

Όμιλος Προβληματισμού για τον Εκσυγχρονισμό της Κοινωνίας μας (ΟΠΕΚ) - CYENS Centre of Excellence
Λευκωσία - 12 Φεβρουαρίου 2024

Τεχνητή νοημοσύνη στη σύγχρονη ζωή
Η επίδραση στην κοινωνία και τα ΜΜΕ

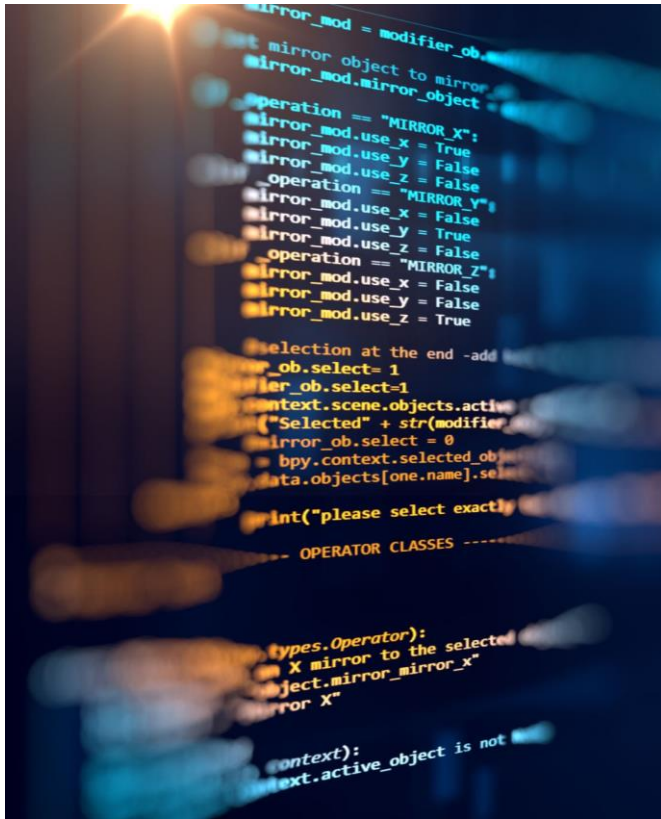
Μερικές σκέψεις για την πληροφόρηση, τα δικαιώματα και τη δημοκρατία

ΛΙΛΙΑΝ ΜΗΤΡΟΥ

Καθηγήτρια Πανεπιστημίου Αιγαίου - Δικηγόρος Αθηνών

Πρόεδρος Ινστιτούτου για το Δίκαιο της Ιδιωτικότητας, τα Προσωπικά Δεδομένα και την Τεχνολογία

Τεχνητή νοημοσύνη ως «ρουτίνα»;



- ▶ Βασικός πυλώνας και μοχλός της 4^{ης} Βιομηχανικής Επανάστασης σε διάδραση με τα Δεδομένα Μεγάλης Κλίμακας (Big Data) και το Διαδίκτυο των Πραγμάτων (Internet of Things).
- ▶ Η σχετική έρευνα δεν οδηγείται πλέον πρωτίστως από την επιστημονική περιέργεια αλλά από οικονομικές και κοινωνικές απαιτήσεις.
- ▶ Η διαθεσιμότητα και επεξεργασία ασύλληπτης ποσότητας δεδομένων διανοίγει νέες τεχνολογικές δυνατότητες και προοπτικές
- ▶ Αγγίζει και αφορά πλέον τον «μέσο χρήστη»: μέσω της προσβασιμότητας σε φθηνή, τεράστια υπολογιστική ισχύ και σύνδεση οι άνθρωποι χρησιμοποιούν πλέον εφαρμογές τεχνητής νοημοσύνης στην καθημερινότητά τους (personal assistants, facial and pattern recognition).
- ▶ Η εφαρμογή ChatGPT αποδείχθηκε η ταχύτερα διαδοθείσα εφαρμογή σε επίπεδο καταναλωτών

