Εισαγωγή στην Τεχνητή Νοημοσύνη

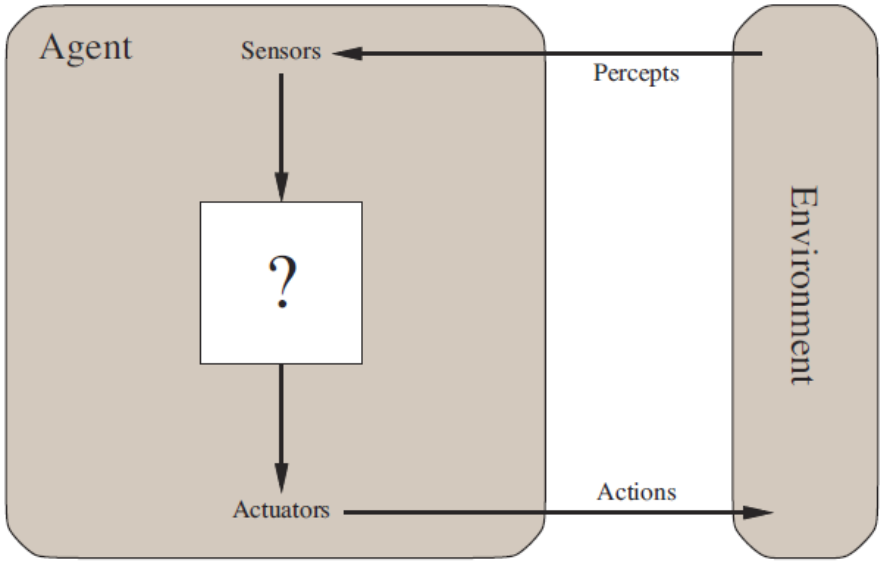
Μανόλης Κουμπαράκης

Τι είναι Τεχνητή Νοημοσύνη?

- **Τεχνητή Νοημοσύνη - ΤΝ (Artificial Intelligence - AI)** είναι η επιστήμη που έχει στόχο την ανάπτυξη **ευφυών πρακτόρων (intelligent agents).**

Πράκτορες

- **Πράκτορας (agent)** είναι οτιδήποτε μπορεί να θεωρηθεί ότι αντιλαμβάνεται το περιβάλλον του μέσω αισθητήρων (sensors) και αλληλεπιδρά με αυτό μέσω μηχανισμών δράσης (effectors - actuators).



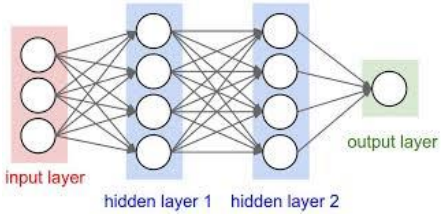
Μηχανική Μάθηση

- **Μηχανική Μάθηση** είναι η επιστήμη του σχεδιασμού αλγορίθμων που μαθαίνουν από δεδομένα.
- Η Μηχανική Μάθηση είναι **η σημαντικότερη σήμερα υποπεριοχή της Τεχνητής Νοημοσύνης.**
- Κάθε μέρα χρησιμοποιούμε **εφαρμογές** που έχουν αναπτυχθεί χρησιμοποιώντας τεχνικές Μηχανικής Μάθησης π.χ.:
  - Αναγνώριση ομιλίας
  - Κατανόηση φυσικής γλώσσας
  - Μηχανική μετάφραση
  - Αυτοοδηγούμενα οχήματα

- Είναι δύσκολο να ονομάσουμε ένα σύστημα **ευφυές** αν δεν μαθαίνει.

Μηχανική Μάθηση

- Μια βασική ιδέα στην Επιβλεπόμενη Μηχανική Μάθηση είναι ότι **εκπαιδεύουμε ένα σύστημα σε μεγάλο όγκο δεδομένων** (π.χ., εικόνες, ήχους ή κείμενο) και το σύστημα μετά μπορεί να λύνει σχετικά προβλήματα με επιτυχία (πολλές φορές μάλιστα, καλύτερα από τους ανθρώπους).



