



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΧΗΜΙΚΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΙΙ



# «Η Τεχνητή Νοημοσύνη ως τεχνολογική επανάσταση: Προσδοκίες, Υποσχέσεις και Κίνδυνοι»

Γιάννης Καλογήρου, Ομότιμος Καθηγητής ΕΜΠ

---

Ενότητα 1<sup>η</sup>: **Τεχνητή Νοημοσύνη: Quo Vadis?** Τεχνολογικές, Οικονομικές, Κοινωνικές, Θεσμικές και Πολιτικές Επιπτώσεις  
**18ο Ετήσιο Σεμινάριο της Ερμούπολης για την Κοινωνία της Πληροφορίας και την Οικονομία της Γνώσης**

Συνεδριακή Αίθουσα Επιμελητηρίου Κυκλάδων, Ερμούπολη,

14 Ιουλίου 2023

# Ταξινόμηση των Τεχνολογικών Επανάστασεων (ΤΕ) στο σύγχρονο κόσμο (1770-σήμερα)

- Ως ΤΕ (Carlota Perez, 2009 & Chris Freeman) ορίζεται **ένα σύνολο ριζικών τεχνολογικών εξελίξεων** (ανακαλύψεις και τεχνικές) που διαμορφώνουν έναν σημαντικό **αστερισμό αλληλοεξαρτώμενων τεχνολογιών**, μια συστάδα (Cluster of Clusters) συστάδων ή ένα σύστημα συστημάτων (System of Systems).
- Το **πλέγμα** των τεχνολογιών που συνδέονται με την Τεχνητή Νοημοσύνη algorithms (TN): Μηχανική Μάθηση, Internet of Things, Robotics, Cloud computing, quantum computing, 3D printing, Big Data Analytics κ.α. επέτρεψαν τη σταδιακή μετατροπή της σε **τεχνολογία γενικού σκοπού με μετασχηματιστική ικανότητα και δυναμική που σταδιακά επηρεάζει όλους τους τομείς και κλάδους της οικονομίας και εν τέλει και της κοινωνίας.**

# Που τοποθετείται η εποχή της Τεχνητής Νοημοσύνης στο πλαίσιο της ταξινόμησης των Τεχνολογικών Επαναστάσεων (ΤΕ) στο σύγχρονο κόσμο (1770-σήμερα)

- Που τοποθετείται η εποχή της Τεχνητής Νοημοσύνης στο πλαίσιο των πέντε ή μάλλον έξι διαδοχικών τεχνολογικών επαναστάσεων (ΤΕ) που εμφανίστηκαν από τη δεκαετία του 1770 έως τη δεύτερη δεκαετία του 21<sup>ου</sup> αιώνα.
- Η προηγούμενη Πέμπτη τεχνολογική επανάσταση των Τεχνολογιών της Πληροφορικής και των Επικοινωνιών [που βασίστηκε στην ανάπτυξη και τη γρήγορη διάχυση της μικροηλεκτρονικής και των υπολογιστών] διαμόρφωσε ένα τεχνολογικό σύστημα γύρω από τους μικροεπεξεργαστές και άλλες κατηγορίες ολοκληρωμένων κυκλωμάτων ημιαγωγών. Η μικροηλεκτρονική ως το πιο χαρακτηριστικό παράδειγμα τεχνολογίας γενικού σκοπού που επηρεάζει όλους τους τομείς και κλάδους της οικονομίας και εν τέλει και της κοινωνίας.