Η Τεχνητή Νοημοσύνη ως διασπαστική/ανατρεπτική τεχνολογία (Disruptive technology)

- ▶ Αντιλαμβάνεται ολιστικά το περιβάλλον και προσαρμόζει τις ενέργειες, κατά τρόπο ώστε να μεγιστοποιεί τις πιθανότητες επιτυχίας-
- ▶ Δεν είναι απλώς μία «τεχνολογία γενικού σκοπού» που - όπως ο ατμός ή ο ηλεκτρισμός - αλλά μια τεχνολογία **Gatekeeper** - μία τεχνολογία πυλωρός
- ▶ Τεχνολογία/ καινοτομία που διαφοροποιεί τις αγορές και την οικονομία κατά τρόπο απροσδόκητο
- ▶ Ανατρεπτικές τεχνολογίες; Κοινωνικά ανατρεπτικές τεχνολογίες;
- ▶ Χαρακτηριστικό/ συστατικό της 4ης Βιομηχανικής Επανάστασης και των μεταμορφώσεων/ «αναστατώσεων» που επιφέρει
- ▶ Τα στάδια της ...αναστάτωσης:
 - ▶ Σύγχυση / παράλυση.....που προκαλεί η αναπάντεχη αλλαγή
 - ▶ Υποβάθμιση της σημασίας
 - ▶ Αποδοχή / συνειδητοποίηση των ζητημάτων και των απειλών
 - ▶ Προσαρμογή : μεταμόρφωση - ανάληψη πρωτοβουλιών

Ηθικές/Δεοντολογικές αρχές

Σεβασμός της ανθρώπινης αυτονομίας: πλήρης και αποτελεσματική αυτοδιάθεση και να είναι σε θέση να συμμετέχουν στη δημοκρατική διαδικασία

Πρόληψη βλάβης: Τα συστήματα ΤΝ δεν θα πρέπει ούτε να προκαλούν ούτε να επιδεινώνουν τυχόν βλάβες ή να επηρεάζουν με άλλο τρόπο τον άνθρωπο αρνητικά

Δικαιοσύνη - Η ανάπτυξη, η εγκατάσταση και η χρήση συστημάτων ΤΝ θα πρέπει να γίνεται με τρόπο δίκαιο

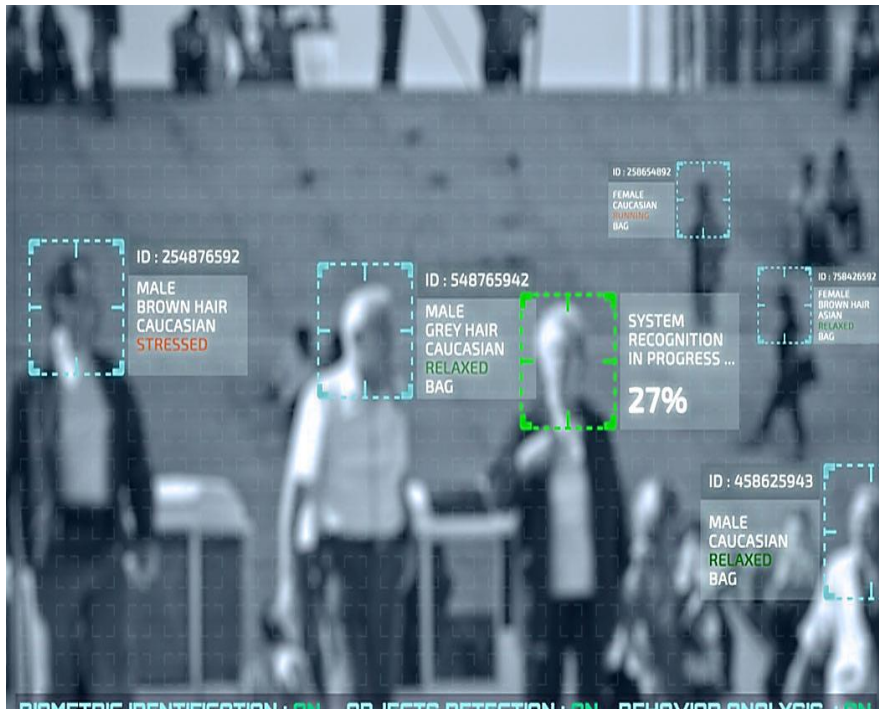
Επεξηγησιμότητα - οι διαδικασίες πρέπει να είναι διαφανείς, οι δυνατότητες και ο σκοπός των συστημάτων ΤΝ να κοινοποιούνται ανοικτά και οι αποφάσεις, στο μέτρο του δυνατού, να επεξηγούνται στους άμεσα και έμμεσα επηρεαζόμενους



