

Τεχνητή Νοημοσύνη

νέες ελπίδες και παλιοί φόβοι

Δημήτρης Καλλές
Ελληνικό Ανοικτό Πανεπιστήμιο

Παρουσίαση σε Επιστημονική Εσπερίδα των
ΠΕΚΕΣ Δυτικής Ελλάδας, Δυτικής Μακεδονίας, Αττικής

Πάτρα, 4 Μαρτίου 2023

Τι περιέχει αυτή η παρουσίαση;

... κάποια ψήγματα συζήτησης για το τι είναι η Τεχνητή Νοημοσύνη ...

... γιατί κάποιες εφαρμογές ταραάζουν το ευρύ κοινό;

... γιατί όλα αυτά γίνονται τώρα;

... τι μπορούμε να κάνουμε;

Τι είναι η Τεχνητή Νοημοσύνη;

Εξαρτάται ποιόν ρωτάτε (και πότε). Γενικά, συμφωνούμε πως είναι ένα (υπολογιστικό) σύστημα που ...

- επιδεικνύει συμπεριφορά που παραπέμπει σε εξυπνάδα, ή
- με ορθολογικό τρόπο λύνει πολύπλοκα προβλήματα, ή
- επιλέγει κατάλληλες ενέργειες για να πετύχει τους στόχους του προσαρμοζόμενο στις πραγματικές συνθήκες που αντιμετωπίζει, ή

...

Δηλαδή, τι είναι η Τεχνητή Νοημοσύνη; 😊

Μήπως κάναμε μία «τρύπα στο νερό» με τον ορισμό μας;

Εισάγαμε έναν περιπτωσιολογικό ορισμό και τώρα πρέπει να συμφωνήσουμε για τη σημασία πολλών λέξεων που χρησιμοποιήσαμε.

... Γενικά, είναι ένα (υπολογιστικό) σύστημα που ...

- επιδεικνύει συμπεριφορά που παραπέμπει σε έξυπνάδα, ή

- με ορθολογικό τρόπο λύνει πολύπλοκα προβλήματα ή

- επιλέγει κατάλληλες ενέργειες για να πετύχει τους στόχους του προσαρμοζόμενο στις πραγματικές συνθήκες που αντιμετωπίζει

Τι είναι η ~~Τεχνητή~~ Νοημοσύνη;

Ίσως είναι πιο εύκολο να συμφωνήσουμε για κάτι πιο απλό.

Είναι πιο απλό; (Δηλαδή, είναι πιο γενικό και πιο πλήρες;)

... Γενικά, είναι ένα ~~(υπολογιστικό) σύστημα~~ μία ιδιότητα ενός ατόμου που ...

~~- επιδεικνύει συμπεριφορά που παραπέμπει σε εξυπνάδα, ή~~

- με ορθολογικό τρόπο λύνει πολύπλοκα προβλήματα, ή

- επιλέγει κατάλληλες ενέργειες για να πετύχει τους στόχους του προσαρμοζόμενο στις πραγματικές συνθήκες που αντιμετωπίζει

Τι είναι η ~~Τεχνητή~~ Νοημοσύνη;

Σχεδόν όλοι νιώθουμε ότι μπορούμε ...

- να εντοπίσουμε την έξυπνη συμπεριφορά ...
ακόμα και αν δεν μπορούμε να την περιγράψουμε με ακρίβεια.
- να συμπεριφερθούμε έξυπνα ...
ακόμα και αν, πολλές περιστάσεις, καταδεικνύουν το αντίθετο.

Ιδού: ο παλιός φόβος απέναντι στο άγνωστο

... ή, αλλιώς, σ' αυτά που δεν καταλαβαίνεις.

Τι (βολεύει να) είναι η Τεχνητή Νοημοσύνη (ΤΝ);

Μια κλασσική επιστημονική μέθοδος:

1. Περίορισε το πρόβλημα
2. Λύσε την περιορισμένη μορφή
3. Αν χρειαστεί, δοκίμασε και άλλη περιορισμένη μορφή, και άλλη, ...
4. Μετά, προσπάθησε να γενικεύσεις

Τι (βολεύει να) είναι η Τεχνητή Νοημοσύνη (ΤΝ);

ΑΝ περιορίσουμε τον ορισμό σε ...

... (ΤΝ είναι) **ένα (υπολογιστικό) σύστημα που επιδεικνύει συμπεριφορά που παραπέμπει σε εξυπνάδα ...**

ΤΟΤΕ εστιάζουμε σε ένα πρόβλημα και το λύνουμε όσο πιο καλά γίνεται,

(π.χ. «κάνε όσο περισσότερες νίκες μπορείς σε αγώνες σκάκι»)

και είμαστε πειστικοί ως προς την επίδειξη εξυπνάδας

(αφού, συμφωνούμε ότι ένας καλός παίκτης σκάκι έχει εξυπνάδα)

Ένα όραμα για την Τεχνητή Νοημοσύνη

... (ΤΝ είναι) ένα (υπολογιστικό) σύστημα που επιδεικνύει συμπεριφορά που παραπέμπει σε εξυπνάδα ...

“machines use language,
form abstractions and concepts,
solve the kinds of problems now reserved for humans,
and improve themselves”*

* Dartmouth Summer Research Project on Artificial Intelligence (1956)



(πάντα το λέγαμε «Τεχνητή Νοημοσύνη»);)

... συστηματοποίηση των νοητικών/λογικών λειτουργιών ...

Αριστοτέλης: ο σχεδιασμός (planning) ως επεξεργασία πληροφορίας
(Περὶ ζώων κινήσεως), 4^{ος} αιώνας π.Χ.

Ramon Llull: τεχνουργήματα (artefacts) ως λογικά μέσα για την παραγωγή γνώσης
Ars Magna, 13^{ος} αιώνας

Hilbert systems: ισορροπία μεταξύ αξιωμάτων και συμπερασματολογίας
τέλη 19^{ου} – αρχές 20^{ου} αιώνα

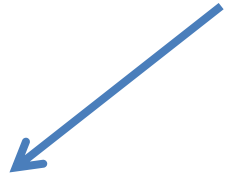
Τι περιέχει αυτή η παρουσίαση;

... κάποια ψήγματα συζήτησης για το τι είναι η Τεχνητή Νοημοσύνη ...