Μηχανική Μάθηση για την Επεξεργασίας Φυσικής Γλώσσας

- Στο μάθημα «Τεχνητή Νοημοσύνη ΙΙ (Deep learning for natural language processing)» δίνουμε μια **σύγχρονη εισαγωγή στη Μηχανική Μάθηση** επικεντρωνόμενοι στις τεχνολογίες των **νευρωνικών δικτύων και των μεγάλων γλωσσικών μοντέλων (LLMs).**
- Μελετούμε επίσης εφαρμογές της Μηχανικής Μάθησης στην **Επεξεργασία Φυσικής Γλώσσας** που είναι μια άλλη σημαντική υποπεριοχή της Τεχνητής Νοημοσύνης.
- Σήμερα τα περισσότερα προβλήματα στην περιοχή της Επεξεργασίας Φυσικής Γλώσσας λύνονται με αλγόριθμους Μηχανικής Μάθησης (και συνήθως **βαθιές αρχιτεκτονικές νευρωνικών δικτύων).**

Ιστορία της ΤΝ

- **Κυοφορία (1943-1955)**
  - Μοντέλα τεχνητών νευρώνων - perceptron (McCulloch and Pitts, 1943). Μάθηση Hebb (Hebb, 1949).
  - Το άρθρο του Alan Turing "Computing Machinery and Intelligence" (1950).
  - Snarc: Ο πρώτος υπολογιστής νευρωνικού δικτύου (Minsky and Edmonds, 1951).



  - Ο Minsky έλαβε επίσης το **Turing Award** το 1969.

Ιστορία της ΤΝ

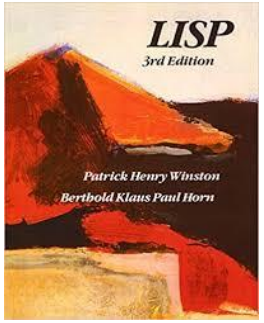


- **Γέννηση (1956)**
  - Το συνέδριο εργασίας στο πανεπιστήμιο Dartmouth το καλοκαίρι του 1956 (McCarthy, Minsky, Newell, Simon).
  - Ο όρος "Artificial Intelligence" υιοθετήθηκε μετά από πρόταση του McCarthy (**Turing Award** winner in 1971).
  - Το πρόγραμμα Logic Theorist (Newell and Simon) έκλεψε την παράσταση. Αμέσως μετά το συνέδριο, το πρόγραμμα αυτό μπόρεσε να αποδείξει τα περισσότερα θεωρήματα του Κεφ. 2 του βιβλίου Principia Mathematica (ένα φημισμένο βιβλίο των Alfred North Whitehead και Bertrand Russell για τη λογική και τα μαθηματικά)!

Ιστορία της ΤΝ



- **Πρώιμος ενθουσιασμός, μεγάλες προσδοκίες (1952-1969)**
  - Logic Theorist, General Problem Solver, Geometry Theorem Prover, game playing, Lisp (from McCarthy), theorem proving, Shakey the robot, micro-worlds, adalines, perceptrons.



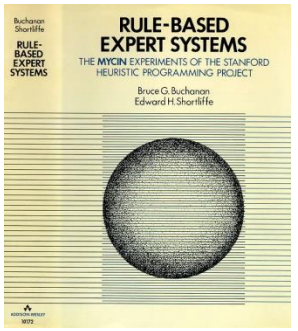
  - Η υπόθεση του **φυσικού συστήματος συμβόλων (physical symbol system hypothesis)** των Newell and Simon: Οποιοδήποτε σύστημα (μηχανή ή άνθρωπος) που εκδηλώνει νοημοσύνη θα πρέπει να λειτουργεί με το χειρισμό δομών δεδομένων που αποτελούνται από σύμβολα. Αυτή είναι η υπόθεση εργασίας για τους ερευνητές της **συμβολικής ΤΝ.**

Ιστορία της ΤΝ

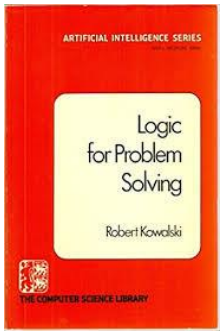


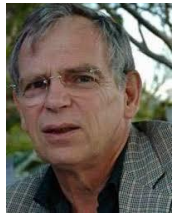
- **Μια δόση ρεαλισμού (1966-1973)**
  - Προγράμματα χωρίς γνώση πεδίου (domain knowledge), προβλήματα δισεπιλυσιμότητας (intractability problems), αρνητικά αποτελέσματα για απλά νευρωνικά δίκτυα.
  - Ακύρωση μεγάλων ερευνητικών έργων σε μηχανική μετάφραση (ΗΠΑ), η έκθεση του Lighthill (Μεγάλη Βρετανία).

Ιστορία της ΤΝ



- **Συστήματα βασισμένα στη γνώση (knowledge-based systems) (1969-1979)**
  - Ο ρόλος της γνώσης πεδίου, τα έμπειρα συστήματα (expert systems) π.χ., το πρόγραμμα MYCIN για την αναγνώριση βακτηρίων που προκαλούν ιώσεις.
  - Γλώσσες αναπαράστασης γνώσης και συλλογιστικής (π.χ., Prolog, frame-based languages).