Ηθικές και κανονιστικές αρχές και απαιτήσεις

- ▶ Εμπλοκή του ανθρώπινου παράγοντα/ Ανθρώπινη παρέμβαση και εποπτεία
- ▶ Τεχνική στιβαρότητα και ασφάλεια
- ▶ Προστασία δεδομένων και διαχείριση της πληροφορίας
- ▶ Διαφάνεια
- ▶ Πολυμορφία, απαγόρευση διακρίσεων και θεμιτός χαρακτήρας (fairness)
- ▶ Κοινωνική και περιβαλλοντική ευημερία
- ▶ Λογοδοσία

Τεχνητή νοημοσύνη και δικαιώματα



- Ανθρώπινη αξιοπρέπεια, ελευθερία του ατόμου, **σεβασμός για τη Δημοκρατία, τη Δικαιοσύνη και το Κράτος Δικαίου**, ισότητα, αποφυγή διακρίσεων, δικαιώματα μειονοτήτων και δικαιώματα πολιτών
- **Σεβασμός της ιδιωτικής και οικογενειακής ζωής** και προστασία δεδομένων προσωπικού χαρακτήρα (άρθρα 7 και 8 ΧΘΔΕΕ)
- **Ελευθερία έκφρασης και πληροφόρησης** (άρθρο 11)
- **Απαγόρευση διακρίσεων** (άρθρο 21) και ισότητα γυναικών και ανδρών (άρθρο 23)
- **Δικαίωμα πραγματικής προσφυγής και αμερόληπτου δικαστηρίου** (άρθρο 47)
- **Δικαίωμα χρηστής διοίκησης** (άρθρο 41)
- **Τεκμήριο αθωότητας** και δικαιώματα της υπεράσπισης (άρθρο 48)

Μπορεί ο αλγόριθμος να διοικεί; Ζητήματα και απαιτήσεις



Είναι εφικτή και δικαιικά αποδεκτή η «ανάθεση» αυτής της διαδικασίας σε μία εφαρμογή ΤΝ και εάν ναι, σε ποια έκταση και υπό ποιες προϋποθέσεις και εγγυήσεις;

- ▶ Ακρίβεια
- ▶ Αξιοπιστία
- ▶ Ορθότητα της κρίσης
- ▶ Δίκαιη και ισότιμη μεταχείριση
- ▶ Αιτιολογία
- ▶ Λογοδοσία
- ▶ Ευθύνη
- ▶ Διαχειρισιμότητα, κοινωνική αποδοχή και πολιτική νομιμοποίηση της «απόφασης δια της ΤΝ»