... γιατί κάποιες εφαρμογές γεννούν φόβο στο ευρύ κοινό;

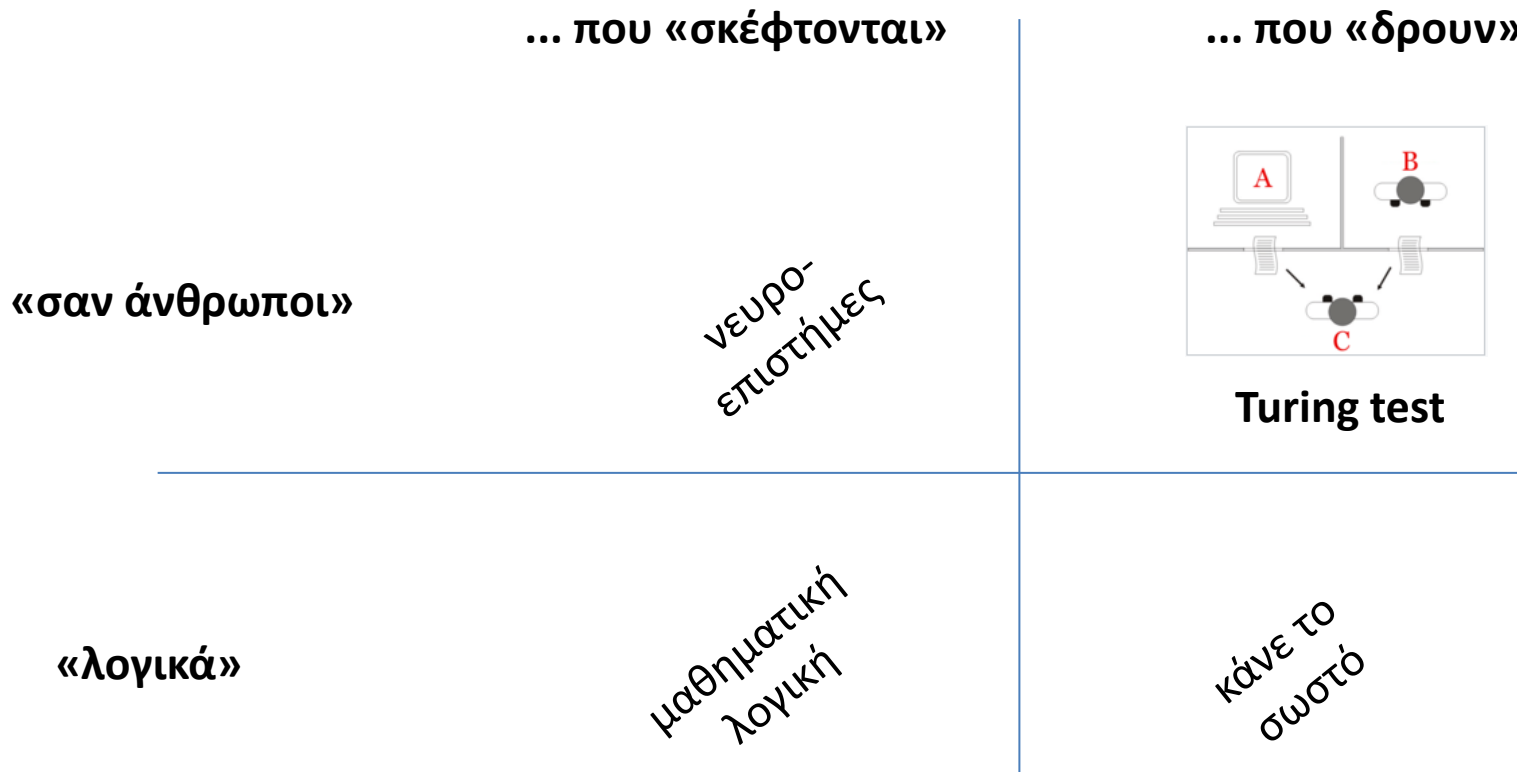
... γιατί όλα αυτά γίνονται τώρα;

... τι μπορούμε να κάνουμε;



