

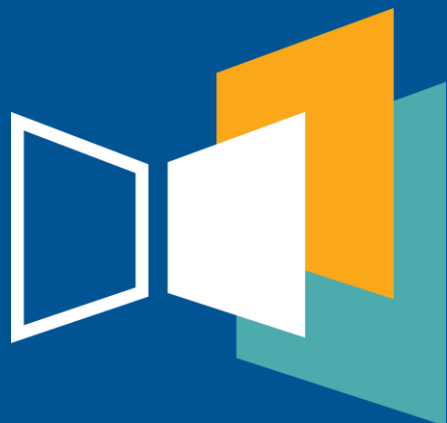


**SMART
ATTICA**

European
Digital
Innovation
Hub



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ



**AI-powering
Greece**

**Εκπαιδευτικό
Πρόγραμμα (1)**

Τετάρτη 12 Φεβρουαρίου 2025

ΤΝ και φιλοσοφία: Από την ανθρώπινη σκέψη στη μηχανική μάθηση

«Ανθρωποκεντρική χρήση τεχνητής νοημοσύνης και βιοηθική»

Δρ Μαρία Κ. Χωριανοπούλου

Επίκουρη Καθηγήτρια Πρακτικής Φιλοσοφίας & Ιατρικής Ηθικής



Co-funded by
the European Union



PROGRAMME 2021 – 2027
COMPETITIVENESS

Εισαγωγή: «Πέρα από τον Αλγόριθμο: Ηθική, Αυτονομία και η Αξία του Ανθρώπινου Παράγοντα στην Τεχνητή Νοημοσύνη»



- ❑ Η ΤΝ πρέπει να υπηρετεί τον άνθρωπο, ενισχύοντας την ευημερία, την ελευθερία και την αξιοπρέπεια.
- ❑ Οι 4 αρχές της βιοηθικής αποτελούν ηθικά θεμέλια για την ανάπτυξη της ΤΝ.
- ❑ Η αλγοριθμική λήψη αποφάσεων εγείρει ζητήματα ευθύνης, διαφάνειας και μεροληψίας.
- ❑ Η προστασία των προσωπικών δεδομένων είναι κρίσιμη για τη διασφάλιση των δικαιωμάτων και της αξιοπρέπειας των χρηστών.
- ❑ Η θεωρία του «extended mind» επαναπροσδιορίζει τα όρια της ανθρώπινης νόησης και την ηθική ευθύνη.
- ❑ Η ανθρώπινη αυτονομία πρέπει να παραμένει κεντρική, ακόμα και σε έναν κόσμο ΤΝ.
- ❑ Η ηθική «από τον σχεδιασμό» (ethics by design) είναι αναγκαία για την πρόληψη ανεπιθύμητων συνεπειών.
- ❑ Η διεπιστημονική συνεργασία είναι απαραίτητη για ηθικές και κοινωνικά βιώσιμες λύσεις.
- ❑ Η εκπαίδευση και η ευαισθητοποίηση σχετικά με την ΤΝ ενισχύουν την υπεύθυνη χρήση της τεχνολογίας.
- ❑ Η βιοηθική προσφέρει έναν ηθικό φάρο για τη διατήρηση της ανθρωποκεντρικής διάστασης της τεχνολογίας.



Τι σημαίνει *ανθρωποκεντρική ΤΝ*:

1. Ο σχεδιασμός και η χρήση της ΤΝ πρέπει να υπηρετούν τον άνθρωπο.
2. Να ενισχύουν την ευημερία, την ελευθερία και την αξιοπρέπεια.

Βιοηθική και ΤΝ: Οι 4 βασικές αρχές

1. Αυτονομία: Σεβόμαστε την ικανότητα του ατόμου να λαμβάνει αποφάσεις.
2. Ευεργεσία: Στόχος η προαγωγή του καλού και της υγείας.
3. Μη βλάβη: Αποφεύγουμε να προκαλέσουμε βλάβη ή αδικία.
4. Δικαιοσύνη: Ισομερής και δίκαιη πρόσβαση στα οφέλη της ΤΝ, χωρίς διακρίσεις.