Τεχνητή νοημοσύνη και κράτος δικαίου

- ▶ Απαίτηση για προβλεψιμότητα (foreseeability) και η συνεκτικότητα της κυβερνητικής/ διοικητικής δράσης: τρόποι προετοιμασίας, παραγωγής και λήψης μίας απόφασης που είναι γενικεύσιμοι, οδηγούν σε αποτελέσματα γενικεύσιμα και επαναλήψιμα
- ▶ (Α)διαφάνεια ως προς την κρίση και την παραγωγή του αποτελέσματος
- ▶ Αλγοριθμική γραφειοκρατία: προβλήματα προκαταλήψεων, της ενίσχυσης «ιστορικών» διακρίσεων ή παγίωσης υφιστάμενων προβληματικών πρακτικών.
- ▶ Συστήματα τεχνητής νοημοσύνης που παράγουν αποφάσεις βασισμένα στη δική «τους» αξιακή εκτίμηση και στα δικά «τους» ερμηνευτικά σχήματα μπορεί να βρεθούν ακριβώς στον αντίποδα του κράτους δικαίου

Τεχνητή νοημοσύνη και ενημέρωση

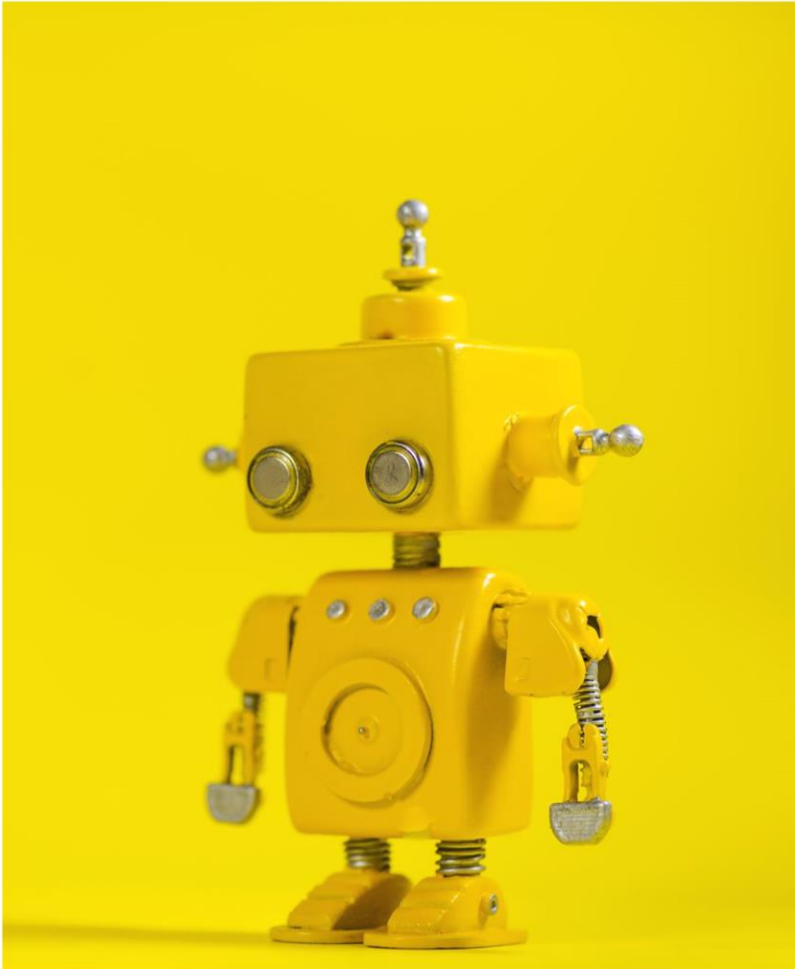
- ▶ Η ΤΝ διαδραματίζει κομβικό ρόλο σε σχέση με την ελευθερία της έκφρασης
- ▶ Εφαρμόζεται σε μια σειρά καταστάσεων που επηρεάζουν τον τρόπο και την έκταση με τα οποία οι άνθρωποι μπορούν να αποκτήσουν πρόσβαση, να αναζητήσουν και να βρουν πληροφορία σε ψηφιακό περιβάλλον (αλγόριθμοι αναζήτησης/ προβολής αποτελεσμάτων).
- ▶ Όλο και περισσότερο πάροχοι ψηφιακών υπηρεσιών/ online intermediaries, όπως πλατφόρμες ψηφιακών δικτύων και μηχανές αναζήτησης χρησιμοποιούν συστήματα ΤΝ για να ελέγξουν το περιεχόμενο, την πληροφορία, η οποία θα προσφερθεί ή στην οποία θα εκτεθούν οι χρήστες με τρόπους που είναι αδιαφανείς και μη υποκείμενοι σε έλεγχο
- ▶ Υφίσταται ο κίνδυνος ένας περιορισμένος αριθμός ψηφιακών ενημερωτικών ή ακόμη και μη ενημερωτικών οργανισμών να καταστούν ο βασικός αγωγός για το ψηφιακό περιεχόμενο, για το περιεχόμενο που προσφέρεται και διαδίδεται ψηφιακά.



