χρήση της TN για καταστροφικούς σκοπούς.

Η θεωρία της ηθικής της μηχανής: Μία από τις πιο ενδιαφέρουσες προσεγγίσεις για την ηθική της TN είναι η θεωρία της ηθικής της μηχανής, η οποία επιχειρεί να διαμορφώσει ένα πλαίσιο για τη λήψη ηθικών αποφάσεων από μηχανές. Ένα χαρακτηριστικό παράδειγμα είναι το ηθικό πρόβλημα του τρόλεϊ (*trolley problem*), το οποίο αφορά την επιλογή ανάμεσα σε δύο καταστάσεις όπου καμία επιλογή δεν είναι ιδανική. Για παράδειγμα, σε ένα σενάριο όπου ένα αυτόνομο όχημα πρέπει να επιλέξει μεταξύ του να προκαλέσει ατύχημα για να προστατεύσει τον οδηγό του ή να βλάψει άλλους στο δρόμο, ποια είναι η σωστή απόφαση; Το πρόβλημα αυτό επιδεικνύει τις

δυσκολίες της ενσωμάτωσης ηθικών αρχών στους αλγορίθμους, καθώς απαιτεί σαφείς κανόνες για καταστάσεις που συχνά εμπλέκουν ηθικές διαισθήσεις και υποκειμενικές κρίσεις.

Η θεωρία της ηθικής της μηχανής προσπαθεί να ενσωματώσει αρχές όπως ο ωφελισμός (η επιλογή που προκαλεί τη μέγιστη ευτυχία) ή η δεοντολογία (η τήρηση κανόνων ανεξάρτητα από το αποτέλεσμα). Παρά τις προσπάθειες, όμως, παραμένει δύσκολο να κωδικοποιηθούν ηθικές έννοιες που συχνά είναι πολύπλοκες και εξαρτώνται από το πλαίσιο.

Συμπερασματικά, εκτιμάται πως η ηθική της ΤΝ απαιτεί συνεχή εξέταση και προσαρμογή, με τη συνεργασία ειδικών από τη φιλοσοφία, την πληροφορική και την κοινωνιολογία· μόνο έτσι μπορεί να εξασφαλιστεί ότι η ανάπτυξη της ΤΝ θα ωφελεί την κοινωνία χωρίς να θυσιάζει θεμελιώδεις ηθικές αξίες.

Συνείδηση & Τεχνητή νοημοσύνη

Ένα από τα πιο θεμελιώδη ερωτήματα στη συζήτηση για την ΤΝ είναι εάν μια μηχανή μπορεί να αποκτήσει συνείδηση. Η συνείδηση, όπως είναι αντιληπτό, είναι η ικανότητα ενός όντος να έχει αυτογνωσία, να βιώνει υποκειμενικές εμπειρίες και να αισθάνεται συναισθήματα. Οι μηχανές, ως προϊόντα προγραμματισμού, εκτελούν εντολές βάσει δεδομένων και αλγορίθμων, αλλά στερούνται της υποκειμενικότητας που χαρακτηρίζει τη συνείδηση.

Παρά τις τεχνολογικές προόδους, δεν υπάρχει σαφής επιστημονική ή φιλοσοφική συναίνεση σχετικά με το εάν η συνείδηση μπορεί να δημιουργηθεί σε μηχανές. Ορισμένοι επιστήμονες υποστηρίζουν ότι, εάν γίνει κατανοητή πλήρως η λειτουργία του ανθρώπινου εγκεφάλου, θα μπορούσαν να προσομοιώσουν τη συνείδηση σε μια μηχανή. Ωστόσο, άλλοι επισημαίνουν ότι η συνείδηση ίσως είναι ένα μοναδικό χαρακτηριστικό της βιολογίας, που δεν μπορεί να αναπαραχθεί τεχνητά.

Οι φιλοσοφικές και ηθικές συνέπειες της πιθανότητας μιας συνειδητής μηχανής είναι τεράστιες. *Αν μια μηχανή αποκτήσει συνείδηση, θα πρέπει να αναγνωριστεί ως ηθικό υποκείμενο με δικαιώματα; Θα είναι υπεύθυνη για τις πράξεις της; Επιπλέον, αν η δημιουργία της συνείδησης συνδέεται με τον πόνο ή τη δυστυχία, θα ήταν ηθικά αποδεκτή μια τέτοια προσπάθεια;* Αυτά τα ερωτήματα ωθούν στην εξέταση της φύσης της συνείδησης, όχι μόνο ως επιστημονικό ζήτημα, αλλά και ως ένα βαθιά ηθικό και κοινωνικό φαινόμενο.