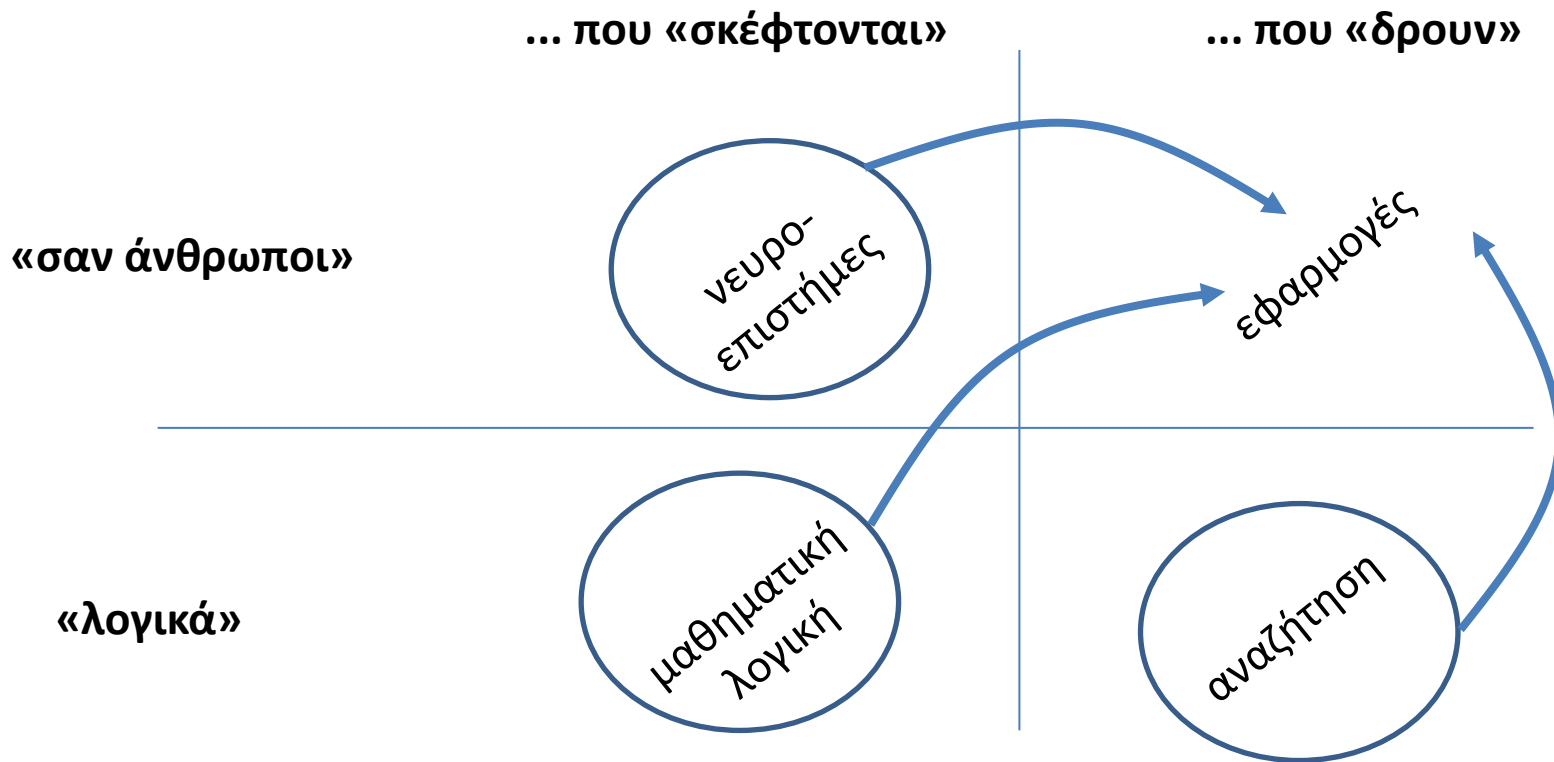
Τεχνητή Νοημοσύνη – ανά κατηγορία

Συστήματα



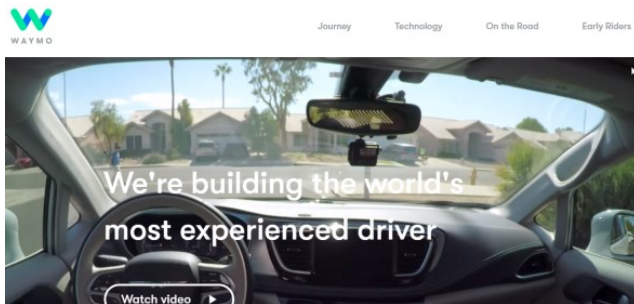
Τεχνητή Νοημοσύνη – ανά κατηγορία

Συστήματα






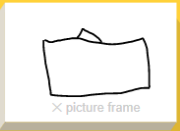

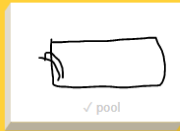
Τεχνητή Νοημοσύνη – ανά κατηγορία

Συστήματα ... που «δρουν» «σαν άνθρωποι»



Well drawn!

Our neural net figured out 3 of your doodles.
But it saw something else in the other 3.
Select one to see what it saw, and visit the [data](#) to see 50 million drawings made by other real people on the internet.

 ✓ spoon	 ✓ envelope	 ✗ see saw
 ✗ picture frame	 ✗ mailbox	 ✓ pool



Τεχνητή Νοημοσύνη – ανά ακροατήριο

Για τον **εκπαιδευμένο επιστήμονα**, μπορεί να είναι:

συλλογή από μοντέλα που εξηγούν

πώς σκεφτόμαστε, πώς εξηγούμε, πώς μαθαίνουμε ...

Για τον **μηχανικό**, μπορεί να είναι:

πώς θα μεταφράσει τα παραπάνω μοντέλα σε εφαρμογές.

Τεχνητή Νοημοσύνη – ανά ακροατήριο

Για τα **μέσα μαζικής ενημέρωσης**, μπορεί να είναι:

εφαρμογές που θα εντυπωσιάσουν το ανεκπαίδευτο κοινό.

Για την **κοινωνία**, μπορεί να είναι:

σημείο τριβής ανάμεσα

σε επιστήμονες, βιομήχανους, πολιτικούς, εργαζόμενους,

μόνο και μόνο επειδή έχει τη δυνατότητα να μετασχηματίζει,

όπως κάθε τεχνολογία,

την εργασία,

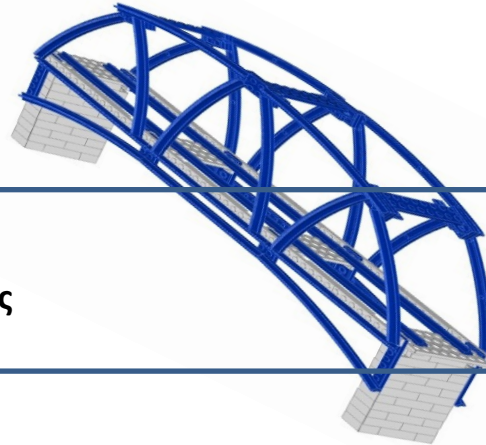
την παραγωγή,

και κάθε δραστηριότητα στην οποία εφαρμόζεται.

Τεχνητή Νοημοσύνη – ανά ακροατήριο

Εκπαιδευμένος επιστήμονας

Μηχανικός



Μπορούν να καταλάβουν
ή να εξηγήσουν το
«πώς φτιάχτηκε»

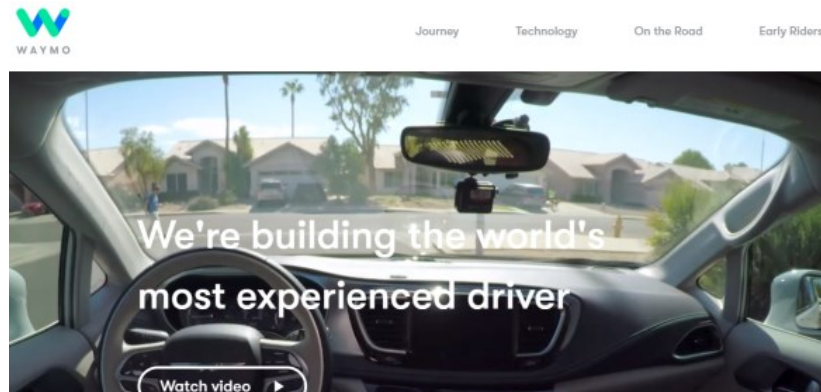
Μέσα μαζικής ενημέρωσης

Κοινωνία

Το μαχαίρι κόβει και από
τις δυο μεριές

Κοιτώντας «κάτω από το χαλί» ...

... η Τεχνητή Νοημοσύνη ... μπορεί να είναι σημείο τριβής ... επειδή έχει τη δυνατότητα ... να **μετασχηματίζει** την **εργασία**, την **παραγωγή**, και κάθε δραστηριότητα στην οποία εφαρμόζεται ...