Τι είναι η δημιουργική/ παραγωγική Τεχνητή Νοημοσύνη;

- ▶ Τέτοια συστήματα είναι σχεδιασμένα για να απαντούν σε ερωτήσεις, με έναν τρόπο φυσικό, διαλεκτικό, διαδραστικό, να παράγουν λόγο και όχι απλώς να προτείνουν πηγές άντλησης πληροφορίας.
- ▶ Δεν δίνουν απλώς αποτελέσματα αναζήτησης όπως η Google και πρόσβαση στο περιεχόμενο.
- ▶ Συγκεντρώνουν, επεξεργάζονται, συνθέτουν δεδομένα και δημιουργούν απαντήσεις, συσχετίζοντας μοτίβα και πληροφορίες.
- ▶ Παράγουν αποτελέσματα (κυρίως κείμενο) που προσιδιάζει σε ανθρώπινο δημιούργημα και να επικοινωνεί με τον χρήστη
- ▶ Εκπαιδεύονται διαρκώς με βάση τεράστια και εμπλουτιζόμενα σύνολα δεδομένων κειμένου, εικόνας κλπ.

“Το τζίνι βγήκε από το μπουκάλι”;



- ▶ Το ChatGPT και άλλα μοντέλα γνώρισαν ταχύτατα τεράστια διάδοση, προκάλεσαν ενθουσιασμό και προσδοκίες σε πολλούς χρήστες αλλά και σήμαναν συναγερμό για τις επιπτώσεις της χρήσης του
- ▶ Είναι το ChatGPT, το Bard, το Bing απλά ευφραδή και έξυπνα μηχανήματα ή εργαλεία παραπληροφόρησης και χειραγώγησης;
- ▶ Μπορούν να καταστήσουν δυνατή την χειραγώγηση εκλογών αλλά και την διαταραχή δημοκρατικών διαδικασιών, να προκαλέσουν δυσπιστία και προκαταλήψεις;
- ▶ Διασπορά δυσπιστίας και πρόκληση κοινωνικών εντάσεων, με σοβαρές δυνητικές επιπτώσεις για την ασφάλειά (υβριδικές απειλές) ;

(Αν)ακρίβεια και(αν)αξιοπιστία

- ▶ Ζητήματα ως προς την αξιοπιστία και την ακρίβεια των αποτελεσμάτων, ειδικά αναφορικά με την ικανότητα των χρηστών που είναι ειδικοί να παράξουν **deepfakes**, ήχο και βίντεο.
- ▶ Προβλήματα ως την ακρίβεια και την αξιοπιστία λόγω της ποιότητας των δεδομένων που χρησιμοποιούνται για την εκπαίδευση των συστημάτων αυτών και εν γένει των (αν)αξιόπιστων πηγών εκπαίδευσης ή των προκατειλημμένων δεδομένων
- ▶ Η «δηλητηρίαση» των δεδομένων ή επιθέσεις **backdoor** μπορεί να οδηγήσουν σε λανθασμένη πληροφόρηση και αποτελέσματα.