Ανθρώπινη & Μηχανική Νοημοσύνη: Η ανάπτυξη της ΤΝ έχει προκαλέσει ανησυχίες και προβληματισμούς σχετικά με την πιθανότητα οι μηχανές να ξεπεράσουν την ανθρώπινη νοημοσύνη ή να αντικαταστήσουν τον άνθρωπο σε διάφορες δραστηριότητες. Η ανθρώπινη νοημοσύνη χαρακτηρίζεται από δημιουργικότητα, ενσυναίσθηση και την ικανότητα να κατανοούμε τον κόσμο μέσα από εμπειρίες. Αντίθετα, η μηχανική νοημοσύνη βασίζεται στην επεξεργασία δεδομένων και την επίλυση προβλημάτων βάσει αλγορίθμων.

Παρ' όλο που οι μηχανές έχουν ήδη ξεπεράσει τον άνθρωπο σε συγκεκριμένους τομείς, όπως η ταχύτητα υπολογισμών ή η ανάλυση μεγάλων όγκων δεδομένων, δεν μπορούν να μιμηθούν πλήρως τη σύνθετη φύση της ανθρώπινης σκέψης. Ωστόσο, η **θεωρία της «τελικής υπερνοημοσύνης» (superintelligence)** εγείρει σημαντικά ερωτήματα. Σύμφωνα με αυτή, η ΤΝ θα μπορούσε να φτάσει σε ένα σημείο όπου θα είναι σε θέση να βελτιώνει μόνη της τον εαυτό

της, οδηγώντας σε μια εκθετική αύξηση της νοημοσύνης της που θα ξεπεράσει κατά πολύ τις ανθρώπινες δυνατότητες.

Η πιθανότητα αυτή προκαλεί ανησυχίες για το εάν οι μηχανές θα μπορέσουν να ελέγξουν πλήρως την κοινωνία ή αν θα γίνουν απειλή για τον άνθρωπο. Ορισμένοι θεωρούν ότι η υπερνοημοσύνη θα μπορούσε να λύσει παγκόσμια προβλήματα, όπως η κλιματική αλλαγή ή οι ασθένειες, ενώ άλλοι προειδοποιούν για τους κινδύνους της, όπως η απώλεια του ανθρώπινου ελέγχου ή η δημιουργία ανισοτήτων.

Η συζήτηση γύρω από την ανθρώπινη και μηχανική νοημοσύνη καλεί την επανεξέταση του τι σημαίνει να είσαι άνθρωπος και πώς διαχειριζόμαστε τις τεχνολογίες του μέλλοντος. Σε κάθε περίπτωση, η ανάπτυξη της ΤΝ πρέπει να συνοδεύεται από αυστηρούς ηθικούς κανόνες, ώστε να διασφαλιστεί ότι θα ωφελήσει την ανθρωπότητα, χωρίς να υπονομεύσει τις θεμελιώδεις αξίες της.

5. Συγκρίσεις, Συμπεράσματα & Συζήτηση

Σύγκριση ανθρώπινης σκέψης και τεχνητής νοημοσύνης

Η ανθρώπινη σκέψη και η ΤΝ αποτελούν δύο διαφορετικά συστήματα επεξεργασίας πληροφοριών, που παρουσιάζουν κοινά χαρακτηριστικά αλλά και θεμελιώδεις διαφορές. Από τη μία πλευρά, και οι δύο βασίζονται στη διαδικασία συλλογής, ανάλυσης και αξιοποίησης δεδομένων, με στόχο την εξαγωγή συμπερασμάτων και την επίλυση προβλημάτων. Ωστόσο, η πηγή της γνώσης, η μέθοδος επεξεργασίας και η ευελιξία τους είναι οι παράγοντες που τα διαφοροποιούν.

Ο άνθρωπος αντλεί τη γνώση του μέσω εμπειριών, διαισθητικής σκέψης και συναισθηματικής κατανόησης. Η ανθρώπινη σκέψη είναι πολυδιάστατη, καθώς συνδυάζει πληροφορίες από διαφορετικά πλαίσια και τις επεξεργάζεται όχι μόνο με βάση τη λογική, αλλά και μέσω της υποκειμενικής εμπειρίας και της φαντασίας. Επιπλέον, η ανθρώπινη σκέψη διαθέτει τη μοναδική ικανότητα να δημιουργεί εντελώς νέες ιδέες και να λειτουργεί έξω από προκαθορισμένα πλαίσια, επιδεικνύοντας δημιουργικότητα.

Αντίθετα, η ΤΝ στηρίζεται αποκλειστικά σε δεδομένα που της παρέχονται και σε προγραμματισμένους αλγόριθμους. Αν και οι τεχνικές μηχανικής μάθησης έχουν βελτιώσει σημαντικά την ικανότητα της ΤΝ να αναγνωρίζει μοτίβα και να εξάγει συμπεράσματα, οι αλγόριθμοί της λειτουργούν μέσα σε αυστηρά καθορισμένα όρια· δεν μπορεί να δημιουργήσει ή να αντιληφθεί κάτι που δεν περιλαμβάνεται στη βάση δεδομένων της ή στο προγραμματιστικό της υπόβαθρο.