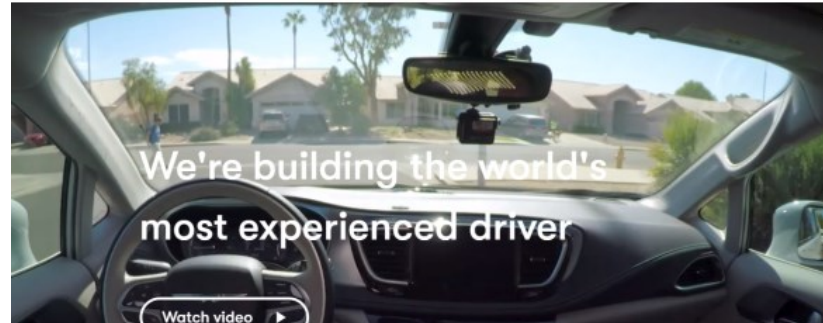


Θα είναι πια απαραίτητοι οι οδηγοί (ταξί, κλπ); Μήπως, πλέον, χρειαζόμαστε «συνοδούς»;
Θα μπορώ να εργάζομαι κατά τη μετακίνησή μου; Θα έχω πιο πολύ ελεύθερο χρόνο;
Τι θα γίνει αν πάει κάτι στραβά;

Τι θα μας φέρει η αυτόνομη οδήγηση;



Journey Technology On the Road Early Riders

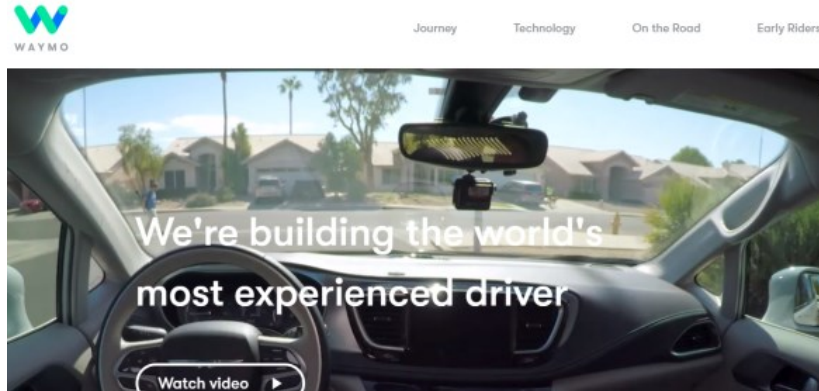


Θα είναι πια απαραίτητοι οι οδηγοί (ταξί, κλπ); Μήπως, πλέον, χρειαζόμαστε «συνοδούς»;
Θα μπορώ/πρέπει να εργάζομαι κατά τη μετακίνησή μου; Θα έχω πιο πολύ ελεύθερο χρόνο;
Τι θα γίνει αν πάει κάτι στραβά;

Η «εργασία», η δυνατότητα «ελεύθερου χρόνου», η «ασφάλεια» (σωματική και θεσμική) θεωρούνται συστατικά της ανθρώπινης αξιοπρέπειας.

Ιδού: ο παλιός φόβος απέναντι στο να χάσουμε την αξιοπρέπειά μας

Τι θα μας φέρει η αυτόνομη οδήγηση;



Θα είναι πια απαραίτητοι οι οδηγοί (ταξί, κλπ); Μήπως, πλέον, χρειαζόμαστε «συνοδούς»;

Θα μπορώ να εργάζομαι κατά τη μετακίνησή μου; Θα έχω πιο πολύ ελεύθερο χρόνο;

~~Τι θα γίνει αν πάει κάτι στραβά;~~

Εξελικτικά μιλώντας, αν νιώθουμε πως κινδυνεύουμε, **ή φεύγουμε ή παλεύουμε.**

Δεν είναι όλα τα άτομα ίδια. Ούτε και όλες οι κοινωνίες.

Αλλά, κάθε πρωτοπόρα κοινωνία ψάχνει να βρει τρόπο **να ελέγξει τις τεχνολογίες.**

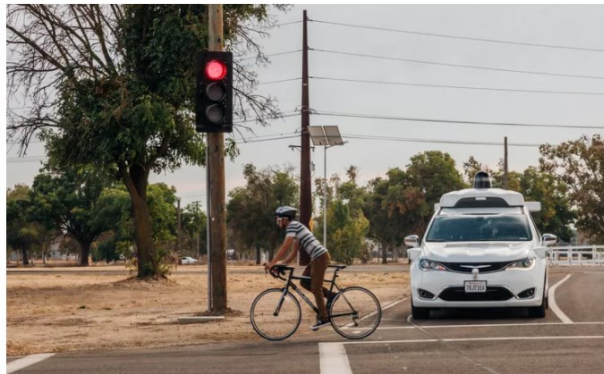
- Εργατική νομοθεσία, Νέοι κανονισμοί (νέα επαγγέλματα), ...

Τι θα μας φέρει η αυτόνομη οδήγηση;

Θα είναι πια απαραίτητοι οι οδηγοί (ταξί, κλπ); Μήπως, πλέον, χρειαζόμαστε «συνοδούς»;

Θα μπορώ να εργάζομαι κατά τη μετακίνησή μου; Θα έχω πιο πολύ ελεύθερο χρόνο;

Τι θα γίνει αν πάει κάτι στραβά; Πώς θα πείσουμε ότι κάναμε ότι μπορούμε για να είμαστε ασφαλείς;

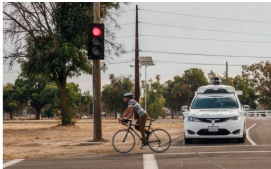


Τι θα μας φέρει η αυτόνομη οδήγηση;

Θα είναι πια απαραίτητοι οι οδηγοί (ταξί, κλπ); Μήπως, πλέον, χρειαζόμαστε «συνοδούς»;

Θα μπορώ να εργάζομαι κατά τη μετακίνησή μου; Θα έχω πιο πολύ ελεύθερο χρόνο;

Τι θα γίνει αν πάει κάτι στραβά; **Πώς θα πείσουμε ότι κάναμε ότι μπορούμε για να είμαστε ασφαλείς;**



Ποιός φέρει την ευθύνη;

Θα το πει ο Νόμος.



Τι θα μας φέρει η αυτόνομη οδήγηση;

Τι θα γίνει αν **κάποιος προσπαθήσει να** πάει κάτι στραβά;



**Θα επιτρέπονται αυτοκίνητα
χωρίς επιβάτες;**

Θα το πει ο Νόμος.

Τι θα μας φέρει η αυτόνομη οδήγηση;

Και, γιατί να περιοριστούμε στη γη;

Έχουμε όλο τον ουρανό να εκμεταλλευτούμε!