Ιστορία της ΤΝ



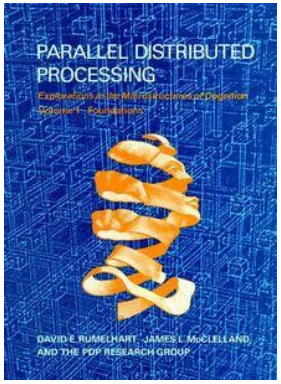
- **Η ΤΝ γίνεται βιομηχανία (1980-σήμερα)**
  - Το πρώτο επιτυχημένο έμπειρο σύστημα: R1 (McDermott, DEC).
  - Η Ιαπωνική δράση “5th Generation project” (1981) και η έμφαση της στο Λογικό Προγραμματισμό.

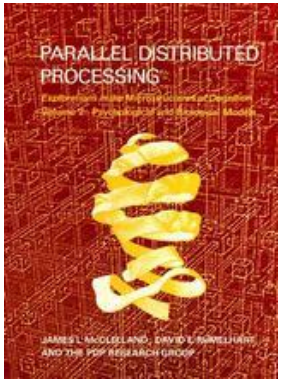


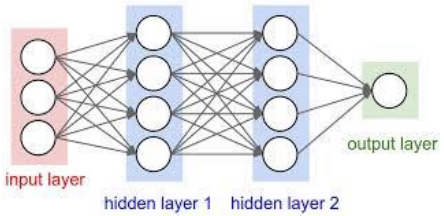
  - Η εταιρία Microelectronics and Computer Technology Corporation (MCC) των ΗΠΑ.
  - Το πρόγραμμα ESPRIT της Ευρωπαϊκής Ένωσης (1983).
  - Η έκθεση Alvey στη Μεγάλη Βρετανία.

Ιστορία της ΤΝ

- **Η επιστροφή των νευρωνικών δικτύων (1986-σήμερα)**
  - Τουλάχιστον τέσσερεις ερευνητικές ομάδες ανακάλυψαν ξανά την μέθοδο μάθησης με **οπισθοδιάδοση (backpropagation)** που είχε προταθεί το 1969.
  - Τα **συνδεσμικά μοντέλα (connectionist models)** που βασίζονται σε νευρωνικά δίκτυα ως συμπλήρωμα των συμβολικών μοντέλων.



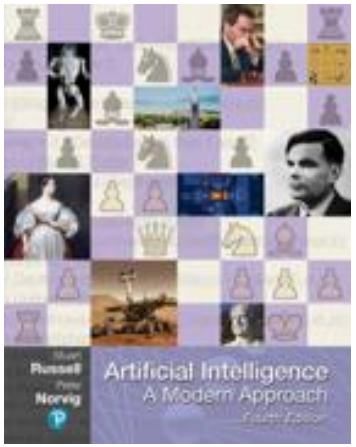




Ιστορία της ΤΝ

- **Η Τεχνητή Νοημοσύνη γίνεται επιστήμη (1987-σήμερα)**
  - Ενδιαφέροντα αποτελέσματα σε αναπαράσταση γνώσης, αναγνώριση ομιλίας, νευρωνικά δίκτυα, εξόρυξη δεδομένων, δίκτυα Bayes, ρομποτική και μηχανική όραση.

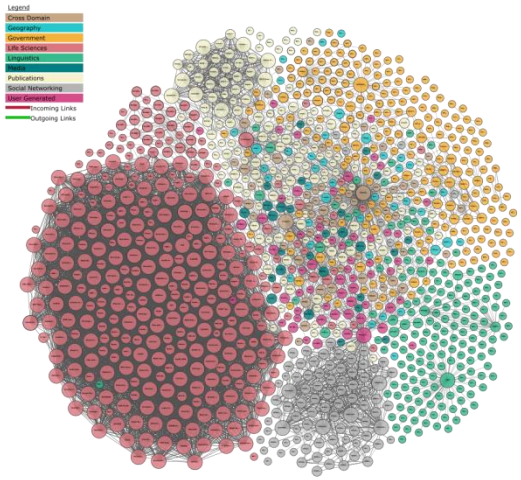
- **Ευφυείς Πράκτορες (1995-σήμερα)**
  - Δείτε το βιβλίο του μαθήματος.