Η προσαρμοστικότητα αποτελεί ακόμη ένα στοιχείο διαφοροποίησης. Ο άνθρωπος μπορεί να αντιδράσει αυθόρμητα σε πρωτόγνωρες καταστάσεις και να αναπτύξει ενστικτώδεις ή δημιουργικές λύσεις. Για παράδειγμα, ένας άνθρωπος που βρίσκεται μπροστά σε μια άγνωστη πρόκληση μπορεί να χρησιμοποιήσει τη φαντασία ή τη διαίσθησή του για να επινοήσει μια λύση που δεν έχει εφαρμοστεί ποτέ ξανά. Η ΤΝ, από την άλλη πλευρά, μπορεί να ανταποκριθεί μόνο εάν η πρόκληση βρίσκεται εντός του πεδίου δεδομένων που έχει εκπαιδευτεί να διαχειρίζεται. Εάν δεν υπάρχουν προηγούμενα δεδομένα ή αν η κατάσταση ξεφεύγει από τα προκαθορισμένα πλαίσια, η ΤΝ δεν μπορεί να προσαρμοστεί.

Επιπλέον, η ανθρώπινη σκέψη περιλαμβάνει συναισθηματική διάσταση, που επηρεάζει την αντίληψη και τις αποφάσεις. Ο άνθρωπος μπορεί να αξιολογήσει μια κατάσταση όχι μόνο με βάση τη λογική αλλά και με βάση το ηθικό και συναισθηματικό πλαίσιο. Για παράδειγμα, η συμπόνια, η ενσυναίσθηση και η διαίσθηση μπορούν να οδηγήσουν σε αποφάσεις που δεν βασίζονται αποκλειστικά σε δεδομένα αλλά σε αξίες και συναισθήματα. Η ΤΝ, από την άλλη πλευρά, στερείται αυτής της συναισθηματικής νοημοσύνης. Αν και μπορεί να αναγνωρίζει συναισθηματικά μοτίβα μέσω αλγορίθμων (π.χ., ανάλυση συναισθημάτων από κείμενα ή εικόνες), δεν μπορεί να τα «αισθανθεί» ή να τα κατανοήσει όπως ένας άνθρωπος.

Τέλος, η ανθρώπινη σκέψη χαρακτηρίζεται από την ικανότητα αυτοκριτικής και αυτογνωσίας. Οι άνθρωποι μπορούν να αξιολογούν τις πράξεις τους, να σκέφτονται για τη δική τους ύπαρξη και να λαμβάνουν αποφάσεις με βάση τις αξίες τους. Η ΤΝ, ωστόσο, στερείται αυτής της ικανότητας· λειτουργεί αποκλειστικά με βάση τους προγραμματιστικούς κανόνες και δεν έχει καμία αίσθηση του «εαυτού» της ή ηθική κρίση.

Συμπερασματικά, ενώ η ανθρώπινη σκέψη και η TN έχουν κοινές βάσεις σε ό,τι αφορά τη διαχείριση δεδομένων, οι θεμελιώδεις διαφορές τους εντοπίζονται στην ευελιξία, τη δημιουργικότητα, την συναισθηματική διάσταση και την ικανότητα αυτογνωσίας. Η ανθρώπινη σκέψη παραμένει ένα σύνθετο και αξεπέραστο φαινόμενο που η TN, αν και χρήσιμη και αποτελεσματική, δεν μπορεί να αντικαταστήσει πλήρως. Βέβαια, αξίζει να σημειωθεί ότι οι πιο ουσιαστικές διαφορές εντοπίζονται στην αντίληψη, το συναίσθημα και τη συνείδηση.

- **Αντίληψη**

Η ανθρώπινη αντίληψη είναι βαθιά συνδεδεμένη με τις αισθήσεις, τις εμπειρίες και τα συναισθήματα. Ο άνθρωπος αντιλαμβάνεται τον κόσμο μέσω των πέντε αισθήσεων (όραση, ακοή, αφή, γεύση, όσφρηση) σε συνδυασμό με τη βιωματική γνώση και την ικανότητα να συνδέει τα ερεθίσματα με προσωπικές εμπειρίες. Αυτή η πολυδιάστατη προσέγγιση επιτρέπει την κατανόηση σύνθετων καταστάσεων, ακόμα και όταν λείπουν σαφή δεδομένα. Επιπλέον, η ανθρώπινη αντίληψη είναι υποκειμενική: διαφορετικοί άνθρωποι μπορεί να αντιληφθούν την ίδια κατάσταση με διαφορετικό τρόπο, ανάλογα με τις εμπειρίες, τις αξίες και τα συναισθήματά τους.