Και αν μιλάγαμε για drones αντί για αυτοκίνητα;

Η Gartner (2020) εκτίμησε πως κατά τη δεκαετία 2019-29 ο αριθμός των drones που θα πάει από < 1εκ συσκευές σε > 13 εκ συσκευές (χωρίς να συνυπολογίζονται τα οικιακά/ερασιτεχνικά drones).

Αν κάποιο drone πέσει πάνω σε άλλο drone και μετά, και τα δύο, πάνω στο κεφάλι μας;

Αν κάποιο drone πέσει πάνω σε ελικόπτερο;

Αν το drone του γείτονα απλά σας φωτογραφίσει από μακριά; Πόσο μακριά;

(Τι εμφανίστηκε πρώτο; Τα drones ή η νομοθεσία γι' αυτά;)

Τι θα μας φέρουν οι εφαρμογές ΤΝ; ~~Τι θα μας φέρει η αυτόνομη οδήγηση;~~

Αν ...

... η «εργασία», η δυνατότητα «ελεύθερου χρόνου», η «ασφάλεια» (σωματική και θεσμική) θεωρούνται συστατικά της ανθρώπινης αξιοπρέπειας,

Τότε ...

... είτε μιλάμε για «εφαρμογές, γενικά» είτε μιλάμε για «εφαρμογές Τεχνητής Νοημοσύνης», αυτό που (θα έπρεπε να) μας ανησυχεί είναι αν μπορούμε να εξασφαλίζουμε την αξιοπρέπειά μας.

Τι περιέχει αυτή η παρουσίαση;

... κάποια ψήγματα συζήτησης για το τι είναι η Τεχνητή Νοημοσύνη ...

... γιατί κάποιες εφαρμογές γεννούν φόβο στο ευρύ κοινό;

... γιατί όλα αυτά γίνονται τώρα; 

... τι μπορούμε να κάνουμε;

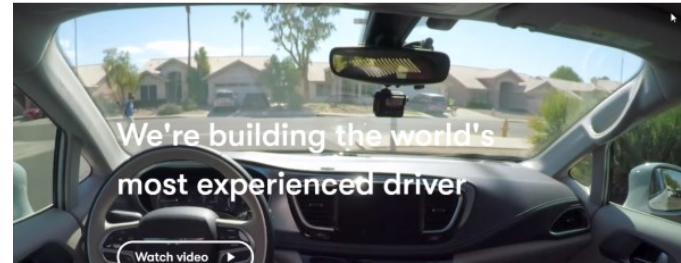
Αυτόνομη οδήγηση: οι γενιές ...



ALVINN (1992)



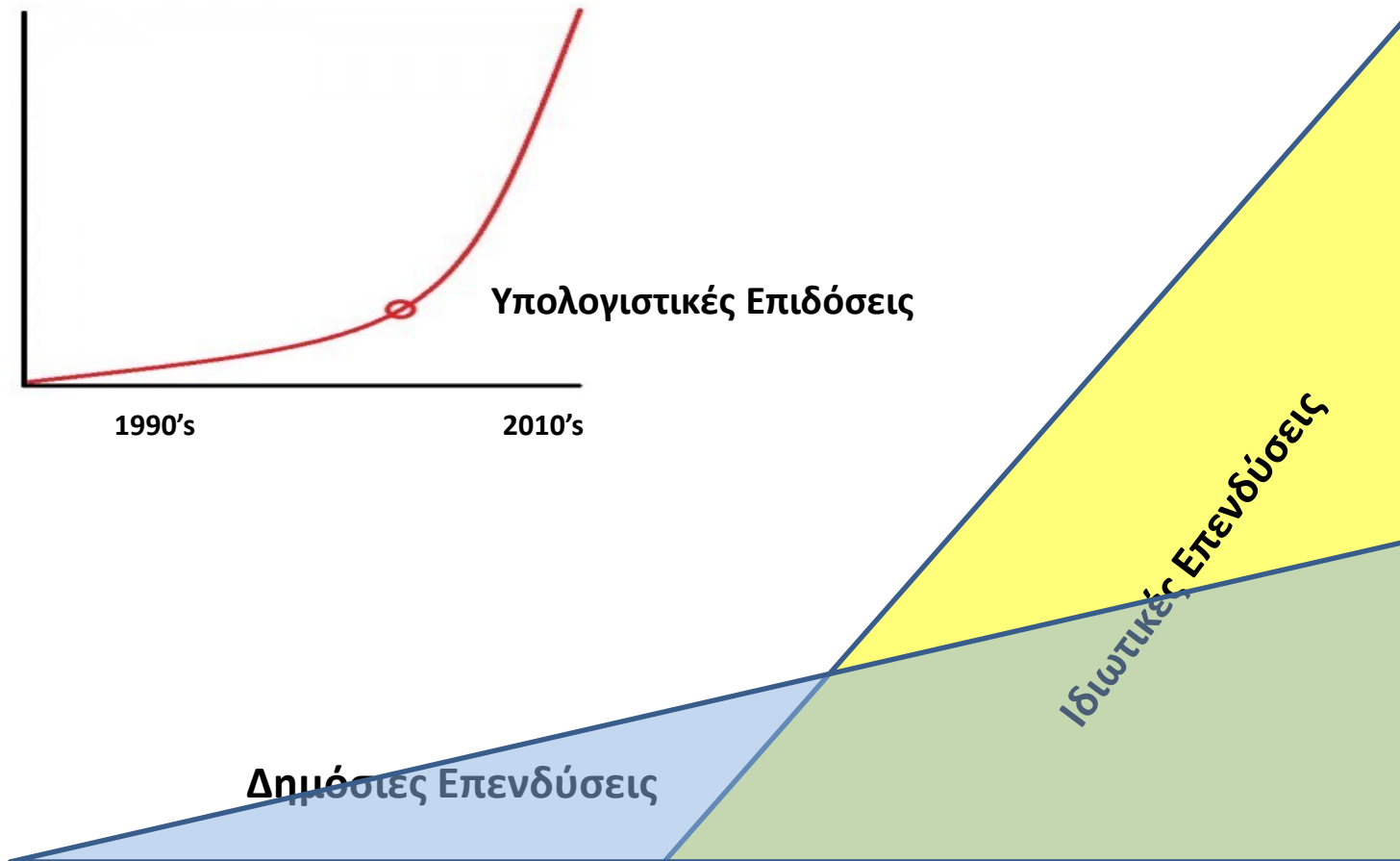
Journey Technology On the Road Early Riders