# Η προσέγγιση της ΤΝ από τη σκοπιά των εποχών της μηχανής (1<sup>η</sup> και 2<sup>η</sup> εποχή των μηχανών που έχει ήδη ξεκινήσει)

- Η τεχνολογία καθαυτή δεν είναι ούτε καλή ούτε κακή. Είναι η **χρήση της τεχνολογίας από τους ανθρώπους** που προσδιορίζει τόσο τη **φύση** της όσο και την **έκταση** των ωφελειών ή/ και των ζημιών» (Chris Freeman, 1997).
  - Στην πρώτη τεχνολογική επανάσταση η **εκμηχάνιση** και η **αξιοποίηση της υδραυλικής ενέργειας** υποκατέστησαν τη **μυϊκή δύναμη** των ζώων και των ανθρώπων και καθιέρωσαν την **άψυχη ενέργεια**.
  - Στη συνέχεια οι **μηχανές υποκατέστησαν επαναλαμβανόμενες διαδικασίες και δραστηριότητες ρουτίνας σχετικά απλές και αλγοριθμικά προγραμματισμένες**.
- Το μεγάλο διακύβευμα είναι η **σχέση ανθρώπου-μηχανής**. Υποκατάσταση ή **Συμπληρωματικότητα**. Η πρώτη εποχή της μηχανής vs δεύτερη εποχή της μηχανής. Ανθρώπινη Κρίση (judgement) και Ενσυναίσθηση μπορεί να σώσουν την παρτίδα.

# Η ανάπτυξη του διανοητικού μονοπωλιακού καπιταλισμούτε (“Intellectual Monopoly Capitalism”)

- Πέντε εταιρείες στις ΗΠΑ και άλλες πέντε στην Κίνα (συμπεριλαμβανομένων και των νεοφυών επιχειρήσεων που έχουν δημιουργηθεί από αυτές, ιδίως στις ΗΠΑ).
- Big Tech race: Οι τεχνολογικοί γίγαντες επιδιώκουν να καλύψουν το χαμένο έδαφος στην κούρσα για την εμπορική αξιοποίηση της τεχνολογίας της γεννητικής (παραγωγικής) Τεχνητή Νοημοσύνης. Η Google άργησε να κυκλοφορήσει το δικό της μοντέλο Chatbot σε σύγκριση με την ανταγωνίστριά της, την Microsoft, η οποία επένδυσε σημαντικά ποσά στην Open AI και η οποία προτίθεται να επενδύσει άλλα 10 δισεκ. στην OpenAI. Στο πλαίσιο αυτό η Google λανσάρισε το Bard chatbot σε μια προσπάθεια να ανταγωνιστεί το δημοφιλές ChatGPT της OpenAI. Ακολουθούν και οι άλλες και βέβαια ο Elon Musk.

# Ο αγώνας δρόμου των μεγάλων χωρών

---

- ΗΠΑ vs Κίνα.
- Που είναι η Ευρώπη;

# Τα μεγάλα διακυβεύματα

1. **Η κοινωνία της επιτήρησης** (AI και αυταρχικά καθεστώτα). Human rights risks of **Facial Recognition AI tech**. **People free from unwarranted government surveillance.**
2. Ο **φυσικός κόσμος** και ο κόσμος των **Κυβερνο-φυσικών Συστημάτων** Cyber Physical Systems (Virtual +Augmented Reality): Η ανάπτυξη και η διαχείρισή τους.
3. **Τρίτο διακύβευμα:** Ο φόβος της απώλειας θέσεων εργασίας και οι επιπτώσεις της νέας οργάνωσης της εργασίας ( Augmentation/ human capabilities vs Automation)
4. **Ίσως** το πιο άμεσο ζήτημα για αντιμετώπιση: Σε ποιο βαθμό η ακραία συγκέντρωση πλούτου και οικονομικής δύναμης που συνδέεται με ψηφιακές καινοτομίες (αρκετοί ερευνητές μιλάνε για έναν **διανοητικό μονοπωλιακό καπιταλισμό**) **μπορεί τελικά να οδηγήσει σε μια παρόμοια συγκέντρωση πολιτικής ισχύος, που τελικώς μπορεί να υπονομεύσει τη δημοκρατία και τους θεσμούς της.**
5. Οι επιπτώσεις στη δημόσια διοίκηση (αποτελεσματικότητα και αποδοτικότητα στην προσφορά δημόσιων υπηρεσιών) καθώς και στην υγεία και την εκπαίδευση (Μεγάλες ευκαιρίες, αλλά και μεγάλοι κίνδυνοι).
6. Η παραγωγικότητα της οικονομίας και η ενδυνάμωση ικανοτήτων.