Αντίθετα, η αντίληψη της TN περιορίζεται σε αντικειμενικά δεδομένα που συλλέγονται μέσω αισθητήρων, καμερών και άλλων τεχνολογικών εργαλείων. Ενώ η TN μπορεί να αναλύσει μεγάλα σύνολα δεδομένων και να εξάγει συμπεράσματα με ακρίβεια, η διαδικασία της δεν περιλαμβάνει προσωπικές εμπειρίες ή ενσυναίσθηση. Για παράδειγμα, μια κάμερα με TN μπορεί να αναγνωρίσει πρόσωπα ή αντικείμενα με υψηλή ακρίβεια, αλλά δεν κατανοεί τη συναισθηματική σημασία που μπορεί να έχει ένα βλέμμα ή μια έκφραση για έναν άνθρωπο. Επομένως, η TN δεν «κατανοεί» το περιβάλλον της όπως ένας άνθρωπος, αλλά το επεξεργάζεται με μηχανικό τρόπο, βασισμένο αποκλειστικά στα δεδομένα που της παρέχονται.

- **Συναίσθημα**

Το συναίσθημα είναι θεμελιώδες στοιχείο της ανθρώπινης ύπαρξης. Οι άνθρωποι όχι μόνο βιώνουν συναισθήματα, αλλά και τα εκφράζουν με τρόπους που επηρεάζουν τη συμπεριφορά τους και τη λήψη αποφάσεων. Η αγάπη, η θλίψη, η χαρά, ο θυμός και πολλά άλλα συναισθήματα παίζουν καθοριστικό ρόλο στις ανθρώπινες σχέσεις καθώς και στις καθημερινές επιλογές. Τα συναισθήματα συνδέονται στενά με την ενσυναίσθηση – την ικανότητα να κατανοούμε και να συμμεριζόμαστε τα συναισθήματα των άλλων. Για παράδειγμα, ένας άνθρωπος μπορεί να προσφέρει βοήθεια σε κάποιον που βρίσκεται σε ανάγκη, όχι επειδή είναι λογικό, αλλά επειδή νιώθει συμπόνια.

Η TN, από την άλλη πλευρά, δεν βιώνει συναισθήματα, καθώς δεν έχει συνείδηση ή νευροβιολογικές διαδικασίες που να τα υποστηρίζουν. Παρ' όλα αυτά, μπορεί να προσομοιώνει συναισθήματα μέσω αλγορίθμων. Για παράδειγμα, ένα chatbot μπορεί να απαντήσει με τρόπο που να φαίνεται φιλικός ή συμπονετικός, αλλά αυτό είναι αποτέλεσμα προγραμματισμού και όχι αυθεντικής συναισθηματικής απόκρισης. Ένα άλλο παράδειγμα είναι η ανάλυση συναισθημάτων σε κείμενα, όπου η TN μπορεί να εντοπίσει αν ένα μήνυμα εκφράζει θυμό, χαρά ή λύπη, χωρίς να κατανοεί ουσιαστικά τα συναισθήματα που υποδηλώνονται.

- **Συνείδηση**

Η ανθρώπινη συνείδηση είναι μία από τις πιο πολύπλοκες και αινιγματικές πτυχές της ύπαρξης και περιλαμβάνει την ικανότητα του ανθρώπου να σκέφτεται για τον εαυτό του, να αντιλαμβάνεται τη

θέση του στον κόσμο και να λειτουργεί με βάση ηθικές αξίες και αυτογνωσία. Μέσω της συνείδησης, οι άνθρωποι μπορούν να αξιολογούν τις πράξεις τους, να λαμβάνουν αποφάσεις με βάση την ηθική και να εξετάζουν τις συνέπειες των επιλογών τους. Επιπλέον, η συνείδηση επιτρέπει στους ανθρώπους να δημιουργούν αφηρημένες ιδέες, να ονειρεύονται το μέλλον και να κατανοούν έννοιες όπως ο χρόνος, η δικαιοσύνη και η ελευθερία.

Αντίθετα, η ΤΝ στερείται συνείδησης· δεν έχει αυτογνωσία, δεν αντιλαμβάνεται τον εαυτό της ως «οντότητα» και δεν διαθέτει ηθική κρίση. Η λειτουργία της ΤΝ βασίζεται αποκλειστικά σε κανόνες, δεδομένα και μαθηματικά μοντέλα. Για παράδειγμα, ένα αυτόνομο όχημα μπορεί να αποφασίσει πώς να αντιδράσει σε μια επείγουσα κατάσταση με βάση αλγορίθμους που δίνουν προτεραιότητα στην ασφάλεια, αλλά η απόφαση αυτή δεν είναι αποτέλεσμα ηθικής κρίσης διότι δεν έχει τη δυνατότητα να εξετάσει φιλοσοφικά ερωτήματα ή να κατανοήσει την έννοια της «υπαρξιακής ευθύνης».