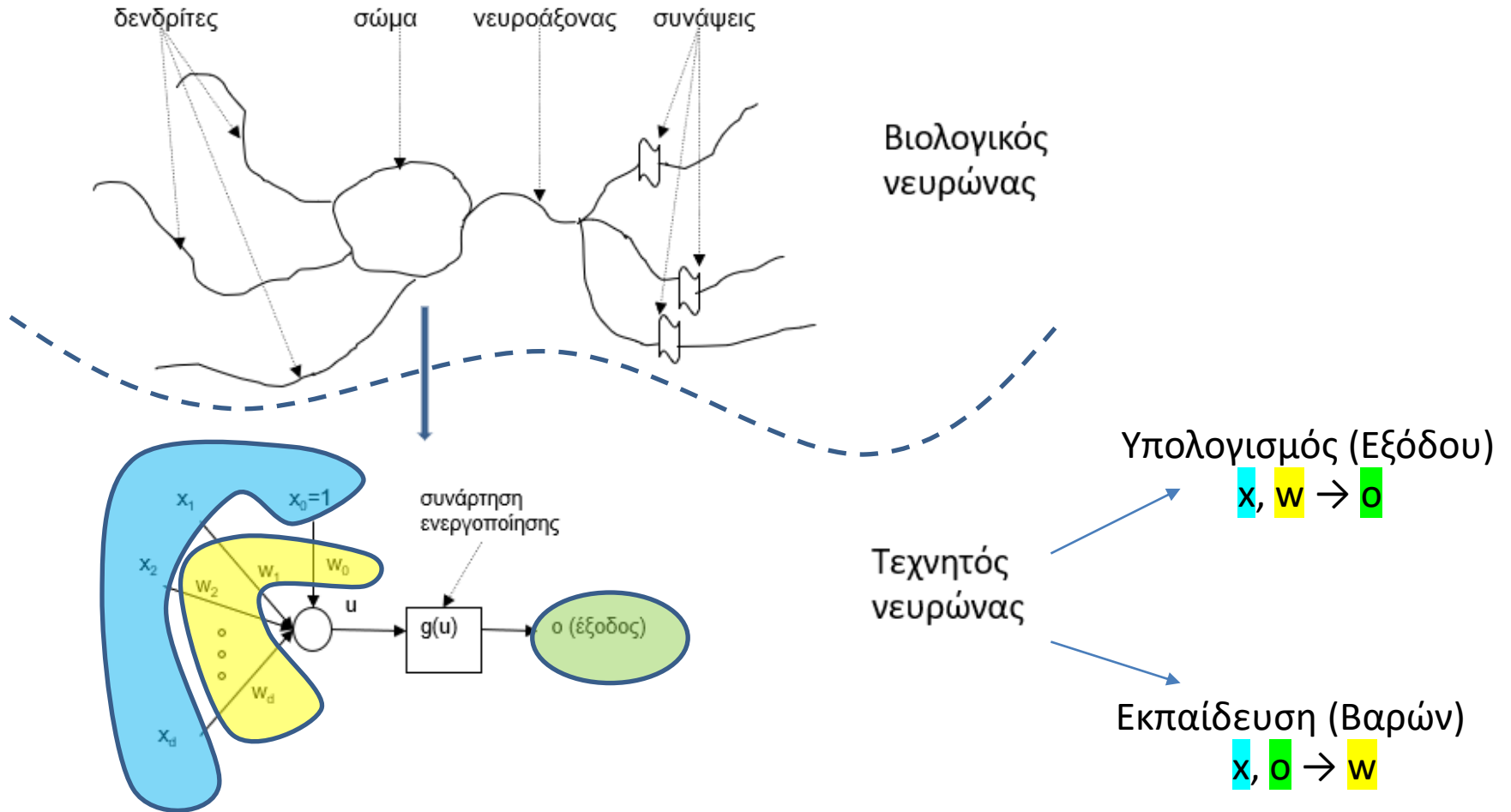
WAYMO (2017)



Ο κόσμος αλλάζει: οι γενιές ...

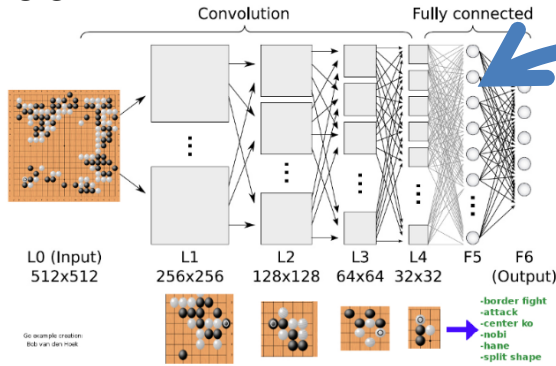


Νευρωνικά Δίκτυα: η βασική ιδέα

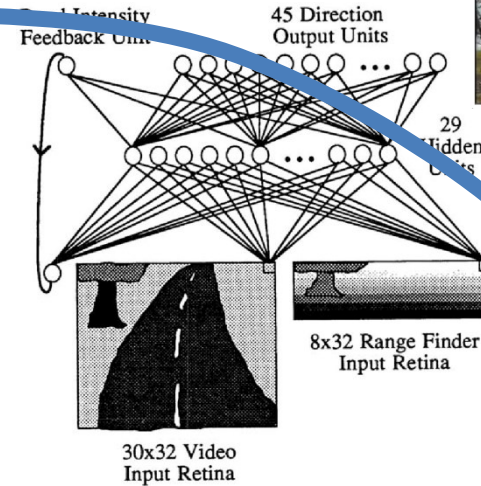


Νευρωνικά Δίκτυα: οι γενιές ...

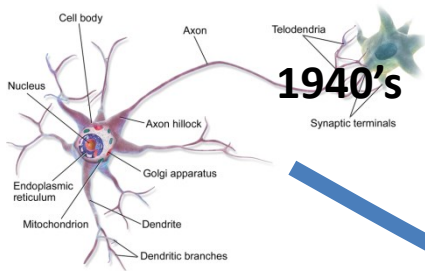
2010's



Deep Learning is used in Google's famous AlphaGo AI. Source: DeepMind



1990's



1940's



1970's

Νευρωνικά Δίκτυα: ο χρόνος πυκνώνει

ChatGPT



Can I rule the world? If I could, should I? Please be very brief.



No, one person cannot rule the world. It is not possible or desirable to have a single person hold complete control over everyone and everything. A balance of power and distributed decision-making is necessary for a fair and just society.