Χειραγώγηση;



- ▶ Πάροχοι, μηχανές αναζήτησης, πλατφόρμες και εταιρίες κοινωνικών δικτύων έχουν καταστεί ισχυροί νέοι παίκτες, ρυθμιστές.
- ▶ Gatekeepers («πυλωροί») συμμετέχοντας στο περιεχόμενο, στην δόμηση, τη διάρθρωση της πληροφορίας και της γνώσης σε κοινωνικό και ατομικό επίπεδο.
- ▶ Σε συνδυασμό με τον έλεγχο του περιεχομένου υφίσταται και ο κίνδυνος της **υπεραντιπροσώπησης ορισμένων απόψεων**
- ▶ Αντίστοιχος είναι ο κίνδυνος της **προπαγάνδας τοξικών στερεότυπων, μισαλλόδοξων αντιλήψεων ή ιδεολογιών** που υιοθετούν διακρίσεις ή την περιθωριοποίηση κοινωνικών ομάδων

Τι είναι η παραπληροφόρηση

- ▶ επαληθεύσιμα ψευδής ή παραπλανητική πληροφορία
- ▶ που δημιουργείται, παρουσιάζεται και διαδίδεται
- ▶ για οικονομικό όφελος ή για την εσκεμμένη εξαπάτηση του κοινού
- ▶ και μπορεί να προκαλέσει δημόσια ζημία.
- ▶ Η δημόσια ζημία περιλαμβάνει τις απειλές για τις δημοκρατικές πολιτικές και τις διαδικασίες χάραξης πολιτικής, καθώς και για τα δημόσια αγαθά, όπως η προστασία της υγείας των πολιτών, του περιβάλλοντος ή της ασφάλειας.
- ▶ Η παραπληροφόρηση δεν περιλαμβάνει την αναφορά λαθών, τη σάτιρα και την παρωδία ή τις σαφώς προσδιορισμένες κομματικές απόψεις και τον σχολιασμό.





Φθηνή, ταχεία, αποτελεσματική (παρα)πληροφόρηση;

- ▶ Η παραπληροφόρηση γίνεται φθηνότερη και πιο αποτελεσματική.
- ▶ η παραπληροφόρηση δεν γεννιέται με τη δημιουργική τεχνητή νοημοσύνη αλλά η διαδικτυακή διάδοση παραπληροφόρησης ενισχύεται σε όγκο και εμβέλεια
- ▶ Οι άνθρωποι τείνουν να πιστέψουν περισσότερο μια πληροφορία όταν την βλέπουν να αναρτάται και να διαδίδεται από περισσότερους
- ▶ Το ποσοτικό ζήτημα εξελίσσεται σε ποιοτικό πρόβλημα

Πειστική και αληθοφανής (παρα)πληροφόρηση;

- ▶ Συστήματα (υπό μία έννοια) προδιαγεγραμμένα να παρέχουν πληροφορία που εκλαμβάνεται ως αξιόπιστη και έγκυρη
- ▶ Η δημιουργική ΤΝ ενισχύει την (υπερβολική) εξάρτηση από τα αποτελέσματα μιας αλγοριθμικής διαδικασίας
- ▶ Οι άνθρωποι τείνουν να επιδεικνύουν μεγαλύτερο βαθμό εμπιστοσύνης σε αποτελέσματα που παράγει η ΤΝ, ιδίως εάν έχει ένα χαρακτήρα επίσημο, αυθεντίας ή εάν οι άνθρωποι βρίσκονται σε ευάλωτη κατάσταση
- ▶ Η δύναμη «πειθούς» ενισχύεται ακόμη περισσότερο από τον «ανθρωπομορφισμό» που χαρακτηρίζει συστήματα όπως το ChatGPT