Συμπερασματικά, οι διαφορές στην αντίληψη, το συναίσθημα και τη συνείδηση αναδεικνύουν την αξεπέραστη πολυπλοκότητα της ανθρώπινης φύσης και, ενώ η ΤΝ μπορεί να επιτύχει εκπληκτικά αποτελέσματα μέσω της τεχνολογίας, παραμένει ένα εργαλείο χωρίς πραγματική κατανόηση ή επίγνωση. Αυτή η θεμελιώδης διαφορά υπενθυμίζει ότι η ΤΝ, όσο προχωρημένη κι αν γίνει, δεν μπορεί να αντικαταστήσει τη μοναδικότητα της ανθρώπινης εμπειρίας και ύπαρξης.

Μελλοντικές προοπτικές: Στρατηγικές & Προκλήσεις για το μέλλον της τεχνητής νοημοσύνης και της φιλοσοφίας

Όπως έχει επισημανθεί και προηγουμένως, η ΤΝ συνεχώς εξελίσσεται, προσφέροντας αμέτρητες δυνατότητες για την κοινωνία, την οικονομία, την τεχνολογία. Ωστόσο, η ταχεία πρόοδος της συνοδεύεται από μια σειρά στρατηγικών και προκλήσεων που απαιτούν σοβαρή σκέψη και προετοιμασία για το μέλλον. Η φιλοσοφία, με την ικανότητά της να εντοπίζει τα θεμελιώδη ερωτήματα γύρω από την ηθική, την κοινωνία και την ανθρώπινη ύπαρξη, διαδραματίζει κρίσιμο ρόλο στην κατεύθυνση ανάπτυξης και χρήσης της ΤΝ.

Αναφορικά με τις στρατηγικές για το μέλλον της ΤΝ, πρώτα και κύρια, είναι απαραίτητο να εξασφαλιστεί η διαφάνεια και η ηθική στον τρόπο που χρησιμοποιούνται οι αλγόριθμοι και τα δεδομένα. Η διαφάνεια επιτρέπει την κατανόηση του τρόπου με τον οποίο η ΤΝ λαμβάνει αποφάσεις, κάτι που είναι ζωτικής σημασίας σε τομείς όπως η αυτόνομη οδήγηση ή οι ιατρικές διαγνώσεις, όπου κάθε απόφαση μπορεί να έχει σοβαρές συνέπειες. Παράλληλα, η ενσωμάτωση ηθικών αρχών στις αλγοριθμικές διαδικασίες θα συμβάλει στην αποφυγή καταχρήσεων της τεχνολογίας και στην ενίσχυση της κοινωνικής δικαιοσύνης.

Μια άλλη στρατηγική είναι η εκπαίδευση και η ευαισθητοποίηση των πολιτών και των εργαζομένων για τις δυνατότητες και τις προκλήσεις της ΤΝ. Η ανάπτυξη ψηφιακών δεξιοτήτων είναι απαραίτητη για να μπορέσουν οι άνθρωποι να προσαρμοστούν στις νέες τεχνολογίες και να αναγνωρίσουν τα οφέλη καθώς και τους κινδύνους τους. Επίσης, η σωστή εκπαίδευση μπορεί να βοηθήσει στην προετοιμασία των εργαζομένων για τις αλλαγές στην αγορά εργασίας που προκαλούνται από την αυτοματοποίηση, μειώνοντας τις αρνητικές κοινωνικές επιπτώσεις.

Η ανάγκη για παγκόσμια συνεργασία και αποτελεσματική ρύθμιση της ΤΝ είναι επίσης επιτακτική. Η διαμόρφωση διεθνών συμφωνιών και κανόνων χρήσης της ΤΝ θα αποτρέψει την ανάπτυξη της με καταχρηστικούς ή καταστρεπτικούς τρόπους, όπως στην περίπτωση των στρατιωτικών εφαρμογών ή της παρακολούθησης των πολιτών. Εν μέσω αυτών των στρατηγικών, η ανθρωποκεντρική προσέγγιση στην ανάπτυξη της ΤΝ είναι κεντρικής σημασίας. Η τεχνολογία πρέπει να εξυπηρετεί τις ανθρώπινες ανάγκες, προωθώντας την ισότητα και την πρόσβαση σε καλύτερες συνθήκες ζωής για όλους, ανεξαρτήτως της οικονομικής ή γεωγραφικής τους θέσης.

Ωστόσο, παρά τις στρατηγικές αυτές, οι προκλήσεις που προκύπτουν από τη χρήση της ΤΝ είναι εξίσου σοβαρές. Ένα από τα μεγαλύτερα προβλήματα που ενδέχεται να προκύψουν είναι η αυξανόμενη ανεργία, καθώς πολλές θέσεις εργασίας, ιδιαίτερα σε τομείς που απαιτούν επαναλαμβανόμενες και προβλέψιμες εργασίες, κινδυνεύουν να αυτοματοποιηθούν. Αυτό μπορεί να οδηγήσει σε οικονομικές ανισότητες και κοινωνικές εντάσεις. Αντίστοιχα, οι κοινωνικές ανισότητες που δημιουργούνται από την ανισότιμη πρόσβαση στην ΤΝ μπορούν να διευρύνουν το χάσμα μεταξύ ανεπτυγμένων και αναπτυσσόμενων χωρών.