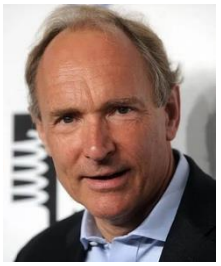


Ιστορία της ΤΝ

- **Σημασιολογικός Ιστός (Semantic Web) και Διασυνδεδεμένα Δεδομένα (Linked Data) (1998-σήμερα)**
  - Δείτε το άρθρο των Tim Berners-Lee, James Hendler και Ora Lassila (2001).
  - Δείτε την ιστοσελίδα του μεταπτυχιακού μαθήματος Τεχνολογίες Γνώσεων και την σχετική ιστοσελίδα του W3C.









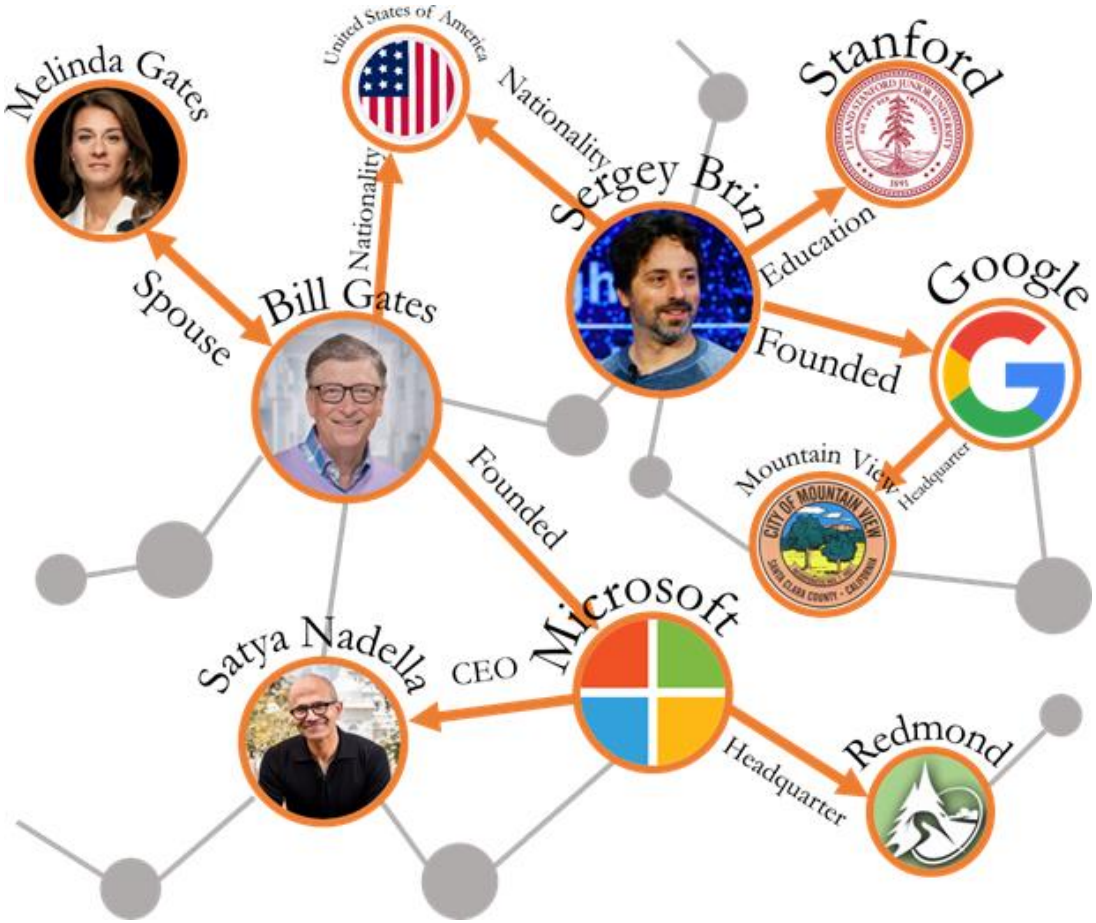


Ιστορία της ΤΝ

- **Χρήση γράφων γνώσεων (knowledge graphs) στις μηχανές αναζήτησης και αλλού (2012-σήμερα)**
  - Δείτε πληροφορίες σχετικά με το Google Knowledge Graph στην ιστοσελίδα: https://support.google.com/knowledgepanel/answer/9787176?hl=en



Παράδειγμα Γράφου Γνώσης



Ιστορία της ΤΝ

- **Λογισμικό επίλυσης προβλημάτων με περιορισμούς (constraint solving software)**
  - Τα προγράμματα επιλυτές (solvers) της Γαλλικής εταιρίας ILOG που τώρα ανήκει στην IBM (https://www.ibm.com/products/ilog-cplex-optimization-studio).