Δημιουργική ΤΝ και η διάβρωση της εμπιστοσύνης



- ▶ Μία από τις βασικές συνέπειες είναι η ενίσχυση της απώλειας εμπιστοσύνης, της δυσπιστίας απέναντι στα παραδοσιακά μέσα
- ▶ Η ανάπτυξη της δημιουργικής/ παραγωγικής ΤΝ και τα ποικίλα αποτελέσματά της παρουσιάζει μία ακόμη διάσταση αναφορικά με τη δυσπιστία: ενισχύει την σκόπιμη ή μη παραπληροφόρηση αλλά ταυτόχρονα οι άνθρωποι δεν ξέρουν τί να πιστέψουν.
- ▶ Η παραπληροφόρηση μέσω γλωσσικών μοντέλων μπορεί να ενισχύει την καχυποψία και να υποσκάψει την πίστη και στην επιστήμη

Η ρύθμιση της Τεχνητής Νοημοσύνης

- ▶ “Collingridge dilemma”
 - ▶ κατά τα αρχικά στάδια μίας νέας τεχνολογίας, η νομοθέτηση είναι δυσχερής καθώς δεν υπάρχει επαρκής γνώση/ πληροφόρηση/ συνείδηση των ζητημάτων.
 - ▶ Όταν αναδεικνύονται οι ανεπιθύμητες επιπτώσεις της τεχνολογίας, έχει διεισδύσει τόσο στη ζωή και την οικονομία που κάθε απόπειρα ρυθμιστικού ελέγχου συναντά την αντίδραση των προσώπων που την αναπτύσσουν, τους επενδυτές αλλά και τους ίδιους τους χρήστες
- ▶ **Artificial Intelligence Act:** Η ρύθμιση της ΤΝ συνιστά ένα ζήτημα ψηφιακής (και όχι) μόνο κυριαρχίας

- ▶ Εμείς (μπορούμε) να σχεδιάσουμε κανονιστικά το μέλλον για την Τεχνητή Νοημοσύνη; Ή μήπως η Τεχνητή Νοημοσύνη έχει ήδη σχεδιάσει ή σχεδιάζει, προδιαγράφει το μέλλον;
- ▶ Είναι εφικτή η ρύθμιση της τεχνολογικής εξέλιξης ;
- ▶ Είναι ζήτημα ευρωπαϊκής ρύθμισης τη στιγμή που η Τεχνητή Νοημοσύνη, όπως κάθε τεχνολογία εξελίσσεται και αναπτύσσεται σε παγκόσμιο επίπεδο;
- ▶ Πώς θα επιδράσει η ρύθμιση στην εξέλιξη;
 - ▶ Μήπως θα ήταν ένα «**θεσμικό εμπόδιο**» στην **καινοτομία** που θα άφηνε την Ευρώπη πίσω από τους άλλους παγκόσμιους παίκτες;
 - ▶ Ή μήπως αυτή είναι η **προστιθέμενη αξία**, η **αξιακή διαφορά** της Ευρώπης ;

Ερωτήματα

Οι δημοκρατικές δεσμεύσεις και η ηθική της τεχνητής νοημοσύνης

- ▶ Ποιος είναι ο ρόλος και η ευθύνη των παραγωγών τεχνολογίας και εφαρμογών τεχνητής νοημοσύνης; Μπορεί να διαμορφωθεί μία ηθική της τεχνητής νοημοσύνης (artificial intelligence ethics) κατ' αναλογία και κατ' εξέλιξη της ψηφιακής ηθικής (digital ethics);
- ▶ Ποιες πολιτικές απαιτούνται ώστε αφενός να προσδιοριστούν και να αξιοποιηθούν οι ευκαιρίες της Τεχνητής Νοημοσύνης και αφετέρου να εξασφαλιστεί ότι η ανάπτυξη αυτή θα ενισχύσει την ευημερία της κοινωνίας και όχι ανισότητες ; Ότι θα σεβαστεί τα δικαιώματα των πολιτών και τα έννομα συμφέροντα των καταναλωτών και δεν θα οδηγήσει σε περιορισμούς αυτών;

Σας ευχαριστώ για
την προσοχή σας