Ακόμα, η ηθική διάσταση της ΤΝ αποτελεί μια από τις μεγαλύτερες προκλήσεις. Ερωτήματα όπως ποιος είναι υπεύθυνος για τις αποφάσεις που λαμβάνονται από την ΤΝ ή πώς μπορεί να διασφαλιστεί ότι οι αποφάσεις αυτές είναι δίκαιες και χωρίς προκαταλήψεις, απαιτούν λεπτομερή και συνεχόμενη συζήτηση. Η ΤΝ ενδέχεται να προκαλέσει σημαντικά ηθικά διλήμματα σε τομείς όπως η στρατιωτική τεχνολογία ή η χρήση της για την παρακολούθηση των πολιτών.

Επιπλέον, η ψηφιακή ανισότητα παραμένει μια σημαντική πρόκληση. Η πρόσβαση στην ΤΝ και στις τεχνολογίες αιχμής είναι άνιση και πολλές χώρες και κοινωνικές ομάδες ενδέχεται να μείνουν πίσω στην «ψηφιακή επανάσταση», ενισχύοντας την οικονομική και κοινωνική ανισότητα.

Τέλος, η προστασία των προσωπικών δεδομένων και η ασφάλεια των πληροφοριών είναι προβλήματα που πρέπει να αντιμετωπιστούν με αυστηρούς κανονισμούς και διαδικασίες. Η χρήση της ΤΝ απαιτεί τη συλλογή και ανάλυση μεγάλου όγκου δεδομένων, γεγονός που μπορεί να οδηγήσει σε σοβαρές παραβιάσεις της ιδιωτικότητας, εάν δεν υπάρξουν επαρκείς προστατευτικές δομές.

Η φιλοσοφία, με τις ηθικές και ανθρωπιστικές της αρχές, παίζει καθοριστικό ρόλο στη διαμόρφωση του μέλλοντος της ΤΝ. Οι φιλοσοφικές θεωρίες για την ηθική, την κοινωνία και την ανθρώπινη φύση μπορούν να προσφέρουν τα απαραίτητα εργαλεία για να κατανοήσουμε τις συνέπειες της τεχνολογίας και να διασφαλίσουμε ότι η ΤΝ αναπτύσσεται με τρόπο που θα ωφελεί την ανθρωπότητα και θα σέβεται τις αξίες της.

Σε αυτό το πλαίσιο, η φιλοσοφία είναι αναγκαία για να καθοδηγήσει τις αποφάσεις μας, προκειμένου η ανάπτυξη της ΤΝ να είναι βιώσιμη και να ανταποκρίνεται στις ηθικές και κοινωνικές ανάγκες του μέλλοντος. Η συνεργασία των φιλοσόφων, των τεχνολόγων, των νομικών και των πολιτικών θα είναι κρίσιμη για την επιτυχή ενσωμάτωση της ΤΝ στην κοινωνία, ώστε να μπορέσουμε να επωφεληθούμε από τις δυνατότητές της, ενώ ταυτόχρονα να αποφύγουμε τις κοινωνικές και ηθικές παγίδες που ενδέχεται να προκύψουν.

Βιβλιογραφία & Πηγές για περαιτέρω μελέτη:

- Οι διαφάνειες των ομιλητών του εκπαιδευτικού προγράμματος.
- Aaronson S. (2013). *Quantum Computing Since Democritus*, Cambridge University Press, Cambridge.
- Adetomiwa F. (2021). *Artificial Intelligence and Human Rights*. SSRN Electronic Journal. 10.2139/ssrn.3964199.
- Agar N. (2020). *How to Treat Machines That Might Have Minds*. *Philosophy & Technology*, 33(2), 269–82.
- Allen C., Smit I., Wallach W. (2005). *Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches*. *Ethics & Information Technology*, 7 (3), 149-155.
- Altman M. C. (2011). *Kant and Applied Ethics: The Uses and Limits of Kant's Practical Philosophy*. Malden, NJ: Wiley-Blackwell.
- Anderson E. (2017). *Private Government: How Employers Rule Our Lives (and Why We Don't Talk about It)*. Princeton, NJ: Princeton University Press.
- Anderson M., Anderson S. (2011). *Machine Ethics*. Cambridge: Cambridge University Press.
- Anderson S. L. (2011). *Machine Metaethics*. In M. Anderson and S. L. Anderson (Eds.), *Machine Ethics*, 21–27. Cambridge: Cambridge University Press.
- Arkin R. (2009). *Governing Lethal Behavior in Autonomous Robots*. Boca Raton, FL: CRC Press.
- Arkin R. (2010). *The Case for Ethical Autonomy in Unmanned Systems*. *Journal of Military Ethics*, 9(4), 332–41.
- Asaro P. M. (2019). *AI ethics in predictive policing: From models of threat to an ethics of care*. *IEEE Technology and Society Magazine*, 38 (2), 40-53.
- Atapattu S. (2015). *Human Rights Approaches to Climate Change: Challenges and Opportunities*. New York: Routledge.
- Awad E., Dsouza S., Kim R., Schulz J., Henrich J., Shariff A., Bonnefon J. F., Rahwan I. (2018). *The Moral Machine Experiment*. *Nature*, 563, 59–64.
- Boden M.A. (2016). *AI: Its Nature and Future*, Oxford University Press, Oxford.
- Borenstein J. Arkin R. (2016). *Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being*. *Science and Engineering Ethics*, 22, 31–46.
- Bostrom N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.
- Bostrom N. (2013). *Existential risk prevention as global priority*. *Global Policy*, 4 (1), 15-31.
- Bostrom N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
- Bryson J. (2012). *A Role for Consciousness in Action Selection*. *International Journal of Machine Consciousness*, 4(2), 471–82.
- Bryson J. (2019). *Patience Is Not a Virtue: The Design of Intelligent Systems and Systems of Ethics*. *Ethics and Information Technology*, 20(1), 15–26.
- Coeckelbergh M. (2010). *Robot rights? Towards a social-relational justification of moral consideration*. *Ethics and Information Technology*, 12 (3), 209-221.
- Dreyfus H.L. (2001). *Τι δεν μπορούν ακόμη να κάνουν οι υπολογιστές; , μετάφραση Πόπη Καρλέτσα, επιμέλεια Κώστας Χατζηκυριάκου, Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο.*

- Gordon, John-Stewart & Nyholm, Sven. (2021). Ethics of Artificial Intelligence. Available here: https://www.researchgate.net/publication/349467117_Ethics_of_Artificial_Intelligence (retrieved 23/1/2025).
- Gunkel, David. (2012). The Machine Question: Critical Perspectives on AI, Robots, and Ethics. 10.7551/mitpress/8975.001.0001.
- Kamm F. M., Rakowski E. (Eds.). (2016). The trolley problem mysteries. New York: Oxford University Press.
- Kurzweil R. (1999). The age of spiritual machines: When computers exceed human intelligence. London: Penguin.
- Kurzweil R. (2005). The singularity is near: When humans transcend biology. London: Viking.
- Kurzweil R. (2012). How to create a mind: The secret of human thought revealed. New York: Viking.
- Müller, Vincent C. (forthcoming 2021), 'Ethics of artificial intelligence', in Anthony Elliott (ed.), The Routledge social science handbook of AI (London: Routledge), 20pp.
- Oluoch I. (2024). Crossing Boundaries: The Ethics of AI and Geographic Information Technologies. ISPRS International Journal of Geo-Information, 13 (3), p. 87. <https://doi.org/10.3390/ijgi13030087>
- Tegmark M. (2015). Το Μαθηματικό Σύμπαν μας, μετάφραση Νίκος Αποστολόπουλος, επιστημονική επιμέλεια Κώστας Κλειΐδης, Εκδοτικός Οίκος Τραυλός, Αθήνα.
- Zerilli J., Knott A., Maclaurin J., Gavaghan C. (2019). Transparency in algorithmic and human decision-making: Is there a double standard? Philosophy & Technology, 32 (4), 661–683.
- Κατεχάκη Β. (2014). Όταν η τεχνητή νοημοσύνη φιλοσοφεί με τον Αριστοτέλη. Διαθέσιμο στο: <https://www.kathimerini.gr/opinion/interviews/563086717/otan-i-technitinoimosyni-filosofei-me-ton-aristoteli/> (προσπελάστηκε 23/1/2025).
- Μαθηματικά, Φιλοσοφία και Τεχνητή Νοημοσύνη – Ένας διάλογος με τον Gian-Carlo Rota και τον David Sharp, μετάφραση: Ανδρομάχη Σπανού, επιστημονική επιμέλεια & πρόλογος: Γιώργος Λ. Ευαγγελόπουλος (2019). Εκδόσεις: Ευρασία, Αθήνα.
- Τάσης Θ. (2019). Ψηφιακός Ανθρωπισμός: Εικονιστικό Υποκείμενο και Τεχνητή Νοημοσύνη. Αθήνα: Αρμός, Αθήνα.
- Χατζηβασιλείου Γ. (2023). Φιλοσοφία της Τεχνητής Νοημοσύνης. Αθήνα: Εκδόσεις Διόπτρα.