Ιστορία της ΤΝ



- **Παιχνίδια (game playing)**
  - IBM's Deep Blue in 1997 (https://www.ibm.com/history/deep-blue)



  - Google's Deep Mind Alpha Go in 2016 (https://deepmind.google/technologies/alphago/. It uses deep neural networks και tree search.

Ιστορία της ΤΝ

- **Δεδομένα μεγάλης κλίμακας (big data)(2007-σήμερα).**
  - Κάθε δύο μέρες παράγουμε τόσα δεδομένα όσα από την εποχή που δημιουργήθηκε ο ανθρώπινος πολιτισμός μέχρι το 2003 (Eric Schmidt, 2010).
  - Χρήση τεχνικών της ΤΝ και ιδιαίτερα Μηχανικής Μάθησης για **ανάλυση μεγάλων δεδομένων (big data analytics).**



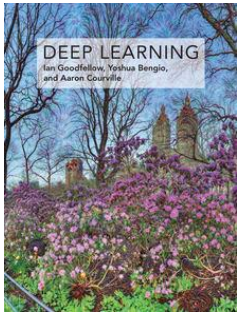
Υποστηρικτικές Τεχνολογίες

- **Ανάπτυξη κλιμακώσιμων υπολογιστικών συστημάτων (clusters and cloud computing, GPUs).**



Ιστορία της ΤΝ

- **Μάθηση σε βάθος (deep learning)(2010-σήμερα)**
  - Δημιουργία μεγάλων συνόλων δεδομένων για την εκπαίδευση νευρωνικών δικτύων (π.χ., Wikipedia, ImageNet, Common Crawl, Pile).
  - Χρήση παράλληλων υπολογιστικών αρχιτεκτονικών και GPUs.
  - Βαθιά νευρωνικά και οι αλγόριθμοι backpropagation και gradient descent για την εκπαίδευση τους.
  - Προγραμματιστικά περιβάλλοντα όπως Keras, TensorFlow, Pytorch και Google Colab.



Ιστορία της ΤΝ

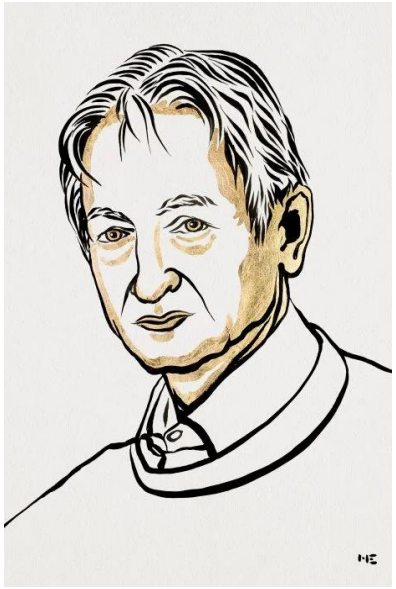
- **Απονομή του Turing Award στους Bengio, Hinton και LeCun το 2019.**



Νόμπελ Φυσικής 2024

- Το Νόμπελ Φυσικής του 2024 απονεμήθηκε στους John Hopfield και Geoff Hinton για τη δουλειά τους στα **Hopfield networks** και τις **Boltzmann machines** αντίστοιχα.





Νόμπελ Χημείας 2024

- Το Νόμπελ Χημείας του 2024 απονεμήθηκε το μισό στον David Baker "for **computational protein design**", και το άλλο μισό από κοινού στους Demis Hassabis και John Jumper "for **protein structure prediction**“.







Turing Award 2024

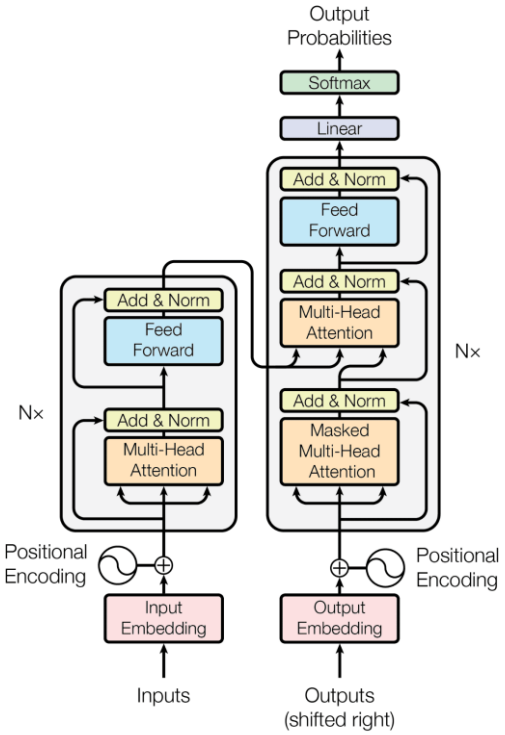
- Το Turing Award του 2024 απονεμήθηκε στους Andrew Barto και Richard S. Sutton “for developing the conceptual and algorithmic foundations of **reinforcement learning (ενισχυτική μάθηση)**”.