- **Δεξιότητες και πόροι** για τον εντοπισμό και την εκμετάλλευση ανοιχτών παραθύρων πολιτικής.
- γ) ΤΝ και διανομή πληροφοριών: προκαλώντας τους πολίτες. Η πρόσβαση σε σχετικά και **αξιόπιστα δεδομένα και πληροφορίες**.
- **δ) Το πρόβλημα της εμπιστοσύνης και της επιτήρησης. (ΤΝ σε δημόσιες και ιδιωτικές μορφές επιτήρησης (public and private forms of surveillance: Challenging trust in the citizen–government relations. Η κοινωνία της επιτήρησης.**
  1. AI and the persuasion industry: Eroding the policy entrepreneurial resources and skills of citizens.
  2. AI and the weaponization of information: Hybrid threats against trust between citizens and democratic institutions.
  3. AI and civic tech: **Engaging citizens in decision-making processes but not without risks.**
  4. Η πρωτοπορία της **Διαύγειας**, ειδικά για την ελληνική πραγματικότητα, ήταν πως για πρώτη φορά δημιουργήθηκε μια διαδικτυακή πλατφόρμα όπου έπρεπε να αναρτώνται όλες οι πράξεις της Δημόσιας Διοίκησης σε μηχαναγνώσιμη μορφή (σ.σ. σε μορφή που επιτρέπει την άμεση ανάγνωση και επεξεργασία τους από υπολογιστές) – προκειμένου να θεωρούνται έγκυρες και εκτελεστές. **Την τελευταία δεκαετία, η Ελλάδα, από πρωτοπόρα χώρα στα ανοικτά δεδομένα, κατέληξε στην κατώτερη βαθμίδα της Ευρώπης**, εμποδίζοντας τόσο την οικονομική ανάπτυξη του ιδιωτικού τομέα όσο και **την άσκηση ελέγχου των εξουσιών, τη διαφάνεια και τη λογοδοσία**. Η φιλόδοξη προσπάθεια που ξεκίνησε στις αρχές της δεκαετίας με τη «Διαύγεια» έμεινε στη μέση και φαίνεται να έχει εγκαταλειφθεί. Κι όμως, **τα ανοικτά δημόσια δεδομένα, ως υποσύνολο των μεγάλων δεδομένων (big data), αποτελούν βασικό πυλώνα του ψηφιακού μετασχηματισμού κράτους και οικονομίας**.
- **An alternative scenario is more regulation to restrict computational propaganda.** Stricter rules on **data protection** and **algorithmic accountability** and **transparency** could also reduce the extent to which machine learning can be abused in political contexts.



- Τα πολλές φορές
- Τα **αυτοματοποιημένα** συστήματα λήψης αποφάσεων που **λειτουργούν με αλγόριθμους** ενέχουν **εξίσου μεγάλες δυνατότητες πρόκλησης ζημιών.**
- **Μεροληψία και ανάγκη για διαφάνεια.** Οι σχεδιαστές που βρίσκονται πίσω από την ανάλυση βάσει δεδομένων αποφασίζουν **ποια δεδομένα θα συλλέξουν και θα συμπεριλάβουν και ποια κριτήρια είναι σχετικά με τη διαδικασία.** Παρόλο που αυτή η **εξελιγμένη στατιστική ανάλυση** αποτελεί **πυλώνα της κοινωνίας του 21ου αιώνα**, οι **τεχνικές διαδικασίες** πίσω από τις αποφάσεις **δεν είναι διαφανείς για τους χρήστες ή και τις ρυθμιστικές αρχές.**
- Το ζήτημα του **τρόπου ρύθμισης της αλγοριθμικής λήψης αποφάσεων και των συναφών πρακτικών δεδομένων** έχει επιπτώσεις σε όλες τις πτυχές της ζωής μας, συμπεριλαμβανομένων των οικονομικών ευκαιριών, της ευημερίας και της ελευθερίας του λόγου.
- Σήμερα και στο κοντινό (?) μέλλον οι τεχνολογίες ΤΝ **μπορούν και να υποκαταστήσουν τους ανθρώπους** και να **αμφισβητήσουν την κυριαρχία του ανθρώπινου είδους στον πλανήτη ή και πάλι να δράσουν συμπληρωματικά.**