Σε αυτό το πλαίσιο, η ΤΝ πρέπει να αναπτύσσεται με ηθικές εγγυήσεις «από τον σχεδιασμό» (ethics by design), ώστε να προλαμβάνει ανεπιθύμητες συνέπειες.



Αλγοριθμική Λήψη Αποφάσεων:

1. Στον χώρο της υγείας, μια λανθασμένη πρόγνωση ή διάγνωση μπορεί να έχει σοβαρές συνέπειες.
2. Τα τυφλά σημεία (bias) στους αλγόριθμους μπορούν να οδηγήσουν σε μεροληπτική μεταχείριση.

Διαφάνεια και Λογοδοσία:

1. Πώς εξηγούμε τη διαδικασία λήψης αποφάσεων του αλγορίθμου;
2. Ποιος αναλαμβάνει την ευθύνη σε περίπτωση βλάβης ή λάθους πρόβλεψης;

Προστασία Προσωπικών Δεδομένων:

1. Ειδικά στον τομέα της υγείας, η διασφάλιση του απορρήτου είναι θεμελιώδης.
2. Οφείλουμε να διασφαλίζουμε ότι δεν παραβιάζονται δικαιώματα ασθενών και πολιτών.

Φιλοσοφικές προεκτάσεις & ο Ανθρώπινος παράγοντας

- **Extended Mind:**
- Η ΤΝ μπορεί να λειτουργήσει ως *προέκταση* των νοητικών μας ικανοτήτων.
- Ωστόσο, αυτό εγείρει ερωτήματα: πού τελειώνει η ανθρώπινη απόφαση και πού αρχίζει η επιρροή της μηχανής;
- **Αναπλασίωση της Ανθρώπινης Ταυτότητας:**
- Η εισαγωγή της ΤΝ αλλάζει σταδιακά την αντίληψή μας για το τι είναι νόηση και δημιουργικότητα.
- Η ηθική διάσταση υπενθυμίζει ότι η ενσυναίσθηση, η αναστοχαστική σκέψη και η ευθύνη δεν αποτελούν αλγοριθμικά προγράμματα που κωδικοποιούνται.

Προτάσεις για Υπεύθυνη Ανάπτυξη ΤΝ



Ethics by Design:

Η ηθική ενσωματώνεται στο στάδιο σχεδιασμού των αλγορίθμων.
Κριτήρια ποιοτικού ελέγχου, διαφάνειας και λογοδοσίας από την αρχή.

Διεπιστημονική Συνεργασία:

Για σύνθετα ζητήματα απαιτείται συνεργασία προγραμματιστών, ιατρών, νομικών, φιλοσόφων και κοινωνιολόγων.
Η διαφορετική οπτική βοηθά να αποφύγουμε περιπτώσεις μεροληψίας και αδικίας.

Δια βίου Εκπαίδευση: Στο ακαδημαϊκό και επαγγελματικό περιβάλλον, προγράμματα σπουδών που εντάσσουν την ηθική της ΤΝ στον πυρήνα τους.

Ενημέρωση και ευαισθητοποίηση του κοινού για τις δυνατότητες και τα ρίσκα της ΤΝ.

ΕΝΔΕΙΚΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ



- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of Biomedical Ethics* (8th ed.). Oxford University Press.
- Dworkin, R. (1977). *Taking Rights Seriously*. Harvard University Press.
- Gewirth, A. (1978). *Reason and Morality*. University of Chicago Press.
- Häyry, M., & Takala, T. (Eds.). (2022). *Ethics of AI in Context*. Springer.
- Jonas, H. (1984). *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. University of Chicago Press.
- Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten (Groundwork of the Metaphysics of Morals)*.
- Kant, I. (1797). *Die Metaphysik der Sitten (The Metaphysics of Morals)*.
- O'Neill, O. (1996). *Towards Justice and Virtue: A Constructive Account of Practical Reasoning*. Cambridge University Press.
- Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.
- Savulescu, J., & Persson, I. (2012). *Unfit for the Future: The Need for Moral Enhancement*. Oxford University Press.



**SMART
ATTICA** European
Digital
Innovation
Hub



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ &
ΗΛΕΚΤΡΟΝΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ



Ευχαριστώ!

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.