Σύγχρονες Τάσεις στη Βαθιά Μηχανική Μάθηση

- **Η αρχιτεκτονική των transformers (2017).**



Google

Σύγχρονες Τάσεις

- **Μεγάλα προεκπαιδευμένα γλωσσικά (ή θεμελιώδη) μοντέλα (LLMs or foundation models)** για την κατανόηση και την παραγωγή φυσικής γλώσσας π.χ., BERT, GPT-2, GPT-3 και άλλα (2018 μέχρι σήμερα).





ChatGPT

- Το Νοέμβριο του 2022 η OpenAI παρουσίασε το **chatbot ChatGPT** που είναι βασισμένο στο InstructGPT.
- Οι ικανότητες του ChatGPT ήταν πραγματικά εντυπωσιακές και το μοντέλο αυτό έλαβε μεγάλη προσοχή από την διεθνή επιστημονική κοινότητα, τα διεθνή μέσα ενημέρωσης και το ευρύτερο κοινό.
- Σήμερα υπάρχουν παρόμοια μοντέλα όπως το **Gemini (Google), Claude (Anthropic), Grok (xAI), Copilot (Microsoft)** και άλλα.

Η Έννοια της Παραίσθησης (Hallucination) στα Μεγάλα Γλωσσικά Μοντέλα

- Όποιος έχει πειραματιστεί με μοντέλα όπως το ChatGPT, θα έχει παρατηρήσει ότι συχνά το μοντέλο αυτό δίνει αληθοφανείς απαντήσεις (π.χ., συντακτικά) αλλά που στην πραγματικότητα είναι λάθος ή δεν βγάζουν καν νόημα. Τότε λέμε ότι το μοντέλο έχει **παραισθήσεις**.
- Το φαινόμενο αυτό ήταν πιο έντονο παλιότερα αλλά τα σημερινά **frontier models** το παρουσιάζουν σε μικρό βαθμό.

Ανοιχτά LLMs

- Εκτός από τα εμπορικά LLMs που αναφέραμε νωρίτερα, υπάρχουν και τα **open-weight** ή **open-source** LLMs.
- Examples:
  - Llama (Meta)
  - Mistral (Mistral.AI)
  - Colibri (Aleph Alpha)
  - DeepSeek (DeepSeek)
  - Olmo (AI2).



- Όλα τα παραπάνω μοντέλα είναι open weight. Μόνο το Olmo όμως είναι πραγματικά open source επειδή όχι μόνο τα βάρη, αλλά και τα δεδομένα εκπαίδευσης και ο κώδικας εκπαίδευσης είναι ελεύθερα διαθέσιμος.

HuggingFace

- Η πλατφόρμα **HuggingFace** είναι σήμερα η πιο δημοφιλής πλατφόρμα για σύνολα δεδομένων και μοντέλα μηχανικής μάθησης (μοντέλο ανοικτού κώδικα). Πρόσφατα αγοράστηκε από την NVIDIA.