ΕΚΠΑΙΔΕΥΤΙΚΟ ΥΛΙΚΟ
ΓΙΑ ΤΟ ΔΙΑΔΙΚΤΥΑΚΟ
ΕΚΠΑΙΔΕΥΤΙΚΟ
ΠΡΟΓΡΑΜΜΑ

12.02.2025 ΑΘΗΝΑ



ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ & ΦΙΛΟΣΟΦΙΑ:

Από την ανθρώπινη σκέψη
στη μηχανική μάθηση



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ

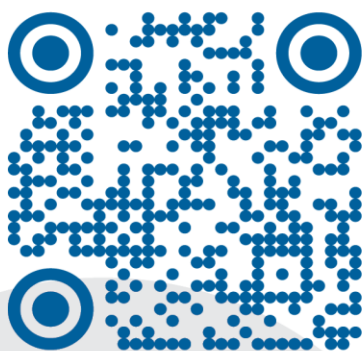
Αυτό το υλικό δημιουργήθηκε στο πλαίσιο του Ευρωπαϊκού Κόμβου Ψηφιακής Καινοτομίας Smart Attica (Smart Attica EDIH), του πρώτου Κόμβου EDIH για την Τεχνητή Νοημοσύνη στην Ελλάδα. Ο Κόμβος αποτελείται από 17 εταιρίες στην Ελλάδα, μεταξύ των οποίων είναι και το Εθνικό Κέντρο Τεκμηρίωσης & Ηλεκτρονικού Περιεχομένου (ΕΚΤ). Το παρόν αποτελεί υποστηρικτικό υλικό για τα εκπαιδευτικά προγράμματα Ιανουαρίου – Μαρτίου 2025 και συγκεκριμένα, για το δεύτερο, με τίτλο «Τεχνητή Νοημοσύνη & Φιλοσοφία: Από την ανθρώπινη σκέψη στη μηχανική μάθηση» που υλοποίησε διαδικτυακά το Εθνικό Κέντρο Τεκμηρίωσης & Ηλεκτρονικού Περιεχομένου (ΕΚΤ), 12/02/2025.

Το έργο Smart Attica υλοποιείται με τη συγχρηματοδότηση της Ευρωπαϊκής Ένωσης. Ωστόσο, οι απόψεις και οι γνώμες που διατυπώνονται εκφράζουν αποκλειστικά τις απόψεις των συντακτών και δεν αντιπροσωπεύουν κατ' ανάγκη τις απόψεις της Ευρωπαϊκής Ένωσης. Ούτε η Ευρωπαϊκή Ένωση ούτε η χορηγούσα αρχή μπορούν να θεωρηθούν υπεύθυνες για αυτές.

Το επικοινωνιακό σκέλος του εκπαιδευτικού προγράμματος επιμελήθηκε η κυρία Κωνσταντίνα Κωνσταντινίου.

Η συλλογή, η αποδελτίωση και η συγγραφή του εκπαιδευτικού υλικού πραγματοποιήθηκε από τη Δρα Μαρία Μαλαγκονιάρη. Ο γραφιστικός σχεδιασμός του υλικού πραγματοποιήθηκε από την κυρία Δήμητρα Πελεκάνου. Την επιστημονική επίβλεψη είχε ο Δρ. Νικόλαος Καραμπέκιος.

- ✦ Περισσότερες πληροφορίες για το ΕΚΤ, δείτε: <https://www.ekt.gr/>
- ✦ Περισσότερες πληροφορίες για τις Υπηρεσίες για Δικτύωση, Συνεργασία & Ανάπτυξη, δείτε: <https://innovation.ekt.gr/>
- ✦ Περισσότερες πληροφορίες για το Smart Attica EDIH στο ΕΚΤ, δείτε: <https://innovation.ekt.gr/smart-attica> & <https://www.ekt.gr/el/projects/28848>



**SMART
ATTICA** European
Digital
Innovation
Hub

210 2204895 | smartattica@ekt.gr
Ζεφύρου 56, 17564 Παλαιό Φάληρο

**ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ**



Με τη συγχρηματοδότηση
της Ευρωπαϊκής Ένωσης



ΠΡΟΓΡΑΜΜΑ

2021 – 2027

ΑΝΤΑΓΩΝΙΣΤΙΚΟΤΗΤΑ

Με τη συγχρηματοδότηση της Ευρωπαϊκής Ένωσης. Ωστόσο, οι απόψεις και οι γνώμες που διατυπώνονται εκφράζουν αποκλειστικά τις απόψεις των συντακτών και δεν αντιπροσωπεύουν κατ' ανάγκη τις απόψεις της Ευρωπαϊκής Ένωσης. Ούτε η Ευρωπαϊκή Ένωση ούτε η χορηγούσα αρχή μπορούν να θεωρηθούν υπεύθυνες για αυτές.