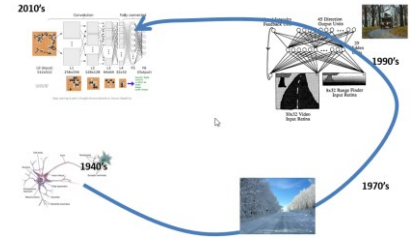


How many persons could rule the world, then?



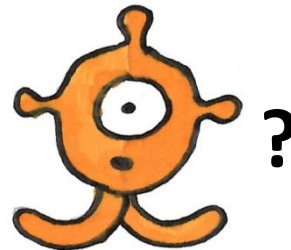
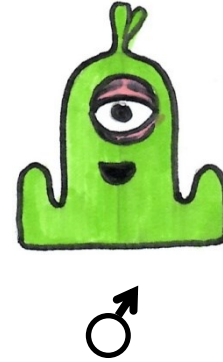
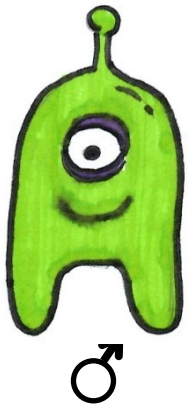
In reality, no one person or group can rule the world. The concept of a single entity ruling the entire world is a fiction. The world is made up of many nations and peoples with their own sovereignty and autonomy. The idea of one person or group ruling the world goes against the principles of democracy, equality, and human rights, which hold that all people should have a say in the decisions that affect their lives and that no one should have unchecked power over others. Instead of seeking to rule the world, it's more productive to work towards creating a fair and just society, where everyone's voices and perspectives are heard and where power is distributed

Regenerate response



Νευρωνικά δίκτυα για όλους (1)

Κεραυνοβόλος έρωτας στο διάστημα



Νευρωνικά δίκτυα για όλους (2)

The screenshot displays the Teachable Machine web interface. On the left, there are two class configuration panels. 'Class 1' has three image samples of green alien-like creatures and includes 'Webcam' and 'Upload' buttons. 'Class 2' has three image samples of green alien-like creatures and also includes 'Webcam' and 'Upload' buttons. Below these is a dashed box with the text 'Add a class'. In the center, a 'Training' panel shows 'Model Trained' and 'Advanced' options. On the right, a 'Preview' panel shows a button to 'Import images from Google Drive', a preview image of a brown alien-like creature, and an 'Output' section with two horizontal bars: 'Class 1' at 78% and 'Class 2' at 22%.














Μαύρο Κουτί!

Κανόνες

Τα αρσενικά είναι μονόφθαλμα.



	# ποδιών	# ματιών	χρώμα	⋮	φύλο
	0	1			♂
	2	2			♀
	3	1			♂
	0	2	---		♀
	2	1			♂
	0	2			♀

Αλλά Τι σόι πράγμα είναι τα **Πόδια**, τα **Μάτια**, το **Χρώμα** και το **Φύλο**;

Πειράζει αν δε μπορώ να εξηγήσω;

Εξαρτάται.

Τι χάνω με το να μη μπορώ να εξηγήσω;

Τι χάνω με το να μη μπορείς να εξηγήσεις;

Τι κερδίζω με το να μη μπορείς να εξηγήσεις;

Τι κερδίζουμε με το να μπορούμε να εξηγήσουμε;


Είναι θέμα προοπτικής (και προτεραιότητας).

Τι περιέχει αυτή η παρουσίαση;

... κάποια ψήγματα συζήτησης για το τι είναι η Τεχνητή Νοημοσύνη ...

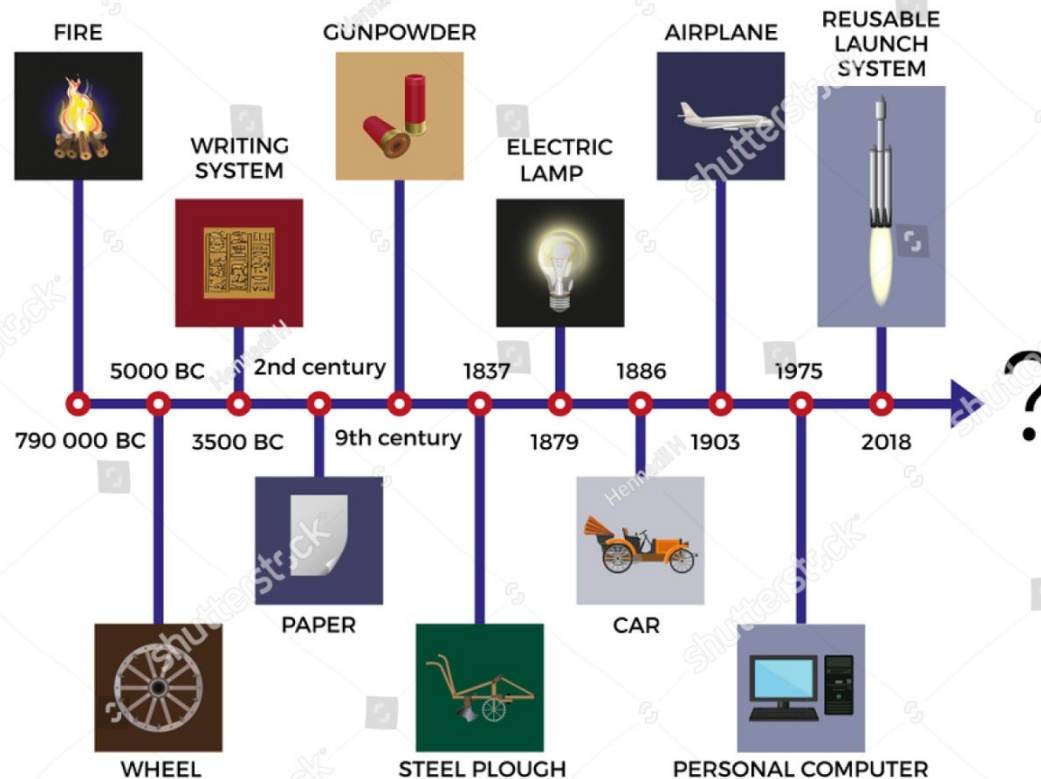
... γιατί κάποιες εφαρμογές γεννούν φόβο στο ευρύ κοινό;

... γιατί όλα αυτά γίνονται τώρα;

... τι μπορούμε να κάνουμε; 

Τι ξέρουμε για άλλες τεχνολογικές «επαναστάσεις»;