Σύγχρονες Τάσεις

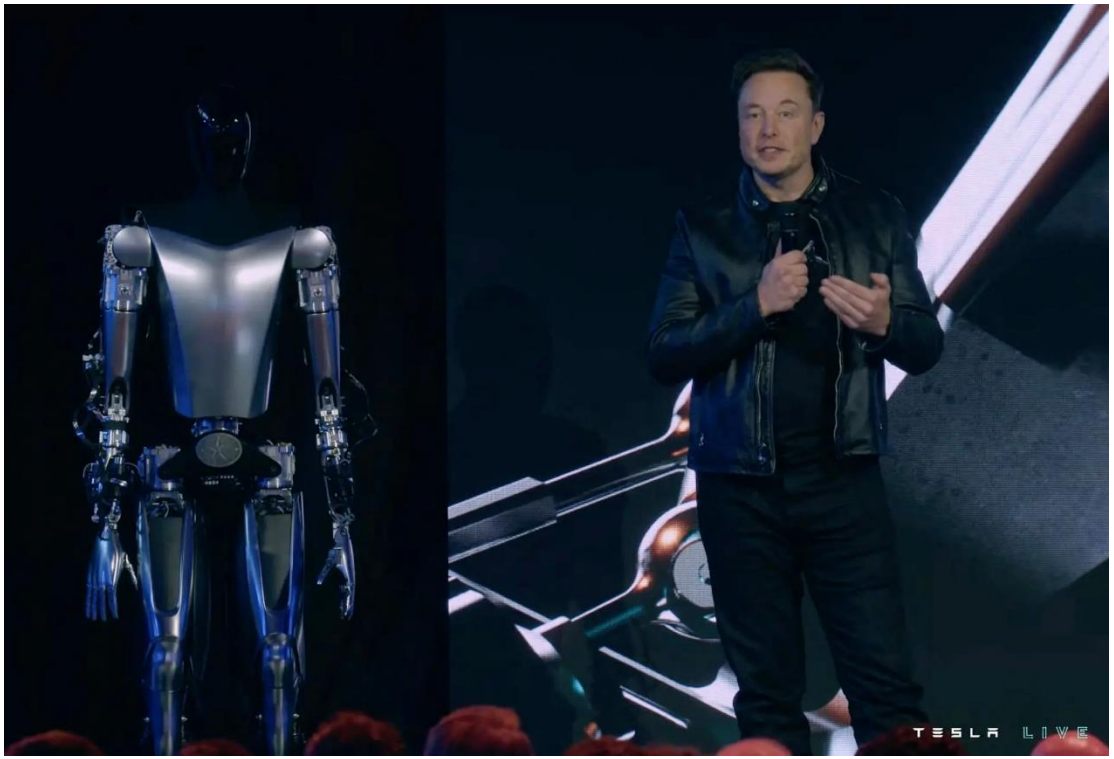
- **Θεμελιώδη μοντέλα (foundation models)** για την παραγωγή **εικόνων, ήχου, βίντεο, κώδικα κλπ.**
- Τα συστήματα αυτά επαυξάνουν σημαντικά την ανθρώπινη δημιουργικότητα οπότε έχουμε φτάσει στην εποχή που οι άνθρωποι δημιουργούν **μαζί** με αυτά τα μοντέλα.

Πρακτορική Τεχνητή Νοημοσύνη (Agentic AI)

- Κατά τη μετάβαση από ένα LLM σε ένα agentic system:
  - **Από την παροχή απαντήσεων στην ανάληψη δράσης** — ένας πράκτορας υπερβαίνει την απλή παραγωγή απαντήσεων: μπορεί να προβαίνει σε **ενέργειες (actions)** για την επίτευξη ενός **στόχου (goal)**.
  - **Σχεδιασμός (planning) και εκτέλεση (execution) εργασιών** — Μπορεί να αναλύσει έναν στόχο σε επιμέρους βήματα, να αποφασίσει για την επόμενη κίνηση και να προσαρμοστεί ανάλογα με τα αποτελέσματα.
  - **Χρήση εργαλείων και αλληλεπίδραση με το περιβάλλον** — Οι πράκτορες μπορούν να αλληλεπιδρούν με API, μηχανές αναζήτησης, βάσεις δεδομένων, λογισμικό ή άλλα συστήματα τεχνητής νοημοσύνης.
  - **Λειτουργία με βαθμό αυτονομίας (autonomy)** — Αντί να απαιτείται εντολή (prompt) για κάθε βήμα, το σύστημα μπορεί να εκτελεί ροές εργασίας πολλαπλών σταδίων, τηρώντας παράλληλα τους καθορισμένους περιορισμούς.
- **Επιγραμματικά:** ένα LLM παράγει περιεχόμενο, ενώ ένας πράκτορας χρησιμοποιεί ένα LLM για να λαμβάνει αποφάσεις, να ενεργεί, να παρατηρεί και να επαναλαμβάνει τη διαδικασία.

Σύγχρονες Τάσεις

- Ανθρωποειδή ρομπότ π.χ., ο Optimus της Tesla.



Σύγχρονες Τάσεις

- **Ρομποτική**
  - Το RoboCup (http://www.robocup.org/) είναι ένας διεθνής διαγωνισμός ρομποτικού ποδοσφαίρου με τον παρακάτω στόχο:
    - By the year 2050, develop a team of fully autonomous humanoid robots that can win against the human world soccer champion team.