# Οι τρεις μεγάλοι εμπειρικοί κανόνες σχετικά με την διαχείριση της τεχνολογικής εξέλιξης (Center for Humane Technology, USA)

1. **Όταν εφευρίσκεται μια νέα ριζική τεχνολογία αποκαλύπτεται μια νέα κατηγορία ευθυνών.** Στην αρχή δεν είναι αυτονόητο ποιες είναι οι συγκεκριμένες νέες ευθύνες, ενώ η νομοθετική ρύθμισή τους ακολουθεί μετά από τη διαπίστωση κάποιου γεγονότος που έχει σημαντικές επιπτώσεις που . Π.χ. Ο κίνδυνος για την ιδιωτικότητα έγινε ορατός όταν οι κάμερες παρακολούθησης βγήκαν μαζικά στην αγορά.
2. **Αν η συγκεκριμένη τεχνολογία προσδίδει δύναμη, θα ξεκινήσει ένας αγώνας δρόμου. Οι εταιρείες της Τεχνητής Νοημοσύνης μπαίνουν σε έναν αγώνα δρόμου για να την αναπτύξουν/ αξιοποιήσουν όσο το δυνατόν γρηγορότερα χωρίς επαρκή μέτρα ασφαλείας.**
3. **Αν δεν συντονιστεί η διαχείριση της προαναφερόμενης εξέλιξης (π.χ. με την αναβάθμιση θεσμών σε έναν κόσμο μετά την ΤΝ), τότε ο αγώνας δρόμου μπορεί να καταλήξει σε τραγωδία.**

# Τι πιστεύουν οι ερευνητές της Τεχνητής Νοημοσύνης στις ΗΠΑ (Η έρευνα του Κέντρου για Ανθρώπινη Τεχνολογία, πριν την κυκλοφορία του ChatGPT, 2022)

- Το 50% των ερευνητών ΤΝ πιστεύει ότι υπάρχει 10% ή μεγαλύτερη πιθανότητα να εξαφανιστούν οι άνθρωποι από την αδυναμία μας να ελέγξουμε την ΤΝ.[Υπαρξιακός Κίνδυνος για το ανθρώπινο είδος].
- Σκεφθείτε αν το 50% των μηχανικών αεροπλάνων πίστευαν ό,τι μια πιθανότητα 10% να πεθάνει κάποιος, θα μπαίνατε στο αεροπλάνο;
- Οι υποστηρικτές της άποψης ό,τι υπάρχει υπαρξιακή απειλή, δεν παραγνωρίζουν τα τεράστια και απίστευτα οφέλη που προκύπτουν και μπορούν να προέλθουν και στο μέλλον από την εξάπλωση της Τεχνητής Νοημοσύνης. Υποστηρίζουν ότι η περιγραφή των ωφελειών και των σχετικών θετικών υποσχέσεων για την ανθρωπότητα είναι υπερβολικές και κυρίως ό,τι ελλοχεύει ο κίνδυνος η τεχνολογία να περάσει σε κακά χέρια.
- Η τεχνολογία καθαυτή δεν είναι ούτε καλή ούτε κακή. Είναι **η χρήση της τεχνολογίας από τους ανθρώπους** που προσδιορίζει τόσο τη **φύση** της όσο και την **έκταση** των ωφελειών ή/ και των ζημιών» (Chris Freeman, 1997).