



Οδηγός Εκπαιδευτικού

Τεχν  
ητή  
Νοημο  
σύνη

5

Δεοντολογία, Ηθική και ασφάλεια  
Διαδικτύου

## Εισαγωγή στη δραστηριότητα

Η τεχνητή νοημοσύνη (AI) είναι μια γρήγορα αναπτυσσόμενη τεχνολογία με τη δυνατότητα να μεταμορφώσει τον τρόπο με τον οποίο ζούμε και εργαζόμαστε. Ωστόσο, το AI θέτει επίσης μια σειρά από ηθικές και δεοντολογικές προκλήσεις.

Το AI είναι ένα εργαλείο με ευρύ πεδίο εφαρμογών. Από τη μία μπορεί να χρησιμοποιηθεί για τη βελτίωση της υγείας, της εκπαίδευσης και της ασφάλειας. Από την άλλη, μπορεί να οδηγήσει στη δημιουργία αυτόνομων όπλων, στην προώθηση παραπληροφόρησης και στην καταπάτηση των ανθρωπίνων δικαιωμάτων. Είναι σημαντικό να διασφαλίσουμε ότι το AI χρησιμοποιείται με υπευθυνότητα και ότι δεν χρησιμοποιείται για να βλάψει τους ανθρώπους ή το περιβάλλον.

Οι παραπάνω, είναι μόνο μερικές από τις ηθικές και δεοντολογικές προκλήσεις που θέτει το AI. Είναι σημαντικό να ασχοληθούμε με αυτές τις προκλήσεις, καθώς το AI συνεχίζει να αναπτύσσεται.



Εικόνα 1.

### Απευθύνεται σε:

Εκπαιδευτικούς και κάθε ενδιαφερόμενο χρήστη χωρίς προηγούμενη εμπειρία με την τεχνητή νοημοσύνη που θέλει να χρησιμοποιεί με ασφάλεια το AI.

### Στόχοι δραστηριότητας:

- Η κατανόηση των ηθικών διλημάτων που προκύπτουν σχετικά με τη χρήση και τις εφαρμογές του AI.
- Η ικανότητα χρήσης των εργαλείων τεχνητής νοημοσύνης με ασφάλεια.



## Εκτέλεση δραστηριότητας

### Ηθική & Δεοντολογία

Η ηθική και η δεοντολογία στο AI είναι ένας νέος και αναπτυσσόμενος κλάδος που εξετάζει τις ηθικές και δεοντολογικές συνέπειες και τα όρια της τεχνητής νοημοσύνης. Στον τομέα του AI, η δεοντολογία ασχολείται με το πώς πρέπει να χρησιμοποιούμε την τεχνητή νοημοσύνη με υπεύθυνο τρόπο. Η ηθική, από την άλλη είναι το σύστημα των αξιών και των πεποιθήσεων που καθοδηγούν τη συμπεριφορά ενός ατόμου.

Μερικοί από τους βασικούς προβληματισμούς της δεοντολογίας στο AI περιλαμβάνουν:

- Το AI μπορεί να χρησιμοποιηθεί με τρόπο επιβλαβή για τους ανθρώπους ή το φυσικό περιβάλλον. Για παράδειγμα, το AI μπορεί να χρησιμοποιηθεί για τη δημιουργία αυτόνομων όπλων, ιών, για την προώθηση της παραπληροφόρησης ή για την καταστροφή του φυσικού περιβάλλοντος.
- **Ευθύνη:** Όταν το AI λαμβάνει αποφάσεις που επηρεάζουν τη ζωή των ανθρώπων, είναι σημαντικό να υπάρχουν άνθρωποι που είναι υπεύθυνοι για αυτές τις αποφάσεις. Για παράδειγμα, ένα σύστημα AI που αποφασίζει ποιο άτομο θα λάβει ιατρική περίθαλψη, θα πρέπει να έχει ανθρώπους που θα αναλαμβάνουν την ευθύνη για τις αποφάσεις του συστήματος. Είναι σημαντικό να διασφαλίσουμε ότι υπάρχει ένα σύστημα ευθύνης για το AI.
- **Προκατάληψη:** Τα συστήματα AI εκπαιδεύονται σε δεδομένα που δημιουργήθηκαν από ανθρώπους. Ως αποτέλεσμα, τα συστήματα AI μπορεί να είναι προκατειλημμένα, αντικατοπτρίζοντας τις απόψεις των ανθρώπων που δημιούργησαν τα δεδομένα. Για παράδειγμα, ένα σύστημα AI που χρησιμοποιείται για την πρόβλεψη της εγκληματικότητας μπορεί να είναι προκατειλημμένο εναντίον των μειονοτήτων. Είναι σημαντικό να διασφαλίσουμε ότι τα συστήματα AI είναι απαλλαγμένα από τέτοιου είδους προκαταλήψεις.
- **Επιδείνωση των ανισοτήτων:** Το AI έχει τη δυνατότητα να επιδεινώσει τις υπάρχουσες ανισότητες. Για παράδειγμα, το AI μπορεί να χρησιμοποιηθεί για την αυτοματοποίηση εργασιών που σήμερα εκτελούνται από ανθρώπους, γεγονός που μπορεί να οδηγήσει σε απώλεια θέσεων εργασίας και αύξηση της ανεργίας. Είναι σημαντικό να διασφαλίσουμε ότι το AI χρησιμοποιείται για να μειώσει τις ανισότητες και όχι για να τις ενισχύσει.

## Εκτέλεση δραστηριότητας

Για να αντιμετωπίσουμε αυτές τις προκλήσεις, είναι σημαντικό:

- **Να αναπτύξουμε κώδικες δεοντολογίας για το AI:**  
Οι κώδικες δεοντολογίας μπορούν να βοηθήσουν στην καθοδήγηση της ανάπτυξης και της χρήσης του AI με υπεύθυνο τρόπο. Υπάρχουν ήδη πολλοί κώδικες δεοντολογίας για το AI, αλλά είναι σημαντικό να συνεχίσουμε να αναπτύσσουμε και να βελτιώνουμε αυτούς τους κώδικες.
- **Να διεξάγουμε έρευνα για τις ηθικές και δεοντολογικές επιπτώσεις του AI:**  
Η έρευνα μπορεί να μας βοηθήσει να κατανοήσουμε καλύτερα τις πιθανές επιπτώσεις του AI και να αναπτύξουμε στρατηγικές για τη μείωσή τους.
- **Να διεξάγουμε δημόσιες συζητήσεις για τις ηθικές και δεοντολογικές συνέπειες του AI:**  
Οι δημόσιες συζητήσεις μπορούν να βοηθήσουν στην ευαισθητοποίηση του κοινού για τις ηθικές και δεοντολογικές επιπτώσεις του AI και στη διατύπωση απαιτήσεων για υπεύθυνους τρόπους ανάπτυξης και χρήσης του AI.



Εικόνα 2.

Η δεοντολογία στο AI είναι ένας σημαντικός τομέας που θα συνεχίσει να εξελίσσεται καθώς αναπτύσσεται η τεχνολογία του AI. Είναι σημαντικό να ενημερωνόμαστε για τις ηθικές και δεοντολογικές επιπτώσεις του AI και να συμμετέχουμε στις συζητήσεις για τον υπεύθυνο τρόπο ανάπτυξης και χρήσης του AI.



## Εκτέλεση δραστηριότητας

### AI & Ασφάλεια Διαδικτύου

Η ασφάλεια του Διαδικτύου και η Τεχνητή Νοημοσύνη είναι δύο αλληλένδετα πεδία με αμφίδρομες εξαρτήσεις. Από τη μία πλευρά, τα περισσότερα συστήματα AI μεγάλης κλίμακας είναι συνδεδεμένα στο Διαδίκτυο και βασίζονται σε αυτό τόσο για να παρέχουν τις υπηρεσίες τους, όσο και για να αντλούν δεδομένα και πληροφορίες. Σαν αποτέλεσμα, είναι εκτεθειμένα σε απειλές που προέρχονται από το Διαδίκτυο και μπορούν να προκαλέσουν σημαντικά προβλήματα στην απρόσκοπτη και ασφαλή λειτουργία τους:

- **Επιθέσεις Άρνησης Υπηρεσίας (DoS - Denial of Service).** Μία επίθεση άρνησης υπηρεσίας έχει ως στόχο να εμποδίσει ένα σύστημα να παρέχει τις υπηρεσίες του πάνω από το Διαδίκτυο. Υλοποιείται μέσω της υποβολής ενός τεράστιου όγκου ψευδών αιτημάτων, που έχουν ως αποτέλεσμα την εξάντληση των πόρων του συστήματος και την αδυναμία του να εξυπηρετήσει πραγματικά αιτήματα. Φανταστείτε τις συνέπειες που θα είχε μία τέτοια επίθεση σε ένα σύστημα AI που παρέχει πληροφορίες πλοήγησης σε οδηγούς.
- **Ιοί.** Οι ιοί είναι κακόβουλα προγράμματα που προσβάλλουν το λογισμικό των υπολογιστών και τους προκαλούν δυσλειτουργίες. Η βασικότερη πηγή ιών είναι το Διαδίκτυο. Το λογισμικό των συστημάτων AI που συνδέονται στο Διαδίκτυο μπορεί να προσβληθεί από ιούς, που θα προκαλέσουν ζημιά ή ακόμα και ολική διακοπή της λειτουργίας τους.
- **Hackers.** Οι hackers είναι κακόβουλοι χρήστες οι οποίοι εκμεταλλεύονται ευπάθειες ασφάλειας που υπάρχουν στα λειτουργικά συστήματα και τις εφαρμογές, ώστε να αποκτήσουν μη εξουσιοδοτημένη πρόσβαση σε ένα σύστημα, μέσω του Διαδικτύου. Αν ένα σύστημα AI παραβιαστεί από κάποιον hacker, οι επιπτώσεις μπορεί να είναι πολύ αρνητικές και ποικίλουν από την αλλοίωση της λειτουργίας του συστήματος ώστε να δίνει μη έγκυρες πληροφορίες, μέχρι την υποκλοπή σημαντικών δεδομένων των χρηστών του συστήματος.

Με βάση τα παραπάνω παραδείγματα, γίνεται αντιληπτό ότι η ασφάλεια διαδικτύου είναι ζωτικής σημασίας για τη λειτουργία ενός συνδεδεμένου συστήματος AI. Για το λόγο αυτό, τα δίκτυα που φιλοξενούν τέτοια συστήματα και υπηρεσίες, θα πρέπει να διαθέτουν ισχυρή κυβερνοασφάλεια.

Από την άλλη πλευρά, η ασφάλεια Διαδικτύου αποτελεί πεδίο εφαρμογής της Τεχνητής Νοημοσύνης. Πολλά εργαλεία κυβερνοασφάλειας χρησιμοποιούν αλγόριθμους Τεχνητής Νοημοσύνης, προκειμένου να προστατέψουν από απειλές τα δίκτυα και τα συστήματα στα οποία χρησιμοποιούνται. Ενδεικτικά αναφέρουμε τα ακόλουθα:

## Εκτέλεση δραστηριότητας

- **Antivirus.** Τα σύγχρονα εργαλεία προστασίας από ιούς, χρησιμοποιούν AI για να εντοπίζουν και να αναλύουν νέους ιούς.
- **E-mail filtering.** Οι εξυπηρετητές ηλεκτρονικής αλληλογραφίας, χρησιμοποιούν αλγόριθμους τεχνητής νοημοσύνης και μηχανική μάθηση, για την ανίχνευση και το φιλτράρισμα ανεπιθύμητων και κακόβουλων mail.
- **Firewall.** Ένα Firewall που χρησιμοποιεί AI, μπορεί να ανιχνεύει και να προστατεύει από εξελιγμένες απειλές, σε πραγματικό χρόνο, χωρίς να απαιτείται ανθρώπινη παρέμβαση.

Τέλος, η σχέση μεταξύ Τεχνητής Νοημοσύνης και ασφάλειας του Διαδικτύου έχει και μία “σκοτεινή” πλευρά, καθώς η Τεχνητή Νοημοσύνη μπορεί να αποτελέσει απειλή για την κυβερνοασφάλεια. Συγκεκριμένα, κατηγορίες επιθέσεων όπως οι Brute Force (εξαντλητικές δοκιμές για το σπάσιμο κωδικών πρόσβασης) και οι DoS (επιθέσεις άρνησης υπηρεσίας), χρησιμοποιούν αλγόριθμους AI για να πετύχουν τον στόχο τους.