Σύγχρονες Τάσεις

- **Αυτοοδηγούμενα οχήματα (self-driving cars)**
  - Η μεγάλη πρόκληση της DARPA's για την κατασκευή αυτόνομων οχημάτων το 2004.
  - Η εμπορική παραγωγή αυτοοδηγούμενων οχημάτων από εταιρίες όπως η Tesla σήμερα (2017 και μετά).





Σύγχρονες Τάσεις

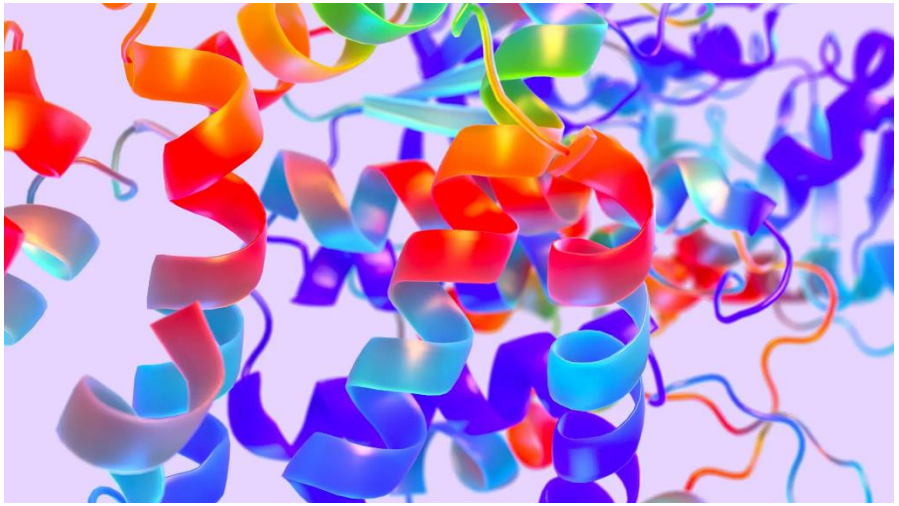
- Οι εταιρίες **Waymo** (πρώην self-driving car project of Google) και άλλες διαθέτουν σήμερα **αυτοοδηγούμενα ταξί** σε πόλεις των ΗΠΑ.



Σύγχρονες Τάσεις

- Χρήση βαθιάς μηχανικής μάθησης για σημαντικά ανοικτά επιστημονικά προβλήματα π.χ., την **πρόβλεψη της δομής των πρωτεϊνών (AlphaFold από την DeepMind).**

Google



Σύγχρονες Εφαρμογές Βασισμένες σε ΤΝ

- **Η τεχνολογία για την οποία θα διαβάσετε αύριο στις εφημερίδες!**



Προκλήσεις

- Προστασία των προσωπικών δεδομένων και ιδιωτικότητα
- Προστασία της πνευματικής ιδιοκτησίας
- Ανάπτυξη έμπιστων συστημάτων Τεχνητής Νοημοσύνης
  - Ασφάλεια
  - Ευθυγράμμιση
  - Δικαιοσύνη και προκαταλήψεις
  - Ερμηνευσιμότητα/Επεξηγησιμότητα
- Περιβαλλοντικό κόστος
- Το μέλλον της εργασίας
- Τεχνητή Νοημοσύνη και εθνική ασφάλεια
  - Κυβερνοασφάλεια
  - Αυτόνομα οπλικά συστήματα
- Τεχνητή Νοημοσύνη και δημόσιο συμφέρον
- Υπαρξιακοί κίνδυνοι από την Τεχνητή Νοημοσύνη
- Ρυθμιστικά πλαίσια

Φιλοσοφικές Ερωτήσεις

- Πόσο κοντά είμαστε στο να φτιάξουμε συστήματα Τεχνητής Νοημοσύνης που να μπορούν να κάνουν ό,τι και ένας άνθρωπος **(human level AI, artificial general intelligence – AGI**);
- Πόσο κοντά είμαστε στο να φτιάξουμε συστήματα Τεχνητής Νοημοσύνης που να ξεπερνούν τον άνθρωπο σε ευφυία και ίσως τον αντικαταστήσουν στον πλανήτη μας (artificial superintelligence);
- Θέλουμε να φτιάξουμε τέτοια συστήματα;

AI@DI

- Δείτε την ιστοσελίδα της ομάδας μου **Artificial Intelligence Team (AITeam).**

ΕΕΤΝ

- Γίνετε μέλη της Ελληνικής Εταιρίας Τεχνητής Νοημοσύνης.



ECAI 2027

- Ελάτε του χρόνου τον Οκτώβριο στο συνέδριο European Conference on AI (ECAI).