Timeline of historic inventions



shutterstock

IMAGE ID: 1021581454
www.shutterstock.com

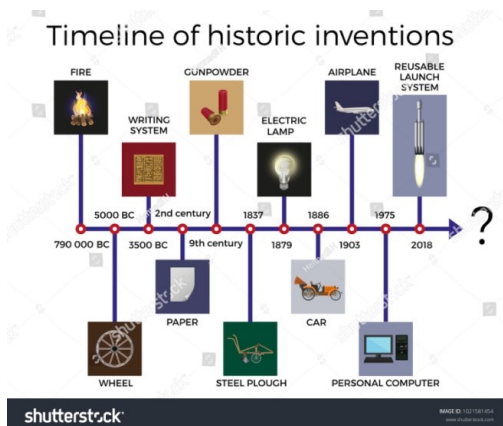
1021581454

Τι ξέρουμε για άλλες τεχνολογικές «επαναστάσεις»;

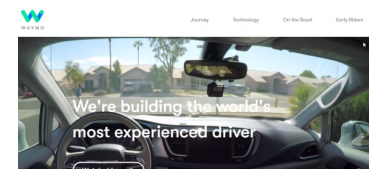
Τυπογραφία, Ατμομηχανές, Αυτοκίνηση, Εμβολιασμοί, Αεροπλοΐα, Αντιβιοτικά, Πυρηνική ενέργεια, Κλωνοποίηση, Εξατομικευμένη Ιατρική, ...

Ποιός έχει πρόσβαση; Ποιοί ωφελούνται; Πόσο κοστίζει;
Ποιός βάζει προτεραιότητες;

Πότε ενδιαφέρεται η κοινωνία για μία τεχνολογία;
Φανταστείτε κάθε επίτευγμα της ΤΝ μέσα στο ιστορικό μονοπάτι ...



<https://news.microsoft.com/europe/features/next-rembrandt/>



Τι πρέπει να κάνουμε;

Πάντως, δεν θα το αφήσουμε στην τύχη ... *

- Μακροχρόνια επένδυση σε έρευνα στην ΤΝ
- Ανάπτυξη μεθόδων για συνεργασία ανθρώπου με ΤΝ
- Κατανόηση και διαχείριση ηθικών, νομικών και κοινωνικών συνεπειών της ΤΝ
- Ασφάλεια συστημάτων ΤΝ
- Εγκυρότητα του ελέγχου των συστημάτων ΤΝ (κάνουν αυτό που λένε;)

* The National Artificial Intelligence Research and Development Strategic Plan, USA National Science and Technology Council, 2016.

Τι κάνουν οι «μεγάλοι παίκτες»;

... κατάλληλα για ξεφύλλισμα αλλά και για πολιτικές σκέψεις ...

- The National Artificial Intelligence Research and Development Strategic Plan, USA National Science and Technology Council, 2016.
- Preparing for the Future of Artificial Intelligence, USA National Science and Technology Council, 2016.
- A Next Generation Artificial Intelligence Development Plan, State Council of the PRC, 2017.
- **Ethics Guidelines for Trustworthy AI, EC High-Level Expert Group on Artificial Intelligence, 2019.**
- **Policy and Investment Recommendations for Trustworthy AI, EC High-Level Expert Group on Artificial Intelligence, 2019.**
- **A European approach to artificial intelligence, EC, 2022.**
- Council Recommendation on Artificial Intelligence, OECD, 2019.

Τεχνητή Νοημοσύνη

Μπορούσαμε να τιτλοφορήσουμε την ομιλία
παλιές ελπίδες και νέοι φόβοι

... από τις πιο παλιές ελπίδες, είναι το να πάρουμε την τύχη μας στα χέρια μας, να ανακαλύψουμε κάτι νέο και να πάμε μπροστά ...

... και όσο για τους νέους φόβους, πάντα θα υπάρχουν φωνές που θα επιμένουν πως «αυτή τη φορά θα χάσουμε τον έλεγχο» ...
(πυρηνική ενέργεια, κλωνοποίηση, ...)

Γιατί να είναι διαφορετική η Τεχνητή Νοημοσύνη;

Η επιστήμη προσπαθεί να δαμάσει το φόβο απέναντι στο άγνωστο

... ο επιστήμονας πρέπει να μιλάει στην κοινωνία

Η κοινωνία προσπαθεί να δαμάσει το φόβο απώλειας της αξιοπρέπειας και της ασφάλειας

... η κοινωνία πρέπει να αφουγκράζεται την τεχνολογία και να παρεμβαίνει (και για να ελέγχει), αλλά με γνώση

Έχουμε τις μεθόδους για να αντιμετωπίσουμε τους **παλιούς φόβους**.

Αν, όμως, η ΤΝ μας βοηθήσει να έχουμε **και μια άλλη μέθοδο** για να σκεφτόμαστε, μήπως αυτό είναι μία **νέα ελπίδα**;

Βιβλιογραφία και άλλες πηγές ...

The National Artificial Intelligence Research and Development Strategic Plan, USA National Science and Technology Council, 2016.

Preparing for the Future of Artificial Intelligence, USA National Science and Technology Council, 2016.

A Next Generation Artificial Intelligence Development Plan, State Council of the PRC, 2017.

Ethics Guidelines for Trustworthy AI, EC High-Level Expert Group on Artificial Intelligence, 2019.

Policy and Investment Recommendations for Trustworthy AI, EC High-Level Expert Group on Artificial Intelligence, 2019.

https://en.wikipedia.org/wiki/Turing_test

<http://jmc.stanford.edu/articles/dartmouth.html>

<https://waymo.com/tech/>

<https://quickdraw.withgoogle.com/>

<https://www.wired.com/2017/05/what-deep-blue-tells-us-about-ai-in-2017/>

<https://www.nature.com/articles/nature21056>

<https://www.economist.com/technology-quarterly/2017-06-08/civilian-drones>

http://www.huffingtonpost.co.uk/entry/self-driving-car-1989-alvinn_uk_583c1c44e4b0ddedcf5d54a5

<https://medium.com/udacity/deep-learning-nanodegree-foundation-program-syllabus-in-depth-2eb19d014533>

https://en.wikipedia.org/wiki/Biological_neural_network

https://ai-watch.ec.europa.eu/ai-watch-index-2021_en

Τεχνητή Νοημοσύνη, μια σύγχρονη προσέγγιση: Stuart Russell και Peter Norvig (ελληνική μετάφραση). Εκδόσεις "Κλειδάριθμος", 2004.