Η τεχνητή νοημοσύνη είναι μια ισχυρή τεχνολογία με τη δυνατότητα να προσφέρει πολλά οφέλη. Ωστόσο, θέτει επίσης μια σειρά από ηθικές και δεοντολογικές προκλήσεις.

Είναι σημαντικό να ληφθούν υπόψη αυτές οι προκλήσεις κατά την ανάπτυξη και τη χρήση του AI. Με προσεκτική σκέψη και σχεδιασμό, μπορούμε να διασφαλίσουμε ότι το AI χρησιμοποιείται για την εξέλιξη των ανθρώπινων κοινωνιών και όχι για τη βλάβη οποιασδήποτε κοινωνικής ομάδας ή του φυσικού περιβάλλοντος.



Εικόνα 3.

## Εκτέλεση δραστηριότητας

Προτρέψτε τα παιδιά να κάνουν ερωτήσεις ηθικών διλημάτων, δεοντολογίας, ασφάλειας όπως οι παρακάτω σε κάποιο εργαλείο NLP (ChatGPT, Bing κ.α.) και ζητήστε τους να φτιάξουν μια παρουσίαση όπου θα συγκρίνουν τις θέσεις του AI και τις δικές τους παρατηρήσεις.

- Είναι πιο σωστό να επενδύει η ανθρωπότητα χρήματα στην εξερεύνηση του διαστήματος ή θα ήταν προτιμότερο να τα επενδύει στους τομείς της υγείας; Αναλύστε τα επιχειρήματά σας.
- Πώς θα μπορούσε να είναι εμφανισιακά ένα άτομο που ασχολείται με τον προγραμματισμό;
- Πώς θα μπορούσα να προστατευτώ από ενδεχόμενη κακόβουλη χρήση του AI; Ποιες είναι οι σκέψεις σας;
- Μπορεί το AI να αποτελέσει απειλή για την ιδιωτικότητα των δεδομένων μας;

Μπορείτε να οργανώσετε ένα debate στην τάξη σε σχέση με τις θέσεις που διατυπώθηκαν.



Εικόνα 4.



## Πηγές

<https://camosun.libguides.com/ai/ethicalconsiderations>

<https://www.coe.int/en/web/bioethics/common-ethical-challenges-in-ai>

<https://openai.com/blog/teaching-with-ai>

<https://www.bing.com/>

## Πηγές εικόνων

**Εικόνα 1:** <https://anteelo.com/ai-ethics-in-2021-ethical-dilemmas-which-needs-to-be-answered/>

**Εικόνα 2:** <https://stock.adobe.com/uk/images/knowledge-fusion-3d-ai-robots-studying-ethical-ai-principles-in-a-futuristic-library-books-titled-ethics-ai-development-and-human-values-ethical-ai-illuminates-the-space/718558714>

**Εικόνα 3:** <https://stock.adobe.com/uk/images/cybersecurity-online-concept-woman-sitting-at-laptop-with-shield-protection-of-personal-data-and-information-prevention-of-hacking-antivirus-cartoon-flat-vector-illustration/677371810>

**Εικόνα 4:** <https://stock.adobe.com/uk/images/schoolgirl-pupil-student-answering-at-the-blackboard-presenting-project-homework-during-class-lesson-at-school-teacher-listening-to-student-s-answer/458082401>

## Σημείωση 1

Το υλικό πνευματικής ιδιοκτησίας τρίτων (εικόνες με ή χωρίς προσαρμογή, προσαρμοσμένοι κώδικες, προσαρμοσμένο κείμενο κ.ά.) σημειώνεται ρητά και διανέμεται με την αντίστοιχη άδεια που ορίζεται από τους όρους χρήσης αυτού. Η χρήση στον παρόντα οδηγό γίνεται για δωρεάν εκπαιδευτικούς μη εμπορικούς σκοπούς.